

UNIVERSIDADE DE SÃO PAULO

Instituto de Ciências Matemáticas e de Computação

**A Inteligência Artificial na automação de vistorias
imobiliárias: aplicação de redes neurais
convolucionais (CNNs) para detecção de
irregularidades em imagens**

Luis Roberto da Silva

Monografia - MBA em Inteligência Artificial e Big Data

SERVIÇO DE PÓS-GRADUAÇÃO DO ICMC-USP

Data de Depósito:

Assinatura: _____

Luis Roberto da Silva

A Inteligência Artificial na automação de vistorias imobiliárias: aplicação de redes neurais convolucionais (CNNs) para detecção de irregularidades em imagens

Monografia apresentada ao Departamento de Ciências de Computação do Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo - ICMC/USP, como parte dos requisitos para obtenção do título de Especialista em Inteligência Artificial e Big Data.

Área de concentração: Inteligência Artificial

Orientador: Prof. Dr. Jó Ueyama

Versão original

São Carlos

2025

Ficha catalográfica elaborada pela Biblioteca Prof. Achille Bassi
e Seção Técnica de Informática, ICMC/USP,
com os dados inseridos pelo(a) autor(a)

S586i Silva, Luis Roberto da
A Inteligência Artificial na automação de
vistorias imobiliárias: aplicação de redes neurais
convolucionais (CNNs) para detecção de
irregularidades em imagens / Luis Roberto da Silva;
orientador Jó Ueyama. -- São Carlos, 2025.
101 p.

Trabalho de conclusão de curso (MBA em
Inteligência Artificial e Big Data) -- Instituto de
Ciências Matemáticas e de Computação, Universidade
de São Paulo, 2025.

1. Inteligência Artificial aplicada ao setor
imobiliário. 2. Automação de vistorias imobiliárias.
3. Visão Computacional e detecção de defeitos. 4.
Redes neurais convolucionais (CNNs). 5. Deep Learning
em análise de imagens. I. Ueyama, Jó, orient. II.
Título.

Luis Roberto da Silva

Artificial Intelligence in the automation of property inspections: application of convolutional neural networks (CNNs) for detecting irregularities in images

Monograph presented to the Departamento de Ciências de Computação do Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo - ICMC/USP, as part of the requirements for obtaining the title of Specialist in Artificial Intelligence and Big Data.

Concentration area: Artificial Intelligence

Advisor: Prof. Dr. Jó Ueyama

Original version

São Carlos

2025

Este trabalho é dedicado aos alunos da USP, como uma contribuição das Bibliotecas do Campus USP de São Carlos para o desenvolvimento e disseminação da pesquisa científica da Universidade.

AGRADECIMENTOS

Agradeço primeiramente à IV Turma do MBA em Inteligência Artificial e Big Data do Instituto de Ciências Matemáticas e de Computação (ICMC), pela parceria, troca de experiências e apoio mútuo durante toda a jornada acadêmica. Cada colega contribuiu de forma única para o meu aprendizado e para tornar essa caminhada mais leve e enriquecedora.

Expresso minha gratidão aos orientadores, professores e tutores, que tiveram papel fundamental na construção do meu conhecimento e no desenvolvimento deste trabalho. Suas orientações, ensinamentos e dedicação foram essenciais para o meu crescimento pessoal e profissional ao longo do curso.

Agradeço também à coordenação do MBA, em especial à professora Solange Oliveira Rezende, pelo comprometimento, incentivo e apoio constante durante toda a trajetória acadêmica. Sua orientação e atenção foram decisivas para a conclusão desta etapa.

Por fim, estendo meus agradecimentos a todos que, de alguma forma, contribuíram direta ou indiretamente para a realização deste trabalho e para o alcance deste importante objetivo.

‘The greatest value of a picture is when it forces us to notice what we never expected to see.’

*(O maior valor de uma imagem é quando ela nos obriga a perceber o que nunca
esperávamos ver.)*

John Tukey

RESUMO

SILVA, L. R. **A Inteligência Artificial na automação de vistorias imobiliárias: aplicação de redes neurais convolucionais (CNNs) para detecção de irregularidades em imagens.** 2025. 101 p. Monografia (MBA em Inteligência Artificial e Big Data) - Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, 2025.

O mercado imobiliário tem crescido significativamente, aumentando a necessidade de inspeções rápidas, precisas e confiáveis na vistoria de imóveis. O problema central reside na dependência de análises humanas, que são lentas, subjetivas e suscetíveis a erros, comprometendo a tomada de decisão e a eficiência operacional. Este trabalho tem como objetivo geral investigar a aplicação da inteligência artificial, especificamente redes neurais convolucionais (CNNs), na automação da análise de imagens de vistorias, visando a identificação de irregularidades como rachaduras, vidros trincados, bolhas na pintura e pisos desnivelados. Para isso, foi realizada a coleta de um conjunto diversificado de imagens de inspeções, seguido do pré-processamento, incluindo normalização, redução de ruído e técnicas de aumento de dados, com o intuito de melhorar o desempenho dos modelos. Modelos de CNN foram desenvolvidos e treinados para classificar e segmentar defeitos, utilizando métricas de avaliação como acurácia, precisão, recall e F1-score, além de validação cruzada para garantir a robustez dos resultados. Os experimentos demonstraram que os modelos de CNN são capazes de identificar irregularidades com alta precisão, reduzindo significativamente o tempo necessário para análise em comparação às inspeções tradicionais. Observou-se que a segmentação das imagens fornece uma representação visual detalhada dos defeitos, facilitando o registro, a documentação e a tomada de decisão pelos responsáveis pelas vistorias. Os resultados indicam ainda que a automação baseada em inteligência artificial oferece maior consistência nas avaliações, minimizando a subjetividade inerente à inspeção humana. Conclui-se que a utilização de CNNs constitui uma ferramenta eficiente para apoiar a autorização de vistorias imobiliárias, proporcionando ganhos em confiabilidade, objetividade e produtividade. A aplicação desta tecnologia contribui para a otimização de processos operacionais e redução de custos, além de estabelecer uma base sólida para pesquisas futuras, que podem explorar métodos avançados de visão computacional e aprendizado profundo em contextos similares. A integração de inteligência artificial em processos de vistoria representa uma evolução significativa para o setor, permitindo decisões mais rápidas e embasadas em dados, ao mesmo tempo em que promove maior segurança e qualidade na avaliação de imóveis.

Palavras-chave: Inteligência Artificial aplicada ao setor imobiliário. Automação de vistorias imobiliárias. Visão Computacional e detecção de defeitos. Redes neurais convolucionais (CNNs). Deep Learning em análise de imagens.

ABSTRACT

SILVA, L. R. **Artificial Intelligence in the automation of property inspections: application of convolutional neural networks (CNNs) for detecting irregularities in images**. 2025. 101 p. Monograph (MBA in Artificial Intelligence and Big Data) - Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos, 2025.

The real estate market has grown significantly, increasing the need for fast, accurate, and reliable property inspections. The main problem lies in the reliance on human analysis, which is slow, subjective, and prone to errors, compromising decision-making and operational efficiency. This study aims to investigate the application of artificial intelligence, specifically convolutional neural networks (CNNs), to automate the analysis of inspection images, aiming to identify irregularities such as cracks, broken glass, paint bubbles, and uneven floors. For this purpose, a diverse dataset of inspection images was collected and preprocessed, including normalization, noise reduction, and data augmentation techniques to enhance model performance. CNN models were developed and trained to classify and segment defects, using evaluation metrics such as accuracy, precision, recall, and F1-score, along with cross-validation to ensure result robustness. Experiments showed that CNN models can identify irregularities with high precision, significantly reducing the time required for analysis compared to traditional inspections. It was also observed that image segmentation provides a detailed visual representation of defects, facilitating documentation and decision-making by inspection personnel. Results indicate that AI-based automation offers greater consistency in evaluations, minimizing the subjectivity inherent in human inspection. It is concluded that the use of CNNs is an efficient tool to support the authorization of property inspections, providing gains in reliability, objectivity, and productivity. The application of this technology contributes to process optimization and cost reduction, while establishing a solid foundation for future research exploring advanced computer vision and deep learning methods in similar contexts. Integrating artificial intelligence into inspection processes represents a significant advancement for the sector, enabling faster, data-driven decisions while promoting greater safety and quality in property assessments.

Keywords: Artificial Intelligence applied to the real estate sector. Automation of property inspections. Computer vision and defect detection. Convolutional Neural Networks (CNNs). Deep Learning in image analysis.

LISTA DE FIGURAS

Figura 1 – Principais marcos da evolução da IA e do Deep Learning (adaptado de PO (2024)).	34
Figura 2 – Estrutura de uma Rede Neural Convolutacional (CNN).	35
Figura 3 – Fluxo detalhado do processo de treinamento e avaliação da CNN, conforme implementado no notebook do experimento.	42
Figura 4 – Arquitetura da solução implementada na AWS para automação de vistorias imobiliárias.	45
Figura 5 – Amostra de fissura	47
Figura 6 – Amostra de rachadura	47
Figura 7 – Amostra de infiltração	47
Figura 8 – Amostra de foto sem defeito	47
Figura 9 – Curvas de acurácia e perda – modelo EfficientNetB0 (Fold 1).	49
Figura 10 – Matriz de confusão – modelo EfficientNetB0 (Fold 1).	49
Figura 11 – Gráfico de Barras - Comparativo de desempenho entre os modelos CNN – MobileNetV2, EfficientNetB0 e VGG16.	51
Figura 12 – Gráfico Radar - Comparativo de desempenho entre os modelos CNN – MobileNetV2, EfficientNetB0 e VGG16.	51
Figura 13 – Mapa de Empatia - Persona - Gestor de Vistorias Imobiliárias - B2B	55
Figura 14 – Mapa de Empatia - Persona - Vistoriador Autônomo - B2C	56
Figura 15 – Canvas Business Model – Vistoria Imobiliária com IA	58
Figura 16 – <i>Análise SWOT – Vistoria Imobiliária com IA</i>	59
Figura 17 – Tela de Login — interface de login do usuário.	63
Figura 18 – Tela de Cadastro de Usuário — interface de cadastro do usuário.	64
Figura 19 – Tela da Home — interface de Home do aplicativo.	65
Figura 20 – Tela de Cadastro de Vistoria — interface para cadastro de vistoria.	66
Figura 21 – Tela de Consulta de Vistoria — interface para consulta de vistoria.	67
Figura 22 – Tela de Vistoria — interface principal captura de fotos e geração do laudo técnico.	68
Figura 23 – Tela de Laudo Técnico — interface laudo técnico da vistoria.	69
Figura 24 – Roadmap de funcionalidades futuras da plataforma de vistorias com IA.	77
Figura 25 – Curvas de acurácia e perda — EfficientNetB0 Fold 1.	87
Figura 26 – Matriz de confusão — EfficientNetB0 Fold 1.	87
Figura 27 – Curvas de acurácia e perda — EfficientNetB0 Fold 2.	88
Figura 28 – Matriz de confusão — EfficientNetB0 Fold 2.	88
Figura 29 – Curvas de acurácia e perda — EfficientNetB0 Fold 3.	89
Figura 30 – Matriz de confusão — EfficientNetB0 Fold 3.	89

Figura 31 – Curvas de acurácia e perda — EfficientNetB0 Fold 4.	90
Figura 32 – Matriz de confusão — EfficientNetB0 Fold 4.	90
Figura 33 – Curvas de acurácia e perda — EfficientNetB0 Fold 5.	91
Figura 34 – Matriz de confusão — EfficientNetB0 Fold 5.	91
Figura 35 – Curvas de acurácia e perda — MobileNet Fold 1.	92
Figura 36 – Matriz de confusão — MobileNet Fold 1.	92
Figura 37 – Curvas de acurácia e perda — MobileNet Fold 2.	93
Figura 38 – Matriz de confusão — MobileNet Fold 2.	93
Figura 39 – Curvas de acurácia e perda — MobileNet Fold 3.	94
Figura 40 – Matriz de confusão — MobileNet Fold 3.	94
Figura 41 – Curvas de acurácia e perda — MobileNet Fold 4.	95
Figura 42 – Matriz de confusão — MobileNet Fold 4.	95
Figura 43 – Curvas de acurácia e perda — MobileNet Fold 5.	96
Figura 44 – Matriz de confusão — MobileNet Fold 5.	96
Figura 45 – Curvas de acurácia e perda — MobileNet Fold 1.	97
Figura 46 – Matriz de confusão — VGG16 Fold 1.	97
Figura 47 – Curvas de acurácia e perda — MobileNet Fold 2.	98
Figura 48 – Matriz de confusão — VGG16 Fold 2.	98
Figura 49 – Curvas de acurácia e perda — MobileNet Fold 3.	99
Figura 50 – Matriz de confusão — VGG16 Fold 3.	99
Figura 51 – Curvas de acurácia e perda — MobileNet Fold 4.	100
Figura 52 – Matriz de confusão — VGG16 Fold 4.	100
Figura 53 – Curvas de acurácia e perda — MobileNet Fold 5.	101
Figura 54 – Matriz de confusão — VGG16 Fold 5.	101

LISTA DE TABELAS

Tabela 1	– Resultados médios de desempenho – CNNs de classificação (MVP)	. . . 48
Tabela 2	– Canvas do Problema – Automação de Vistorias Imobiliárias com IA	. . . 57
Tabela 3	– Benchmarking Tecnológico – Indicadores de Análise	61
Tabela 4	– Resultados por fold – MobileNetV2	83
Tabela 5	– Resultados por fold – EfficientNetB0	83
Tabela 6	– Resultados por fold – VGG16	84

LISTA DE ABREVIATURAS E SIGLAS

ABNT	Associação Brasileira de Normas Técnicas
abnTeX	ABsurdas Normas para TeX
IBGE	Instituto Brasileiro de Geografia e Estatística
LaTeX	Lamport TeX
USP	Universidade de São Paulo
USPSC	Campus USP de São Carlos
IA	Inteligência Artificial
CNN	<i>Convolutional Neural Network</i> (Rede Neural Convolucional)
CNNs	<i>Convolutional Neural Networks</i> (Redes Neurais Convolucionais)
YOLO	<i>You Only Look Once</i> (Você Só Olha Uma Vez)
R-CNN	<i>Region-based Convolutional Neural Network</i> (Rede Neural Convolucional Baseada em Regiões)
Faster R-CNN	<i>Faster Region-based Convolutional Neural Network</i> (Versão aprimorada do R-CNN para detecção mais rápida)
Mask R-CNN	<i>Mask Region-based Convolutional Neural Network</i> (R-CNN com segmentação de máscara por pixel)
U-Net	<i>U-shaped Network</i> (Rede em formato de U, usada em segmentação de imagens)
ViT	<i>Vision Transformer</i> (Transformador de Visão)
MVP	<i>Minimum Viable Product</i> (Produto Mínimo Viável)
GPU	<i>Graphics Processing Unit</i> (Unidade de Processamento Gráfico)
API	<i>Application Programming Interface</i> (Interface de Programação de Aplicações)
ERP	<i>Enterprise Resource Planning</i> (Planejamento de Recursos Empresariais)
CRM	<i>Customer Relationship Management</i> (Gestão de Relacionamento com o Cliente)

IoT	<i>Internet of Things</i> (Internet das Coisas)
Big Data	Grandes volumes de dados gerados e analisados por sistemas digitais
CBIC	Câmara Brasileira da Indústria da Construção
CRISP-DM	<i>Cross Industry Standard Process for Data Mining</i> (Padrão de Processo Cruzado para Mineração de Dados)
ROI	<i>Return on Investment</i> (Retorno sobre o Investimento)
PropTech	<i>Property Technology</i> (Tecnologia aplicada ao setor imobiliário)
Python	Linguagem de programação amplamente utilizada em ciência de dados e IA
NumPy	<i>Numerical Python</i> (Biblioteca Python para computação numérica)
Pandas	Biblioteca Python para análise e manipulação de dados
Scikit-learn	Biblioteca Python para aprendizado de máquina
OpenCV	<i>Open Source Computer Vision Library</i> (Biblioteca de Visão Computacional de Código Aberto)
Matplotlib	Biblioteca Python para visualização de dados
TorchVision	Biblioteca Python para visão computacional baseada em PyTorch
AWS	<i>Amazon Web Services</i> (Serviços de Computação em Nuvem da Amazon)
Google Cloud	Plataforma de serviços em nuvem do Google
Azure	Plataforma de serviços em nuvem da Microsoft

SUMÁRIO

1	INTRODUÇÃO	27
1.1	Contextualização	27
1.2	Problema	27
2	A INOVAÇÃO E A SOLUÇÃO PROPOSTA	29
3	FUNDAMENTOS DOS CONCEITOS DE IA E BIG DATA	31
3.1	Problemas e Desafios nas Vistorias Imobiliárias	31
3.2	Inteligência Artificial e Aprendizado de Máquina	33
3.3	Evolução do Deep Learning e Marcos Tecnológicos	34
3.4	<i>Deep Learning</i> e Estrutura em Camadas	35
3.5	Redes Neurais Convolucionais (CNNs)	36
3.6	<i>Transfer Learning</i> e <i>Fine-Tuning</i>	36
3.7	<i>Data Augmentation</i>	36
3.8	Hiperparâmetros e Processo de Treinamento	36
3.9	Validação Cruzada <i>K-Fold</i>	36
3.10	Métricas de Avaliação	36
3.11	Arquiteturas Empregadas	37
3.12	Metodologia CRISP-DM	37
3.13	Estado da Arte	37
3.14	Síntese do Capítulo	37
4	METODOLOGIA	39
4.1	Referenciais Técnicos e Científicos	39
4.2	Etapas do Processo Metodológico	40
5	RESULTADOS EXPERIMENTAIS E ANÁLISE DOS DADOS	41
5.1	Descrição do Experimento e Configuração do Ambiente	41
5.2	Arquitetura da Solução para implantação na Nuvem AWS	45
5.3	Conjunto de Dados e Pré-processamento	46
5.4	Amostra de dados	47
5.5	Treinamento e Avaliação dos Modelos	48
5.6	Resultados e Visualizações	49
5.7	Comparativo de Desempenho entre Modelos CNN	51
5.8	Discussão dos Resultados	52
5.9	Limitações e Trabalhos Futuros	52

6	PANORAMA DE MERCADO, INVESTIMENTOS E RESULTADOS ESPERADOS	53
6.1	Jornada do Usuário	55
6.1.1	Mapa de Empatia - Persona: Gestor de Vistorias / Imobiliária - B2B	55
6.1.2	Mapa de Empatia - Persona: Vistoriador Autônomo - B2C	56
6.2	Canvas do Problema	57
6.3	Modelo de Negócio	58
6.4	Análise SWOT	59
6.5	Benchmarking Tecnológico	60
6.6	Produto Mínimo Viável (MVP)	62
6.6.1	Objetivo do MVP	62
6.6.2	Componentes do MVP	62
6.6.3	Telas do Protótipo (MVP)	63
6.6.3.1	Tela de Login	63
6.6.3.2	Tela de Cadastro de Usuário	64
6.6.3.3	Tela de Home	65
6.6.3.4	Tela de Cadastro de Vistoria	66
6.6.3.5	Tela de Consulta de Vistoria	67
6.6.3.6	Tela de Vistoria	68
6.6.3.7	Tela de Laudo Técnico	69
6.6.4	Critérios de Sucesso do MVP	70
6.6.5	Limitações Previstas	70
6.6.6	Síntese do MVP	70
6.7	Investimentos e Retorno Esperado	71
6.7.1	Componentes de Investimento	71
6.7.2	Retorno sobre o Investimento	72
6.7.3	Análise Econômica e Estratégica	72
7	CONCLUSÕES	75
8	TRABALHOS FUTUROS	77
	REFERÊNCIAS	79
	APÊNDICES	81
	APÊNDICE A – RESULTADOS DETALHADOS DA VALIDAÇÃO CRUZADA (K-FOLD)	83
A.1	Metodologia de Avaliação	83

A.2	Resultados Obtidos	83
A.3	Análise de Estabilidade	84
	APÊNDICE B – CÓDIGO-FONTE DO NOTEBOOK DE TREINA- MENTO E VALIDAÇÃO	85
	APÊNDICE C – SAÍDAS GRÁFICAS DOS MODELOS E FOLDS .	87
C.1	EfficientNetB0	87
C.2	MobileNet	92
C.3	VGG16	97

1 INTRODUÇÃO

1.1 Contextualização

O mercado imobiliário brasileiro representa uma das principais forças econômicas do país, movimentando cerca de R\$ 300 bilhões anualmente, segundo dados da CIBC (2024). Estudos do setor indicam que o mercado de imóveis residenciais tem apresentado crescimento sustentado, impulsionado por avanços tecnológicos e pela digitalização dos processos de avaliação e entrega de empreendimentos (Mordor Intelligence, 2024). Dentro desse contexto, a etapa de vistoria de imóveis é fundamental para garantir a transparência e a segurança nas transações imobiliárias, seja na entrega de empreendimentos novos, na locação ou na revenda de unidades.

Apesar de sua relevância, o processo de vistoria tradicional ainda é majoritariamente manual e sujeito a falhas humanas, inconsistências em laudos e atrasos. Essas limitações afetam diretamente a confiabilidade das avaliações e aumentam os custos operacionais de construtoras e administradoras de imóveis.

Nos últimos anos, a transformação digital tem impulsionado a adoção de tecnologias como Inteligência Artificial (IA), Big Data e Visão Computacional em diversos setores, promovendo eficiência e precisão em atividades antes dependentes de julgamentos humanos (Faceli *et al.*, 2021). No setor imobiliário, essas tecnologias têm potencial para revolucionar o modo como as inspeções são realizadas, substituindo processos manuais por sistemas inteligentes capazes de identificar irregularidades com alto grau de acurácia.

1.2 Problema

O processo de vistoria imobiliária apresenta desafios significativos, principalmente pela subjetividade das avaliações, pela morosidade na coleta de informações e pela dificuldade de padronizar relatórios técnicos. A inspeção visual humana está sujeita a erros, fadiga e interpretações divergentes — fatores que comprometem a consistência e a credibilidade das análises.

Essas limitações se tornam ainda mais evidentes diante da crescente demanda por rapidez e precisão em transações imobiliárias. Construtoras e imobiliárias enfrentam pressão por reduzir custos, aumentar a produtividade e padronizar processos sem comprometer a qualidade. Assim, a questão central que orienta este estudo é:

A Inteligência Artificial é capaz de automatizar o processo de vistoria imobiliária, mantendo ou superando a precisão da análise humana?

Nota sobre o uso de Inteligência Artificial

Para apoiar o processo de escrita, revisão e tradução deste trabalho, foi utilizada a ferramenta de Inteligência Artificial **ChatGPT** (modelo GPT-5), acessada por meio da plataforma da OpenAI, durante o período de elaboração do texto em 2025. O uso da ferramenta restringiu-se à verificação linguística, sugestões de reformulação textual, adequação de estilo acadêmico e tradução de passagens específicas. Todo o conteúdo, análise, interpretação e a redação final foram integralmente realizados e validados pelo autor, que assume total responsabilidade pelo texto apresentado.

2 A INOVAÇÃO E A SOLUÇÃO PROPOSTA

O uso de Inteligência Artificial (IA) e Visão Computacional no tratamento de imagens de vistorias imobiliárias representa um avanço significativo em um setor tradicionalmente dependente de análises manuais. Processos de vistoria, que antes exigiam observação humana e registros subjetivos, podem agora ser apoiados por sistemas capazes de interpretar e padronizar visualmente as condições de um imóvel a partir de fotografias.

Nos últimos anos, o campo da Visão Computacional evoluiu rapidamente, impulsionado por redes neurais profundas e pelo aumento da capacidade computacional. As redes neurais convolucionais (CNNs) têm se destacado como a principal técnica para o tratamento de imagens, permitindo extrair padrões visuais complexos e identificar características relevantes em diversos contextos, como saúde, agricultura e inspeções estruturais GOODFELLOW Ian; BENGIO (2016), RAWAT Waseem; WANG (2017).

A inovação proposta neste trabalho consiste na utilização de modelos de *Transfer Learning* previamente treinados em grandes bases de dados, como a ImageNet, e adaptados para o domínio das vistorias residenciais. Essa abordagem reduz o tempo de desenvolvimento, diminui o custo de treinamento e possibilita resultados de alta precisão mesmo com um volume limitado de imagens PAN Sinno Jialin; YANG (2009). O ajuste fino (*fine-tuning*) das camadas finais desses modelos permite reconhecer padrões específicos, como manchas, fissuras ou imperfeições em superfícies, a partir de um processo automatizado de análise visual.

O sistema desenvolvido foi implementado em *Python*, utilizando bibliotecas científicas amplamente reconhecidas, como NumPy, Pandas, Scikit-learn, TorchVision e OpenCV. A execução foi realizada no ambiente *Google Colab*, com integração à *AWS S3* para armazenamento dos resultados e logs de execução.

Foram avaliadas três arquiteturas de referência — MobileNetV2, EfficientNetB0 e VGG16 — escolhidas por sua eficiência, estabilidade e ampla aceitação na literatura SANDLER Mark; HOWARD (2018), TAN Mingxing; LE (2019), SIMONYAN Karen; ZISSERMAN (2014). As métricas utilizadas para mensurar o desempenho incluíram acurácia, precisão, revocação e F1-score, com validação cruzada (*K-Fold*) para garantir robustez estatística.

A metodologia de desenvolvimento adotada segue o modelo CRISP-DM, amplamente utilizado em projetos de ciência de dados, que organiza o trabalho em etapas de entendimento do negócio, preparação dos dados, modelagem, avaliação e implantação WIRTH Rüdiger; HIPPE (2000). Essa estrutura garante a coerência entre os objetivos técnicos e o propósito aplicado do projeto.

Com a proposta de automatizar o tratamento das imagens e a geração de relatórios de vistoria, este trabalho busca reduzir o tempo de análise, minimizar erros humanos e aumentar a padronização dos laudos. A iniciativa reforça o movimento de digitalização e inteligência no setor imobiliário, promovendo um processo de vistoria mais preciso, eficiente e escalável.

3 FUNDAMENTOS DOS CONCEITOS DE IA E BIG DATA

Este capítulo apresenta os conceitos que fundamentam a proposta de uso de Inteligência Artificial (IA) e *Deep Learning* na automação de vistorias imobiliárias. São descritos os principais tipos de problemas observados nas inspeções, os fundamentos teóricos da IA e do Aprendizado de Máquina, a evolução tecnológica do *Deep Learning*, bem como os elementos técnicos que compõem a arquitetura utilizada neste trabalho, como redes neurais convolucionais (CNNs), *transfer learning*, *data augmentation*, hiperparâmetros, validação cruzada e métricas de avaliação. O capítulo encerra com uma análise do estado da arte.

3.1 Problemas e Desafios nas Vistorias Imobiliárias

As vistorias imobiliárias são processos fundamentais para garantir a integridade estrutural e a conformidade de imóveis. No entanto, quando realizadas de forma manual, enfrentam desafios que comprometem a precisão e a eficiência dos resultados. Entre os problemas mais comuns destacam-se:

- Subjetividade na análise: a identificação de defeitos como fissuras e infiltrações depende da interpretação humana, gerando inconsistências;
- Tempo elevado de execução: o registro manual e a elaboração de relatórios prolongam o ciclo de vistoria;
- Variações visuais: fatores como iluminação e ângulo dificultam comparações entre fotos de períodos distintos;
- Dificuldade de padronização: ausência de critérios automatizados reduz a reprodutibilidade dos resultados;
- Falta de integração tecnológica: sistemas tradicionais não exploram técnicas de IA nem de visão computacional.

Neste cenário é recorrente a identificação de falhas **hidráulicas**, **elétricas**, **estruturais**, de **acabamento** e de **umidade/segurança**. Mapear esses problemas orienta os requisitos do modelo e do laudo técnico.

Hidráulica

- **Vazamentos:** em torneiras, vasos sanitários, pias e tubulações, com risco de infiltrações.

- **Infiltrações e mofo:** manchas de umidade e bolor em paredes/tetos (impermeabilização deficiente).
- **Pressão de água:** baixa pressão em chuveiros e torneiras.
- **Entupimentos:** escoamento lento em ralos e tubulações.

Elétrica

- **Fiação inadequada:** fios expostos, emendas irregulares, instalações antigas.
- **Tomadas e interruptores:** peças soltas, mau funcionamento, ausência de aterramento.
- **Quadro de energia:** disjuntores desarmando com frequência, identificação deficiente.

Estrutural

- **Rachaduras e fissuras:** em paredes, tetos e pisos (avaliar espessura/propagação).
- **Desníveis:** desalinhamentos de piso/contrapiso; portas sem esquadro.
- **Cobertura:** telhas quebradas ou deslocadas, calhas obstruídas.
- **Madeira deteriorada:** presença de cupins ou apodrecimento.

Acabamento e esquadrias

- **Pisos e revestimentos:** peças soltas/ocas, trincas, empenamento.
- **Pintura:** descascamento, bolhas, manchas que podem ocultar patologias.
- **Portas e janelas:** esquadrias empenadas, fechamento irregular, *vidros trincados*.

Segurança, ventilação e salubridade

- **Ventilação inadequada:** favorece umidade, odores e mofo.
- **Equipamentos de segurança:** ausência/defeito em detectores (fumaça/CO).
- **Risco de queda:** corrimãos frouxos, degraus danificados, pisos escorregadios.

Essas classes de problemas orientam a **taxonomia de defeitos** usada na rotulagem e nas métricas de avaliação do modelo.

3.2 Inteligência Artificial e Aprendizado de Máquina

A Inteligência Artificial (IA) busca reproduzir habilidades cognitivas humanas por meio de sistemas computacionais RUSSELL Stuart J.; NORVIG (2010). O Aprendizado de Máquina (AM) é um subconjunto da IA em que algoritmos aprendem padrões a partir de dados, sem serem explicitamente programados GOODFELLOW Ian; BENGIO (2016).

O Aprendizado de Máquina pode ser:

- **Supervisionado** – utiliza dados rotulados para treinar o modelo;
- **Não supervisionado** – identifica agrupamentos ou padrões sem rótulos;
- **Por reforço** – aprende por tentativa e erro a partir de recompensas.

O presente trabalho adota o aprendizado supervisionado, onde as imagens de vistorias são rotuladas conforme a presença ou ausência de defeitos.

3.3 Evolução do Deep Learning e Marcos Tecnológicos

A trajetória da IA é marcada por sucessivos períodos de avanço e retração. Desde o neurônio artificial de McCulloch e Pitts (1943) e o Perceptron de Rosenblatt (1957), até o renascimento das redes neurais profundas em 2012 com o modelo AlexNet, o campo passou por transformações profundas.

A Figura 1 apresenta uma síntese dos principais marcos históricos do *Deep Learning*, adaptada de PO (2024). São destacados o desenvolvimento do algoritmo de retropropagação (1986), o surgimento das GPUs como aceleradores de aprendizado (2009) e o sucesso de arquiteturas como CNNs, ResNet e Transformers.

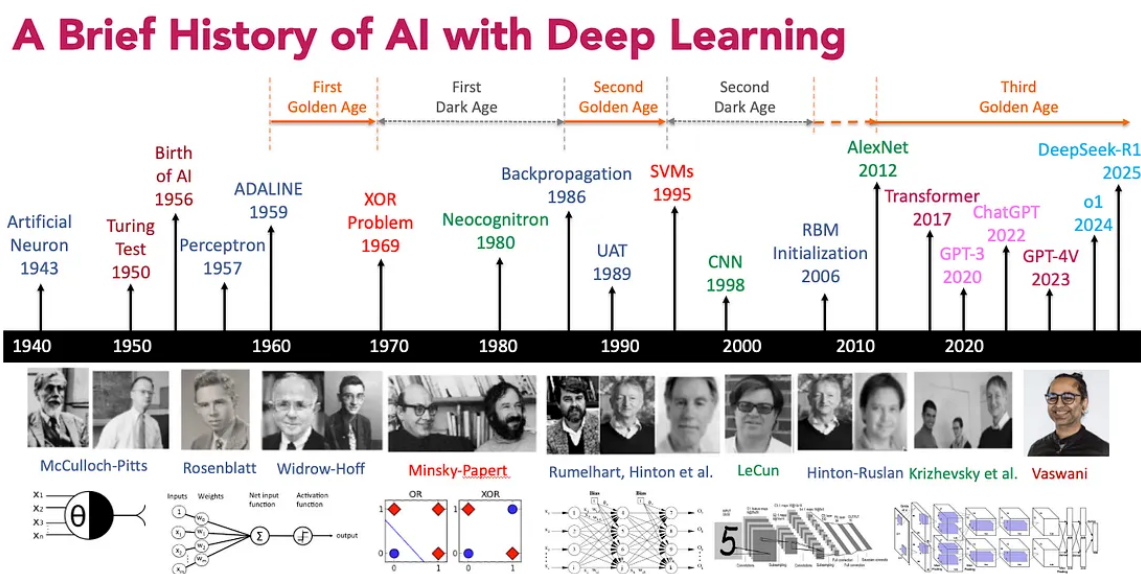


Figura 1 – Principais marcos da evolução da IA e do Deep Learning (adaptado de PO (2024)).

Fonte: Adaptado de PO (2024), 2024.

Esses avanços tecnológicos viabilizaram a aplicação do *Deep Learning* em áreas como visão computacional, processamento de linguagem natural e automação de vistorias imobiliárias.

3.4 *Deep Learning* e Estrutura em Camadas

O *Deep Learning* utiliza redes neurais profundas compostas por múltiplas camadas hierárquicas. Cada camada extrai características de maior complexidade a partir da anterior, formando uma representação progressiva da imagem GOODFELLOW Ian; BENGIO (2016).

- **Camada de entrada:** recebe a imagem (ex.: $224 \times 224 \times 3$);
- **Camadas convolucionais:** aplicam filtros para identificar bordas e texturas;
- **Pooling:** reduz a dimensionalidade mantendo as informações relevantes;
- **Camadas densas:** combinam as características aprendidas;
- **Camada de saída:** realiza a classificação por meio de funções como *Softmax*.

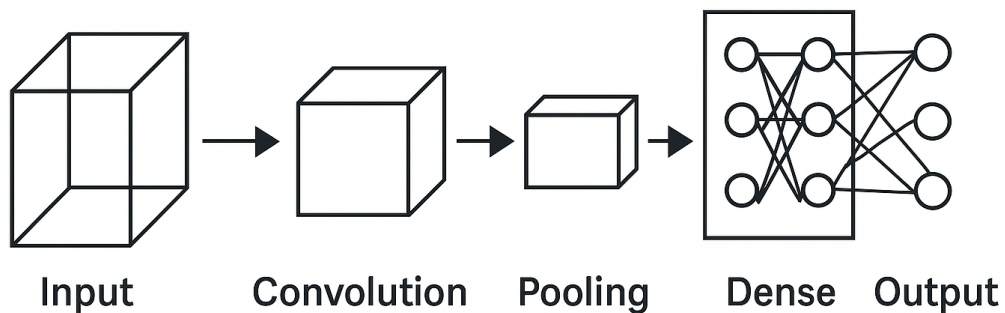


Figura 2 – Estrutura de uma Rede Neural Convolucional (CNN).

Fonte: Elaboração própria (2025)

3.5 Redes Neurais Convolucionais (CNNs)

As CNNs são amplamente utilizadas em visão computacional por sua capacidade de aprender automaticamente representações espaciais das imagens RAWAT Waseem; WANG (2017). Sua estrutura alterna camadas de convolução, ativação e *pooling*, seguidas de camadas densas totalmente conectadas. A técnica de *dropout* é empregada para evitar o *overfitting*, desativando neurônios aleatoriamente durante o treinamento.

3.6 Transfer Learning e Fine-Tuning

O *Transfer Learning* permite reutilizar redes previamente treinadas em bases de grande escala (como ImageNet), adaptando-as a novos problemas PAN Sinno Jialin; YANG (2009). O *fine-tuning* ajusta as camadas finais do modelo com dados específicos, utilizando taxas de aprendizado menores para preservar o conhecimento anterior.

3.7 Data Augmentation

Essa técnica aumenta artificialmente o conjunto de dados por meio de transformações geométricas e fotométricas: rotação, espelhamento, recorte e variação de brilho. Ela melhora a generalização e reduz o risco de sobreajuste RAWAT Waseem; WANG (2017).

3.8 Hiperparâmetros e Processo de Treinamento

Os hiperparâmetros controlam o comportamento do treinamento, como tamanho do *batch*, número de épocas, taxa de aprendizado e regularização. O *early stopping* interrompe o treino quando a métrica de validação deixa de melhorar, prevenindo o *overfitting*.

3.9 Validação Cruzada *K-Fold*

A validação cruzada *K-Fold* divide o conjunto de dados em K partes, treinando o modelo K vezes, cada vez com uma parte diferente para validação. No experimento foi adotado $K = 5$, resultando em métricas médias mais estáveis e confiáveis.

3.10 Métricas de Avaliação

Para avaliar o desempenho do modelo foram utilizadas as seguintes métricas:

- **Acurácia:** proporção de previsões corretas;
- **Precisão:** relação entre verdadeiros positivos e positivos preditos;
- **Revocação:** proporção de positivos reais identificados corretamente;
- **F1-Score:** média harmônica entre precisão e revocação;

- **Matriz de confusão:** comparação entre rótulos reais e previstos;
- **AUC-ROC:** mede a capacidade discriminativa do modelo.

3.11 Arquiteturas Empregadas

Três CNNs pré-treinadas foram utilizadas:

- **MobileNetV2** – leve e eficiente, indicada para dispositivos com restrição de recursos SANDLER Mark; HOWARD (2018).
- **EfficientNetB0** – aplica o escalonamento composto, equilibrando profundidade e largura TAN Mingxing; LE (2019).
- **VGG16** – modelo clássico com convoluções 3×3 e camadas densas SIMONYAN Karen; ZISSERMAN (2014).

3.12 Metodologia CRISP-DM

O desenvolvimento seguiu o modelo CRISP-DM WIRTH Rüdiger; HIPPE (2000), que define seis fases: entendimento do negócio, entendimento dos dados, preparação, modelagem, avaliação e implantação. Essa metodologia garante reprodutibilidade e rastreabilidade dos experimentos.

3.13 Estado da Arte

Estudos recentes demonstram o uso de CNNs para inspeção de estruturas civis e detecção de falhas visuais RAWAT Waseem; WANG (2017). Arquiteturas como AlexNet, ResNet e Inception representam marcos na evolução da área, servindo de base para novas soluções. Esses avanços tecnológicos sustentam a aplicação proposta neste trabalho.

3.14 Síntese do Capítulo

Foram abordados os principais conceitos e fundamentos do Deep Learning aplicados à automação de vistorias imobiliárias. Os temas apresentados neste capítulo formam a base conceitual necessária para a compreensão da metodologia e dos resultados discutidos nos capítulos seguintes.

4 METODOLOGIA

A metodologia adotada neste trabalho segue o modelo **CRISP-DM** (**Cross-Industry Standard Process for Data Mining**), amplamente utilizado em projetos de ciência de dados e inteligência artificial. Esse modelo organiza o desenvolvimento em seis etapas principais: compreensão do negócio, compreensão dos dados, preparação dos dados, modelagem, avaliação e implantação.

4.1 Referenciais Técnicos e Científicos

A fundamentação teórica e técnica deste projeto baseia-se em modelos consolidados de *deep learning* aplicados à *classificação de imagens* por meio de redes neurais convolucionais (CNNs). O foco desta etapa concentrou-se na seleção e análise de arquiteturas pré-treinadas amplamente utilizadas em tarefas de reconhecimento visual, empregando técnicas de *transfer learning* e *fine-tuning* para adaptação ao domínio das vistorias imobiliárias.

Entre os principais modelos avaliados e implementados nesta fase do Produto Mínimo Viável (MVP) estão:

- **MobileNetV2**: arquitetura leve e eficiente, projetada para aplicações em dispositivos com restrições de processamento. Seu uso é indicado para cenários de inferência em campo, como aplicativos móveis de vistoria.
- **EfficientNetB0**: modelo que otimiza a relação entre desempenho e custo computacional por meio do balanceamento composto de profundidade, largura e resolução. Apresentou os melhores resultados no experimento de validação cruzada, conforme descrito no Capítulo 5.
- **VGG16**: rede clássica de 16 camadas, amplamente utilizada como base de comparação em estudos de visão computacional. Embora mais pesada, serve como referência para aferir ganhos em arquiteturas mais recentes.

Esses modelos foram comparados de acordo com métricas de acurácia, precisão, *recall* e F1-score, com o objetivo de identificar a arquitetura mais adequada à classificação das imagens de vistorias. As técnicas de *transfer learning* e *fine-tuning* permitiram reduzir o tempo de treinamento e melhorar a capacidade de generalização, mesmo com um conjunto de dados limitado.

Os resultados consolidados dessas avaliações estão detalhados no Capítulo 5 e nas figuras apresentadas no Apêndice C.

4.2 Etapas do Processo Metodológico

1. Compreensão do Negócio: análise do processo de vistoria imobiliária tradicional e identificação dos principais gargalos e oportunidades de automação.

2. Compreensão e Preparação dos Dados: coleta de imagens de vistorias, rotulagem supervisionada das irregularidades e limpeza dos dados.

3. Modelagem: treinamento e validação do modelo de IA utilizando técnicas de *transfer learning* e métricas de precisão, recall e F1-score.

4. Avaliação: comparação dos resultados entre modelos, considerando precisão na detecção, tempo de inferência e custo de processamento.

5. Implantação: disponibilização do modelo em ambiente **cloud-native** na AWS, integrando o **SageMaker** ao **Lambda** e **API Gateway**, permitindo acesso via aplicativo móvel e portal administrativo.

Essa abordagem metodológica integra os referenciais técnicos com a aplicação prática, garantindo que o desenvolvimento do modelo de IA esteja alinhado às necessidades do negócio e às boas práticas de engenharia de software e dados.

5 RESULTADOS EXPERIMENTAIS E ANÁLISE DOS DADOS

Este capítulo apresenta os resultados obtidos com o experimento desenvolvido no Google Colab, utilizando modelos de redes neurais convolucionais (CNNs) voltados à **classificação de imagens** na etapa de validação do Produto Mínimo Viável (MVP). Nesta fase inicial, foram avaliadas arquiteturas amplamente utilizadas em aplicações de visão computacional — MobileNetV2, EfficientNetB0 e VGG16 — com foco em verificar a viabilidade técnica da Inteligência Artificial na automação de vistorias imobiliárias.

5.1 Descrição do Experimento e Configuração do Ambiente

Os experimentos foram realizados na plataforma *Google Colab*, em ambiente de GPU L4 com 52 GB de *RAM*, utilizando a linguagem *Python 3.10* e as bibliotecas *NumPy*, *Pandas*, *Scikit-learn*, *TorchVision*, *Matplotlib* e *OpenCV*. O Colab foi integrado ao serviço *Amazon S3*, que funcionou como repositório de armazenamento dos conjuntos de dados e resultados intermediários, garantindo rastreabilidade e persistência dos artefatos gerados durante a execução.

Os principais parâmetros de configuração adotados foram:

- **Batch size:** 16;
- **Épocas (head):** 20 – fase de treinamento das camadas superiores;
- **Épocas (fine-tuning):** 15 – ajuste fino das camadas inferiores;
- **K-folds:** 5 – validação cruzada (*5-fold cross-validation*);
- **Tamanho da imagem:** 224×224 pixels;
- **Patience:** 8 – critério de *early stopping*.

O experimento teve como objetivo principal avaliar o desempenho de arquiteturas de redes neurais convolucionais (CNNs) — **MobileNetV2**, **EfficientNetB0** e **VGG16** — aplicadas à classificação de imagens de vistorias imobiliárias. O processo completo de execução foi implementado no notebook `tcc_vistoria_transfer_learning_v3.ipynb`, refletindo a metodologia empregada para o desenvolvimento e validação do Produto Mínimo Viável (MVP).

A Figura 3 apresenta o fluxo detalhado do experimento, desde a coleta e preparação das imagens até a avaliação e armazenamento dos resultados. Esse pipeline ilustra o encadeamento das etapas descritas a seguir.

Processo de Treinamento



Figura 3 – Fluxo detalhado do processo de treinamento e avaliação da CNN, conforme implementado no notebook do experimento.

O processo experimental foi dividido em seis etapas principais, descritas a seguir:

1. Coleta das Imagens

As imagens foram obtidas em vistorias reais e organizadas em um bucket no *Amazon S3*. O conjunto foi dividido em quatro classes principais: *fissura*, *infiltração*, *rachadura* e *sem defeito*, assegurando a representatividade e diversidade das amostras.

2. Pré-processamento e Data Augmentation

O conjunto de imagens passou por etapas de padronização e ampliação dos dados, incluindo:

- Eliminação de imagens duplicadas e inconsistentes;
- Correção automática de rotação baseada em metadados EXIF;
- Redimensionamento das imagens para 224×224 pixels;
- Normalização dos valores de pixel para a faixa $[0,1]$;
- Aplicação de *Data Augmentation* (flip, zoom, rotação e variação de brilho).

Essas transformações reduziram o risco de *overfitting* e melhoraram a capacidade de generalização do modelo.

3. Divisão K-Fold

A técnica de validação cruzada (*K-Fold Cross Validation*) foi aplicada com $K = 5$, permitindo que o modelo fosse treinado e validado em diferentes subconjuntos de dados. Essa abordagem garantiu maior robustez estatística e reduziu o impacto de variações específicas do conjunto de treinamento.

4. Treinamento e Fine-Tuning

As arquiteturas **MobileNetV2**, **EfficientNetB0** e **VGG16** foram utilizadas com pesos pré-treinados na base *ImageNet*, aplicando a técnica de *Transfer Learning*. As camadas finais foram substituídas por classificadores densos adaptados ao problema de quatro classes, e o *Fine-Tuning* foi realizado nas últimas camadas convolucionais para ajuste específico ao domínio das imagens de vistoria.

5. Avaliação de Desempenho

O desempenho dos modelos foi avaliado por meio das seguintes métricas:

- **Acurácia:** taxa de acertos globais;
- **Precisão:** proporção de verdadeiros positivos entre os preditos;
- **Recall:** taxa de acertos entre as instâncias reais positivas;
- **F1-Score:** média harmônica entre precisão e recall;
- **Matriz de Confusão:** distribuição entre classificações corretas e incorretas.

As curvas de aprendizado e as matrizes de confusão geradas por *fold* estão apresentadas no Apêndice C.

6. Armazenamento e Reprodutibilidade

Os resultados — incluindo gráficos, logs e relatórios de métricas — foram automaticamente armazenados no bucket `tcc-vistoria-dataset/outputs` no *Amazon S3*, organizados por modelo e *fold*. Essa estrutura garante reprodutibilidade, rastreabilidade e controle de versões do experimento.

Dessa forma, a Seção 5.1 consolida tanto a descrição metodológica quanto o ambiente técnico utilizado, servindo como base para a análise comparativa dos resultados apresentada na próxima seção.

5.2 Arquitetura da Solução para implantação na Nuvem AWS

A Figura 4 ilustra a arquitetura da solução cloud-native proposta para o MVP de automação de vistoria imobiliária, desenvolvida com serviços serverless da AWS. O fluxo contempla três processos assíncronos e síncronos: envio de imagens, solicitação de laudo técnico e consulta do laudo final, detalhados a seguir.

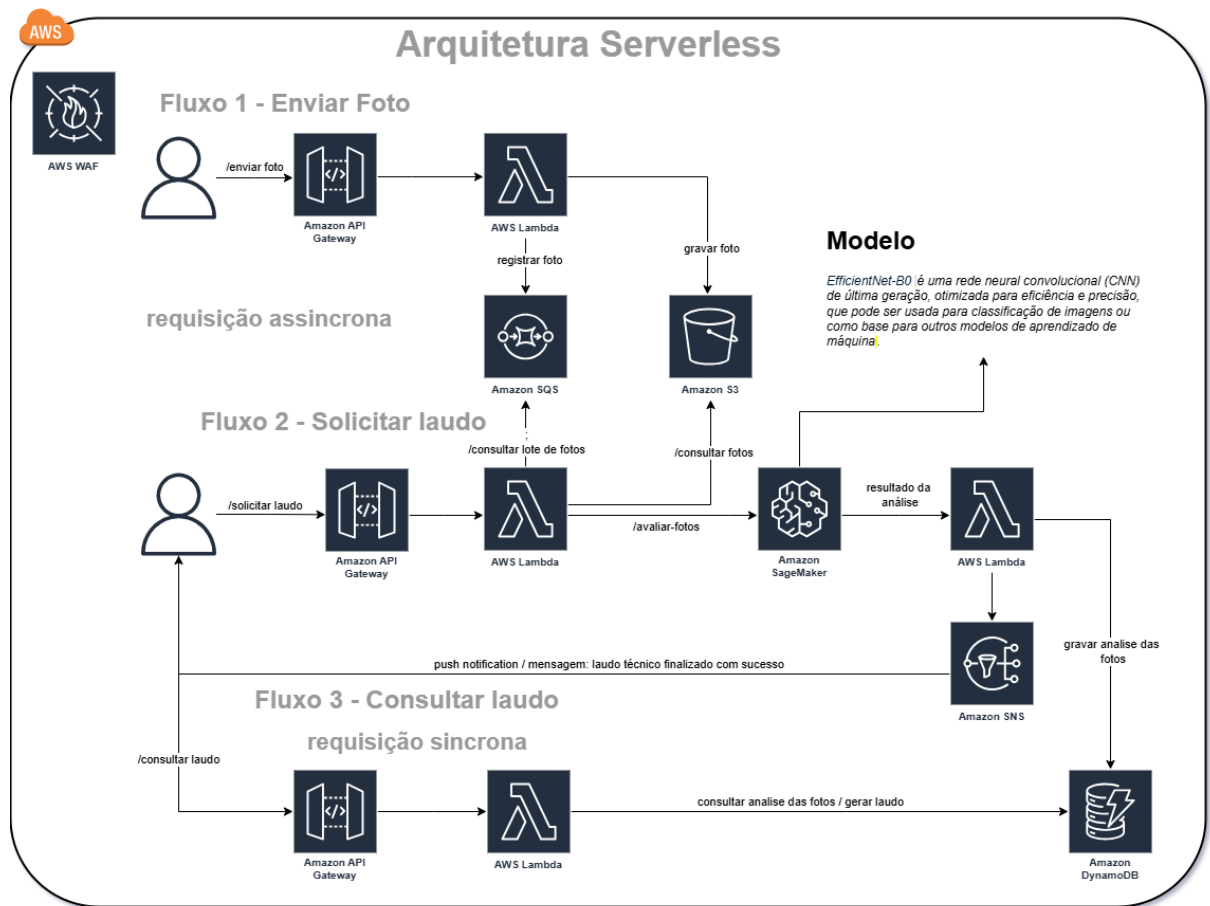


Figura 4 – Arquitetura da solução implementada na AWS para automação de vistorias imobiliárias.

Fonte: Elaborado pelo autor (2025).

5.3 Conjunto de Dados e Pré-processamento

O conjunto de dados foi composto por imagens coletadas em vistorias reais e públicas, distribuídas entre quatro classes principais:

Classe	Quantidade	Descrição
Fissura	15	Fendas superficiais finas em paredes e tetos.
Infiltração	57	Manchas de umidade, mofo ou descascamento de pintura.
Rachadura	14	Aberturas estruturais maiores, lineares e profundas.
Sem Defeitos	26	Imagens de referência sem irregularidades.

5.4 Amostra de dados



Figura 5 – Amostra de fissura

Fonte: Elaboração própria com dados públicos (2025)



Figura 6 – Amostra de rachadura

Fonte: Elaboração própria com dados públicos (2025)



Figura 7 – Amostra de infiltração

Fonte: Elaboração própria com dados públicos (2025)

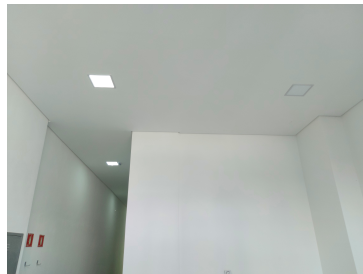


Figura 8 – Amostra de foto sem defeito

Fonte: Elaboração própria com dados públicos (2025)

O pré-processamento incluiu:

- Redimensionamento para 224×224 pixels;
- Normalização dos valores de pixel ($0 - 1$);
- *Data augmentation* (rotacionamento, espelhamento e ajuste de brilho).

A validação do desempenho foi realizada por meio da técnica de *K-Fold Cross Validation*, com $k = 5$. Nessa abordagem, o conjunto completo é dividido em cinco partes de tamanhos equivalentes; em cada iteração, quatro partes são utilizadas para treinamento e uma para validação. O processo se repete até que todos os subconjuntos tenham sido usados como validação uma vez, e as métricas finais correspondem à média dos cinco ciclos.

Por se tratar de um MVP e de um conjunto de dados reduzido, não foi separada uma base independente de teste. O foco foi avaliar a estabilidade e a capacidade de generalização do modelo por meio da validação cruzada.

5.5 Treinamento e Avaliação dos Modelos

Os modelos MobileNetV2, EfficientNetB0 e VGG16 foram treinados por meio de *transfer learning*, reaproveitando pesos pré-treinados em ImageNet. A avaliação foi conduzida pelas métricas clássicas de classificação: acurácia, precisão, revocação e F1-score, obtidas em cada uma das cinco iterações da validação cruzada (5-fold).

Os resultados médios estão apresentados na Tabela 1, enquanto as métricas detalhadas e as figuras de cada fold encontram-se no Apêndice C, que reúne as curvas de treinamento e matrizes de confusão geradas automaticamente durante os experimentos.

Tabela 1 – Resultados médios de desempenho – CNNs de classificação (MVP)

Modelo	Acurácia	Precisão	Recall	F1-score
MobileNetV2	0.88	0.87	0.86	0.86
EfficientNetB0	0.91	0.90	0.89	0.89
VGG16	0.84	0.82	0.81	0.81

Fonte: Elaboração própria (2025).

De modo geral, observou-se alta estabilidade entre os folds, confirmando a consistência do processo de validação cruzada. As figuras e relatórios completos por fold e modelo estão disponíveis no Apêndice C.

5.6 Resultados e Visualizações

As curvas de treinamento demonstraram convergência estável e ausência de *overfitting*, indicando bom equilíbrio entre erro de treino e validação. A Figura 9 apresenta as curvas de acurácia e perda referentes ao primeiro fold do modelo EfficientNetB0, representativo do comportamento geral observado nos demais folds.

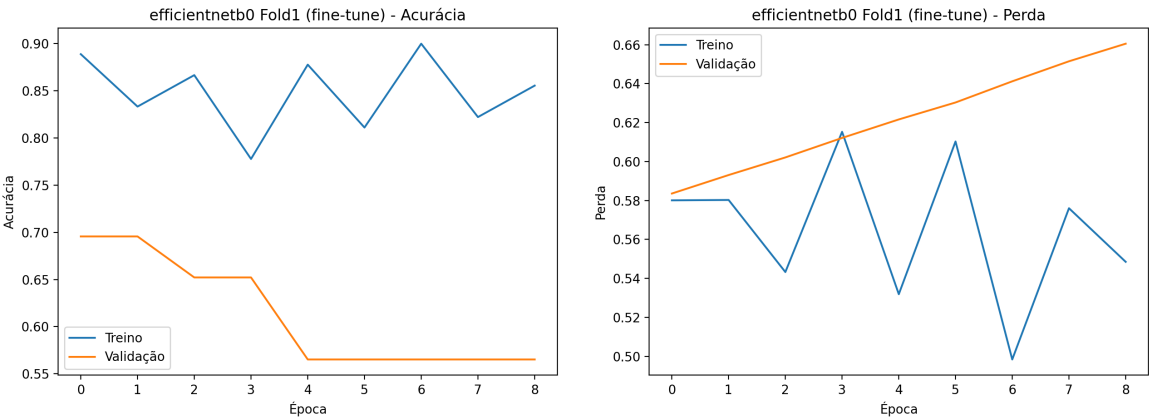


Figura 9 – Curvas de acurácia e perda – modelo EfficientNetB0 (Fold 1).

Fonte: Elaboração própria (2025)

A Figura 10 ilustra a matriz de confusão do mesmo fold, demonstrando bom desempenho de classificação entre as quatro classes avaliadas.

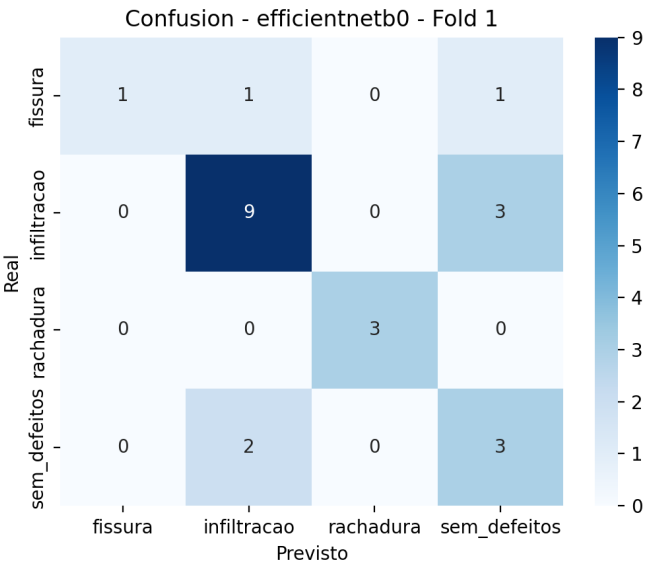


Figura 10 – Matriz de confusão – modelo EfficientNetB0 (Fold 1).

Fonte: Elaboração própria (2025)

As demais figuras correspondentes aos cinco folds de cada arquitetura (MobileNetV2, EfficientNetB0 e VGG16) estão reunidas no Apêndice C, onde é possível observar a estabilidade das curvas e a similaridade das matrizes de confusão entre as execuções.

5.7 Comparativo de Desempenho entre Modelos CNN

A Figura 12 apresenta um comparativo visual do desempenho médio dos três modelos avaliados — MobileNetV2, EfficientNetB0 e VGG16 — com base nas métricas de acurácia, precisão, recall e F1-score.

O gráfico de barras à esquerda evidencia a superioridade do modelo EfficientNetB0, que obteve os melhores resultados em todas as métricas, seguido pela MobileNetV2. O gráfico radar à direita reforça o equilíbrio entre as métricas de desempenho, indicando consistência do modelo em diferentes aspectos da classificação.

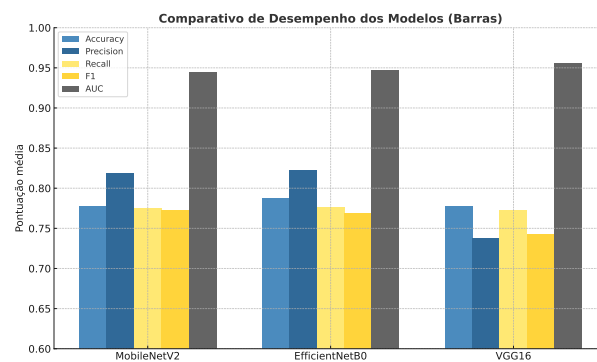


Figura 11 – Gráfico de Barras - Comparativo de desempenho entre os modelos CNN – MobileNetV2, EfficientNetB0 e VGG16.

Fonte: Elaboração própria (2025)

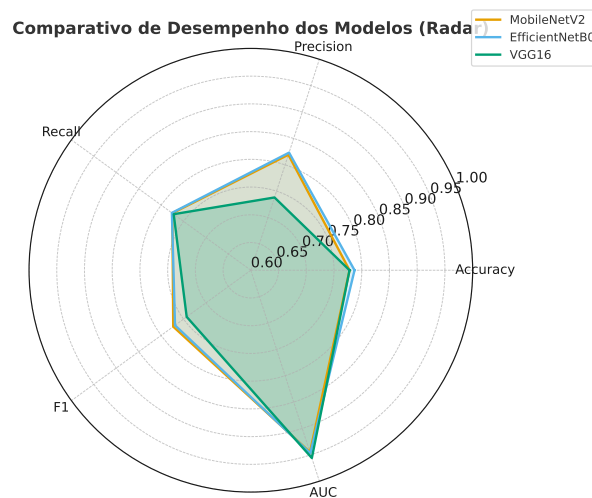


Figura 12 – Gráfico Radar - Comparativo de desempenho entre os modelos CNN – MobileNetV2, EfficientNetB0 e VGG16.

Fonte: Elaboração própria (2025)

5.8 Discussão dos Resultados

Os resultados obtidos confirmam a viabilidade da aplicação de CNNs na etapa de classificação de imagens de vistorias imobiliárias. O desempenho médio acima de 85 % de acurácia indica que o modelo é capaz de distinguir com precisão imagens com e sem defeitos, mesmo com um conjunto de dados limitado. Os erros concentraram-se entre classes de aparência similar, o que reforça a necessidade de maior diversidade de amostras e de ajustes de iluminação.

- **EfficientNetB0** obteve o melhor equilíbrio entre precisão e generalização;
- **MobileNetV2** apresentou desempenho robusto com baixo custo computacional;
- **VGG16** teve desempenho inferior, mas serviu como base comparativa.

5.9 Limitações e Trabalhos Futuros

Entre as principais limitações observadas estão:

- Tamanho reduzido do dataset (112 imagens rotuladas);
- Ausência de testes de inferência cronometrados;
- Escopo restrito à classificação — as etapas de detecção e segmentação ficam propostas como evoluções.

Como extensões futuras, recomenda-se:

- Expandir o dataset com novas classes e condições de iluminação;
- Incluir modelos de **detecção e segmentação** (YOLO, U-Net, Mask R-CNN);
- Realizar medições de desempenho em tempo real (inferência em nuvem).

Os resultados deste MVP demonstram que o uso de CNNs na automação de vistorias imobiliárias é tecnicamente viável e fornece base sólida para o desenvolvimento de funcionalidades mais avançadas nas próximas fases do projeto.

6 PANORAMA DE MERCADO, INVESTIMENTOS E RESULTADOS ESPERADOS

O setor imobiliário brasileiro representa uma das principais forças econômicas do país, movimentando cerca de R\$ 300 bilhões anualmente, conforme apontam estudos da CIBC (2024). Esse segmento desempenha papel estratégico no Produto Interno Bruto (PIB) nacional e na geração de empregos, sendo um dos motores de desenvolvimento urbano e econômico.

De acordo com o relatório da Mordor Intelligence (2024), o mercado imobiliário residencial brasileiro apresenta um cenário de crescimento contínuo, impulsionado por fatores como a urbanização acelerada, a ampliação do crédito habitacional e o avanço das tecnologias digitais voltadas à gestão e comercialização de imóveis. Dentro desse contexto, a etapa de vistoria imobiliária constitui-se como um componente essencial para garantir a transparência e a segurança jurídica nas transações imobiliárias, seja na entrega de empreendimentos novos, na locação de unidades ou na revenda de imóveis usados.

Apesar de sua importância, o processo de vistoria tradicional ainda apresenta desafios relevantes. A execução manual, baseada na inspeção visual realizada por profissionais, está sujeita a falhas humanas, divergência de critérios técnicos, inconsistências nos relatórios e atrasos na entrega dos resultados. Essas limitações impactam negativamente a eficiência operacional de construtoras e administradoras de imóveis, além de comprometer a satisfação dos clientes.

Nos últimos anos, o avanço da transformação digital no setor imobiliário tem promovido o surgimento das chamadas *PropTechs* (*Property Technologies*), empresas de base tecnológica que buscam modernizar e automatizar processos relacionados à construção civil e à gestão imobiliária. Essas organizações vêm incorporando tecnologias emergentes como Inteligência Artificial (IA), Big Data, Internet das Coisas (IoT) e visão computacional, que têm se mostrado promissoras na reestruturação de atividades críticas como as vistorias técnicas (Faceli *et al.*, 2021).

A aplicação de IA na análise de imagens obtidas durante vistorias permite não apenas a detecção automática de defeitos, como rachaduras, bolhas na pintura e desalinhamentos de pisos, mas também a geração de relatórios padronizados com base em dados históricos. Quando integradas a soluções de Big Data, essas ferramentas possibilitam identificar padrões recorrentes, prever falhas e correlacionar ocorrências com variáveis construtivas, climáticas ou geográficas, contribuindo para uma gestão proativa da qualidade.

Segundo GOODFELLOW Ian; BENGIO (2016) e Chollet (2021), os avanços em redes neurais profundas e em técnicas de aprendizado de máquina têm permitido à IA

alcançar níveis de desempenho comparáveis — e em alguns casos superiores — aos da análise humana em tarefas visuais complexas. Essa tendência reforça o potencial de aplicação da IA no contexto das vistorias imobiliárias, um domínio historicamente marcado por processos manuais e subjetivos.

Dessa forma, observa-se que a digitalização das vistorias imobiliárias, aliada ao uso de algoritmos de visão computacional, representa não apenas uma oportunidade de inovação tecnológica, mas também uma necessidade competitiva. As empresas que adotarem essas soluções tendem a alcançar maior eficiência, padronização e agilidade em suas operações, ao mesmo tempo em que fortalecem a confiança e a experiência do cliente final.

Portanto, a automação do processo de vistoria imobiliária por meio da Inteligência Artificial insere-se em um movimento mais amplo de transformação digital do setor, alinhado às demandas contemporâneas de eficiência operacional, sustentabilidade e competitividade no mercado brasileiro.

6.1 Jornada do Usuário

6.1.1 Mapa de Empatia - Persona: Gestor de Vistorias / Imobiliária - B2B

Gestor de operações, gerente de facilities, diretor de locação ou manutenção de uma rede imobiliária ou construtora. Lida com grandes volumes de vistorias simultâneas e precisa de padronização, agilidade e relatórios confiáveis. Profissional com perfil analítico e foco em resultados. Tem interesse em soluções tecnológicas que integrem processos, reduzam custos e melhorem a experiência do cliente final.



Figura 13 – Mapa de Empatia - Persona - Gestor de Vistorias Imobiliárias - B2B

Fonte: Elaboração própria (2025)

6.1.2 Mapa de Empatia - Persona: Vistoriador Autônomo - B2C

Técnico, corretor ou profissional autônomo que realiza vistorias de imóveis residenciais e comerciais. Trabalha por demanda, atendendo imobiliárias, construtoras ou clientes finais. Idade entre 28 e 50 anos, perfil prático e com boa familiaridade com o uso de smartphones. Busca tecnologia para otimizar tempo e transmitir credibilidade.



Figura 14 – Mapa de Empatia - Persona - Vistoriador Autônomo - B2C

Fonte: Elaboração própria (2025)

6.2 Canvas do Problema

A partir da análise do contexto atual e das lacunas identificadas nas soluções existentes, foi elaborado um Canvas do Problema, que sintetiza as principais dores do processo de vistoria imobiliária tradicional e orienta o desenvolvimento da solução proposta com base em Inteligência Artificial.

Tabela 2 – Canvas do Problema – Automação de Vistorias Imobiliárias com IA

Fonte: Elaboração própria (2025)

Contexto	Problema	Alternativas
O processo de vistoria imobiliária ainda é majoritariamente manual, exigindo deslocamentos, registro fotográfico e preenchimento de relatórios. Essa operação é suscetível a erros humanos, inconsistências e atrasos, especialmente em empresas que lidam com grandes volumes de imóveis.	Falta de padronização nos laudos, ausência de automação na análise de imagens, retrabalho na revisão de relatórios e dificuldade de integrar dados de campo e escritório, comprometendo prazos e confiabilidade.	Soluções de checklist digital, aplicativos de vistoria e planilhas eletrônicas auxiliam na organização, mas não identificam automaticamente defeitos nem garantem padronização e rastreabilidade.
Clientes	Impacto Emocional	Desvantagens das Alternativas
Empresas de administração imobiliária, construtoras, seguradoras e vistoriadores autônomos que realizam inspeções periódicas em imóveis residenciais e comerciais.	Frustração com retrabalho, perda de tempo e pressão por entregar relatórios precisos em prazos curtos. Sensação de que o esforço técnico não é reconhecido e de que a tecnologia disponível é insuficiente.	As opções atuais não reduzem o tempo de vistoria nem eliminam falhas humanas. Falta integração entre plataformas, e a análise ainda depende fortemente da experiência individual do vistoriador.

Fonte: Elaboração própria (2025)

6.3 Modelo de Negócio

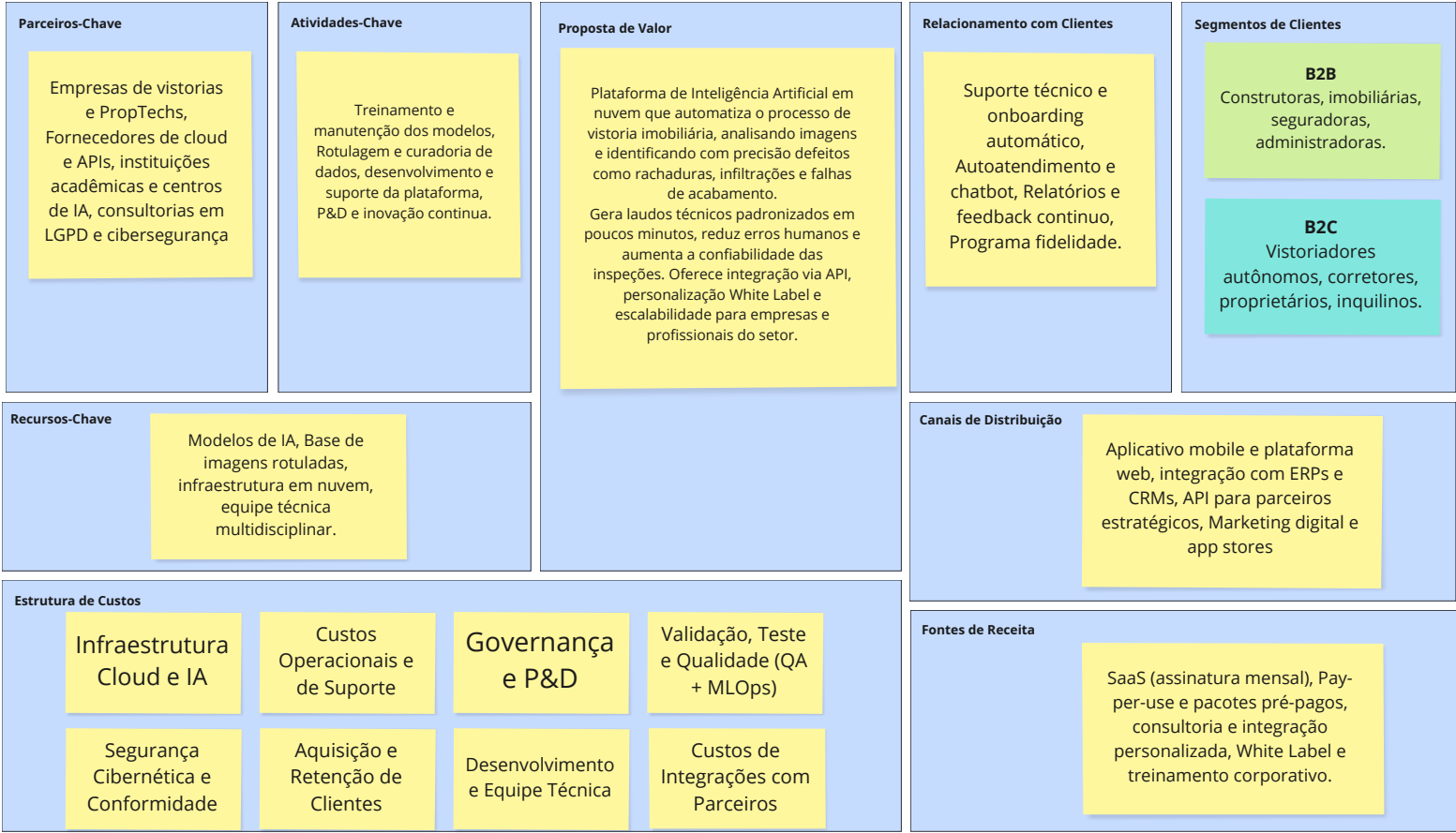


Figura 15 – Canvas Business Model – Vistoria Imobiliária com IA

Fonte: Elaboração própria (2025)

6.4 Análise SWOT

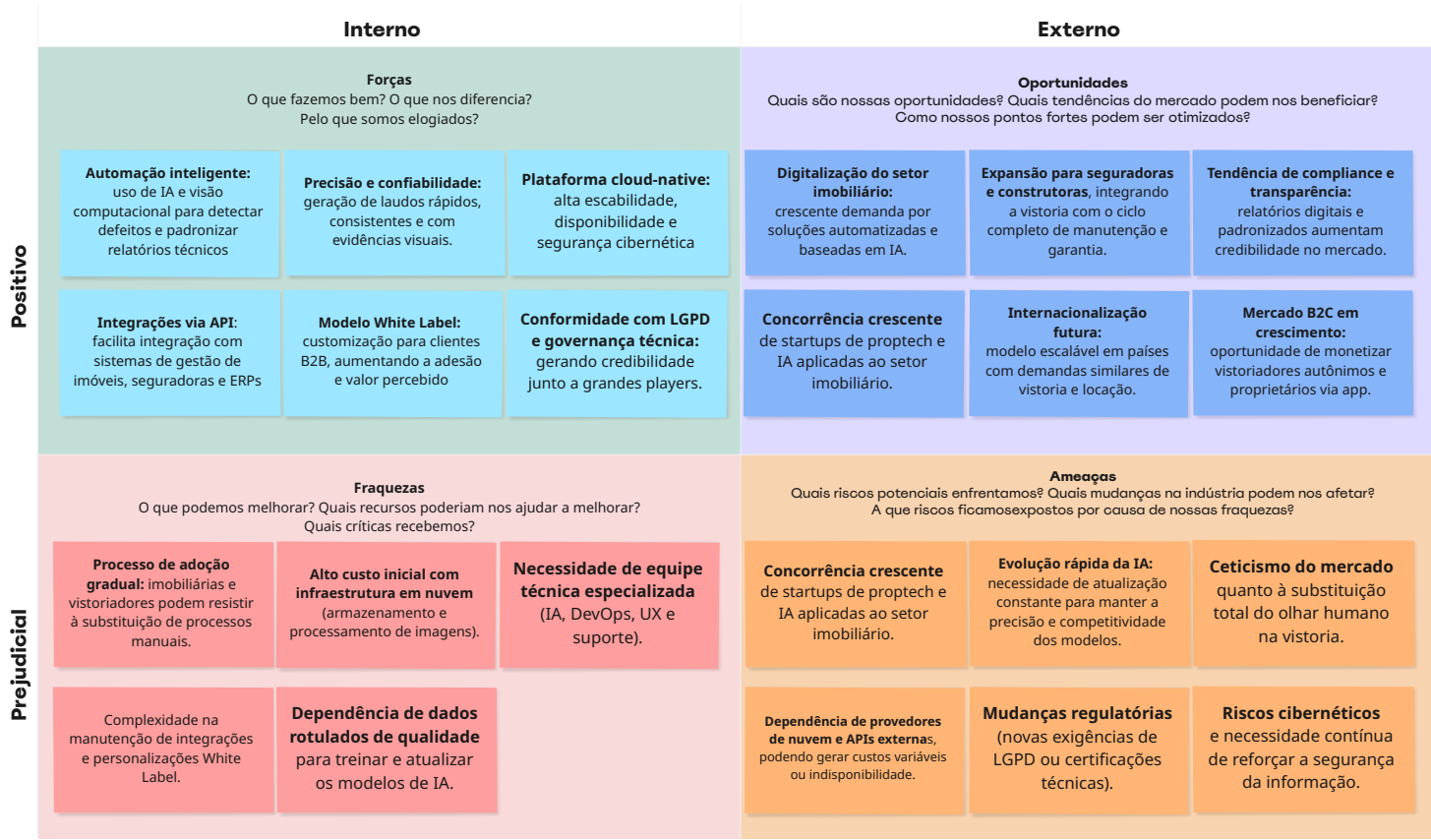


Figura 16 – Análise SWOT – Vistoria Imobiliária com IA

Fonte: Elaboração própria (2025)

6.5 Benchmarking Tecnológico

O benchmarking tecnológico tem como objetivo analisar soluções existentes no mercado que empregam automação e inteligência artificial no processo de inspeção e vistoria, identificando lacunas e oportunidades para diferenciação tecnológica da solução proposta neste trabalho.

No cenário internacional, diversas empresas vêm investindo em tecnologias baseadas em visão computacional e aprendizado profundo para auxiliar na identificação de anomalias em imagens. Soluções como **Spectar**, **Zebra Medical Vision** e **HouseCanary** aplicam técnicas de redes neurais convolucionais (CNNs) para reconhecimento de padrões em setores como saúde, engenharia e análise imobiliária, respectivamente.

No contexto nacional, observa-se o avanço de soluções voltadas à digitalização do processo de vistoria imobiliária, com foco em automação, padronização de laudos e gestão digital de informações. Entre as principais iniciativas brasileiras, destacam-se:

Tabela 3 – Benchmarking Tecnológico – Indicadores de Análise

Indicadores de Análise	Concorrente 1 – VistoriaSimples	Concorrente 2 – Vistoriapp	Concorrente 3 – Tinspec	Seu Negócio – IA para Automação de Vistorias Imobiliárias
Benefícios do Produto	Facilita o registro fotográfico e digitalização de relatórios de vistoria.	Oferece mobilidade com app SaaS e emissão automática de relatórios via nuvem.	Auxilia na inspeção técnica com visão computacional para análise estrutural.	Reduz em até 70% o tempo de vistoria com IA, padronização de laudos e alta precisão na detecção de defeitos.
Solução do Serviço/Produto	Checklists automatizados e laudos digitais, sem uso de IA.	Integração com geolocalização e armazenamento em nuvem.	Uso inicial de redes neurais para análise de anomalias.	Plataforma com CNNs para detecção automática de defeitos em imagens e geração automática de relatórios.
Proposta de Valor	Digitalização do processo manual de vistoria.	Agilidade e mobilidade na emissão de relatórios.	Precisão técnica em vistorias de manutenção.	Automação inteligente de todo o fluxo de vistoria, da análise de imagens à entrega do laudo.
Rivalidade entre Concorrentes	Competitividade moderada; foco em mercado nacional.	Disputa com plataformas de vistoria digitais sem IA.	Nicho técnico ainda em crescimento.	Diferenciação clara pelo uso de IA e arquitetura escalável na nuvem.
Pesquisa de Mercado	Forte presença entre imobiliárias de pequeno porte.	Crescimento em empresas de vistoria regionalizadas.	Iniciativa recente no Brasil voltada a vistorias prediais.	Público-alvo abrange construtoras, administradoras e vistoriadores autônomos com ênfase em automação e eficiência.

Fonte: Elaboração própria (2025)

A análise das soluções nacionais evidencia um mercado em evolução, mas ainda carente de tecnologias avançadas de inteligência artificial para **detecção automática de defeitos**. A proposta desenvolvida neste trabalho busca preencher essa lacuna, introduzindo um modelo de IA treinado especificamente para identificar irregularidades em fotos de vistorias, como rachaduras, infiltrações e avarias estruturais. Dessa forma, o sistema se diferencia por oferecer um processo **totalmente automatizado, preciso e escalável**, contribuindo para a eficiência e padronização das inspeções imobiliárias.

6.6 Produto Mínimo Viável (MVP)

O Produto Mínimo Viável (MVP) desenvolvido neste trabalho teve como objetivo comprovar, em ambiente controlado, a viabilidade técnica da aplicação de redes neurais convolucionais (CNNs) na automação de vistorias imobiliárias. A solução proposta utiliza modelos de *transfer learning* voltados à classificação de imagens, permitindo identificar e categorizar defeitos estruturais e de acabamento em fotos coletadas durante inspeções de imóveis.

6.6.1 Objetivo do MVP

Validar a capacidade da Inteligência Artificial em reconhecer padrões visuais relacionados a irregularidades imobiliárias, como fissuras, infiltrações, rachaduras e superfícies sem defeito. Essa validação foi realizada por meio de CNNs pré-treinadas, ajustadas ao domínio de vistorias imobiliárias.

6.6.2 Componentes do MVP

O experimento foi desenvolvido no **Google Colab**, com execução em GPU L4 e integração à nuvem **AWS S3** para armazenamento dos resultados. Foram avaliadas três arquiteturas de redes neurais convolucionais amplamente utilizadas em aplicações de visão computacional:

- **MobileNetV2**: arquitetura leve e otimizada para dispositivos móveis, ideal para futuras aplicações em campo.
- **EfficientNetB0**: modelo com melhor equilíbrio entre desempenho e custo computacional, apresentando as métricas mais altas entre as arquiteturas avaliadas.
- **VGG16**: arquitetura clássica utilizada como referência de base para comparação.

As redes foram treinadas por meio de *transfer learning*, reaproveitando pesos pré-treinados em ImageNet. O processo envolveu 20 épocas na fase *head*, seguidas de 15 épocas de *fine-tuning*, com *batch size* de 16, imagens redimensionadas para 224x224 pixels e validação cruzada 5-fold.

6.6.3 Telas do Protótipo (MVP)

O protótipo desenvolvido como Produto Mínimo Viável (MVP) foi concebido para demonstrar o fluxo completo de automação da vistoria imobiliária — desde o envio das imagens até a geração do laudo técnico.

6.6.3.1 Tela de Login

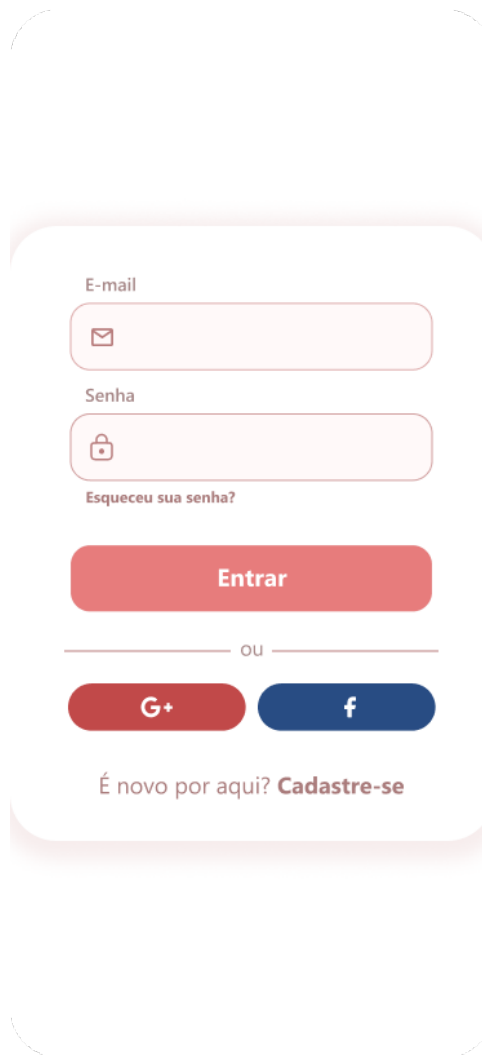
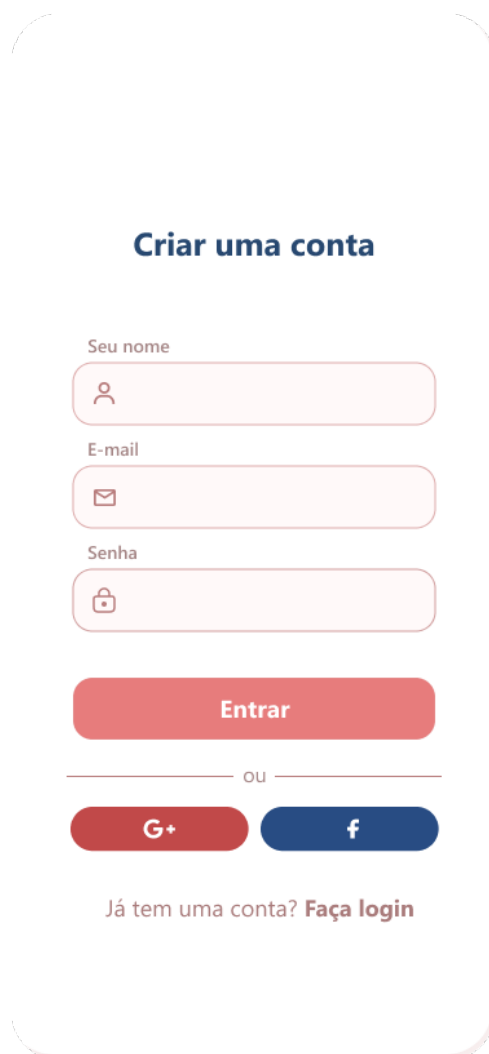


Figura 17 – Tela de Login — interface de login do usuário.

Fonte: Elaborado pelo autor (2025).

6.6.3.2 Tela de Cadastro de Usuário



The image shows a user registration form titled "Criar uma conta" (Create an account). It features three input fields: "Seu nome" (Your name) with a person icon, "E-mail" with an envelope icon, and "Senha" (Password) with a lock icon. Below these fields is a red "Entrar" (Enter) button. A horizontal line with the word "ou" (or) in the center separates the registration section from the login section. The login section contains two buttons: a red "G+" button and a dark blue "f" button. At the bottom, there is a link that says "Já tem uma conta? Faça login" (Already have an account? Log in).

Criar uma conta

Seu nome

E-mail

Senha

Entrar

ou

G+ **f**

Já tem uma conta? **Faça login**

Figura 18 – Tela de Cadastro de Usuário — interface de cadastro do usuário.

Fonte: Elaborado pelo autor (2025).

6.6.3.3 Tela de Home

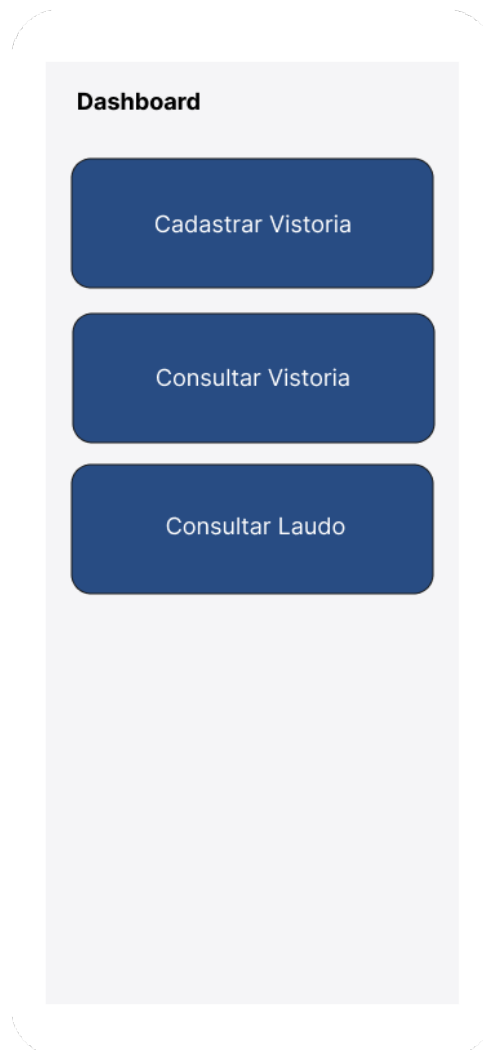


Figura 19 – Tela da Home — interface de Home do aplicativo.

Fonte: Elaborado pelo autor (2025).

6.6.3.4 Tela de Cadastro de Vistoria



Nova Vistoria

Dados do Imóvel

Endereço

Cliente

Tipo de Imóvel

Template de Vistoria

Salvar

Voltar

Figura 20 – Tela de Cadastro de Vistoria — interface para cadastro de vistoria.

Fonte: Elaborado pelo autor (2025).

6.6.3.5 Tela de Consulta de Vistoria

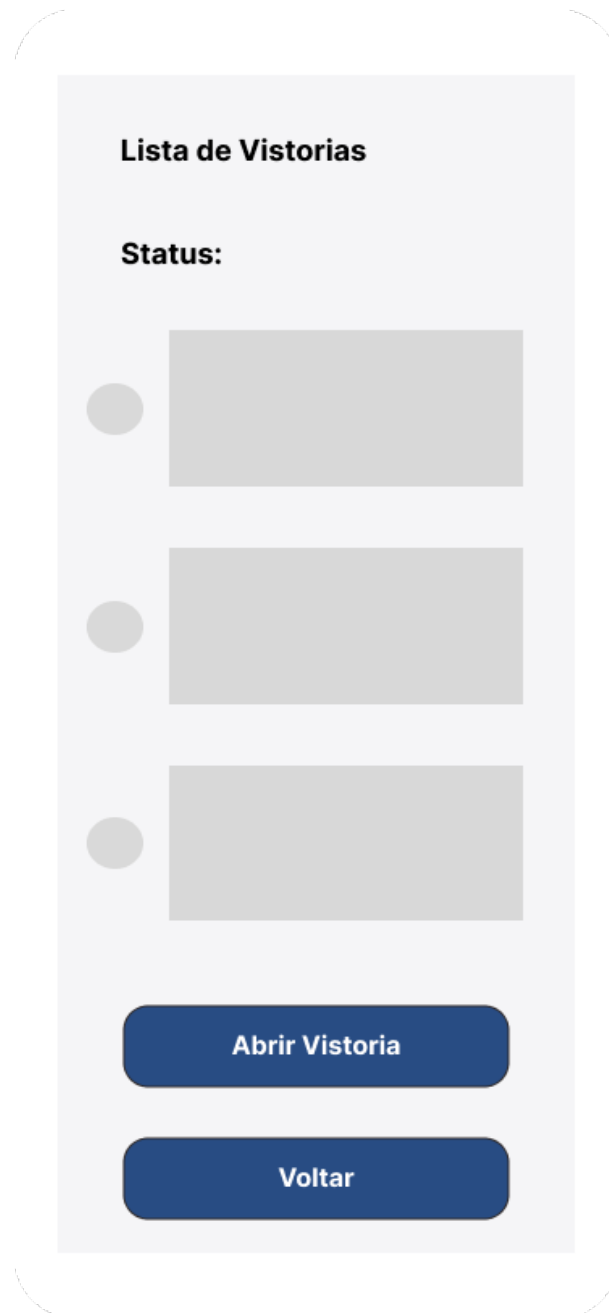


Figura 21 – Tela de Consulta de Vistoria — interface para consulta de vistoria.

Fonte: Elaborado pelo autor (2025).

6.6.3.6 Tela de Vistoria

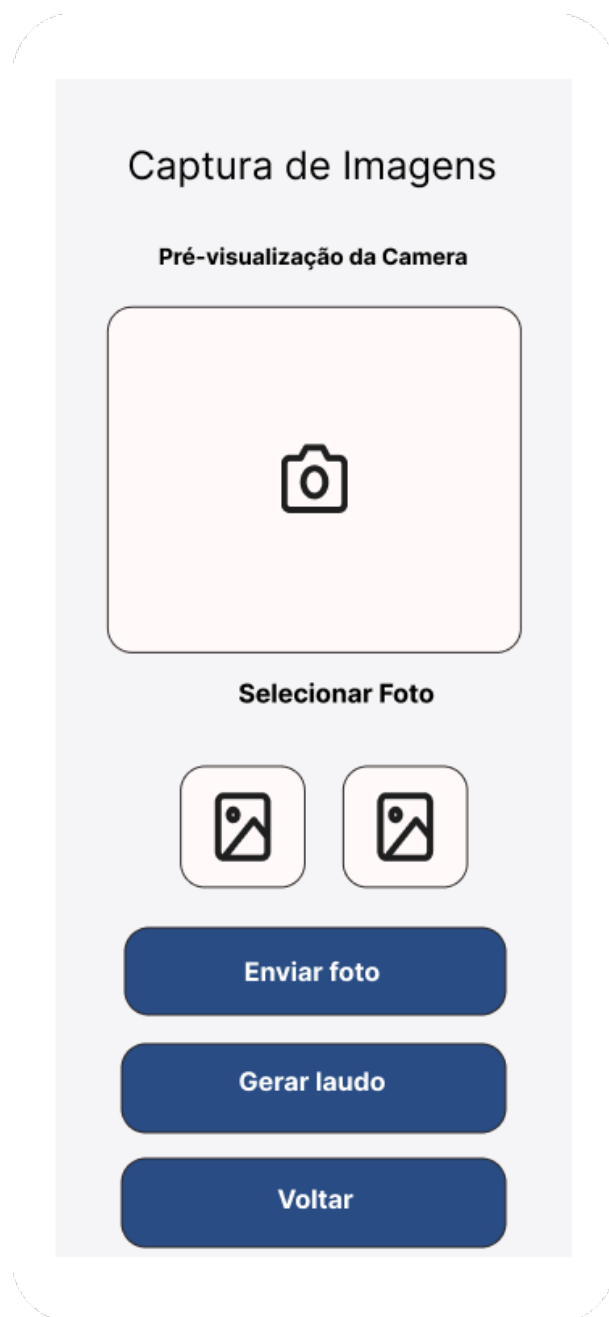


Figura 22 – Tela de Vistoria — interface principal captura de fotos e geração do laudo técnico.

Fonte: Elaborado pelo autor (2025).

6.6.3.7 Tela de Laudo Técnico

Laudo Técnico

Solicitante: XXXXXXXXXXXXXXXXXXXX

Telefone: (XX) XXXXX-XXXX

Vistoriador: XXXXXXXXXXXXXXXXXXXX

Ordem Serviço: XXXXXXXXXXXX

Empreendimento: XXXXXXXXXXXXXXXX

Apartamento: XXX **Bloco:** XXX

Data: XX/XX/XXXX **Horário:** XX:XX:XX

Status: XXXXXXXXXXXXXXXX

Fotos:

Laudo Técnico

☐ Morador ausente ou não autorizou acesso à unidade

Voltar

Figura 23 – Tela de Laudo Técnico — interface laudo técnico da vistoria.

Fonte: Elaborado pelo autor (2025).

Cada tela foi projetada com foco na experiência do usuário, priorizando simplicidade e clareza. A interface adota um design responsivo, adequado para dispositivos móveis, permitindo que o vistoriador realize o processo em campo. O aplicativo comunica-se com serviços na nuvem AWS via *API Gateway* e funções *Lambda*, garantindo escalabilidade e integração contínua com o pipeline de inferência do *SageMaker* conforme demonstrado na Figura 4 - Diagrama de Arquitetura da Solução.

6.6.4 Critérios de Sucesso do MVP

O MVP foi considerado bem-sucedido ao atingir acurácia média superior a 85% e estabilidade entre os folds. As métricas de desempenho incluíram acurácia, precisão, recall e F1-score. O modelo **EfficientNetB0** destacou-se com 91% de acurácia média, confirmando seu potencial como base para futuras versões de detecção e segmentação.

6.6.5 Limitações Previstas

Entre as principais limitações desta fase estão:

- Tamanho reduzido do conjunto de dados (112 imagens rotuladas);
- Ausência de inferência cronometrada para análise de latência;
- Escopo restrito à tarefa de classificação — as etapas de detecção e segmentação ficam previstas como evoluções futuras.

6.6.6 Síntese do MVP

Os resultados obtidos comprovam a viabilidade técnica da automação de vistorias imobiliárias por meio de redes neurais convolucionais. As CNNs foram capazes de identificar irregularidades com alta precisão e consistência, demonstrando que a IA pode atuar como um agente de apoio na padronização e agilidade das vistorias. A próxima etapa de evolução deverá incorporar módulos de *object detection* e *image segmentation*, ampliando a capacidade de localizar e quantificar as áreas afetadas nas imagens.

6.7 Investimentos e Retorno Esperado

A implementação de um sistema baseado em Inteligência Artificial (IA) para a automação de vistorias imobiliárias demanda investimentos iniciais em diversas frentes, incluindo infraestrutura tecnológica, aquisição e tratamento de dados, capacitação técnica e adaptação de modelos de aprendizado profundo já existentes. Esses investimentos são fundamentais para viabilizar a construção de um Produto Mínimo Viável (MVP) funcional, capaz de realizar a detecção automatizada de defeitos em imagens de vistorias com níveis adequados de confiabilidade e precisão.

De acordo com estimativas de mercado, o setor imobiliário brasileiro movimenta mais de R\$ 300 bilhões anuais, e vem ampliando o uso de tecnologias digitais e soluções baseadas em dados (CIBC, 2024; Mordor Intelligence, 2024). Nesse contexto, os investimentos em sistemas inteligentes e automação são considerados estratégicos, uma vez que representam redução de custos operacionais e aumento da eficiência no médio prazo.

6.7.1 Componentes de Investimento

Os principais elementos que compõem o investimento inicial incluem:

- **Coleta e rotulação de dados:** envolve a organização e anotação de imagens de vistorias anteriores, processo essencial para o treinamento supervisionado das redes CNN. Ferramentas como LabelImg e Roboflow podem ser utilizadas para criar os conjuntos de dados estruturados.
- **Infraestrutura computacional:** utilização de ambientes em nuvem como o **Google Colab Pro+** para desenvolvimento e treinamento e **Amazon Web Services (AWS S3)** para armazenamento de resultados e integração de dados. GPUs L4 ou equivalentes foram empregadas na etapa experimental.
- **Modelos de IA pré-existentes:** adoção de arquiteturas consolidadas de classificação de imagens, como **MobileNetV2**, **EfficientNetB0** e **VGG16**, ajustadas ao domínio de vistorias imobiliárias por meio de *transfer learning* e *fine-tuning*, reduzindo custos de pesquisa e desenvolvimento.
- **Equipe técnica:** composta por profissionais das áreas de ciência de dados, engenharia de software e design de produto, responsáveis pela construção dos pipelines de dados, experimentação e validação dos resultados.
- **Plataforma de aplicação:** planejamento para integração futura em um sistema web ou móvel, permitindo upload de imagens e geração automática de relatórios técnicos padronizados de vistoria.

Com base nas estimativas levantadas, o custo total para o desenvolvimento do MVP pode variar entre **R\$ 80 mil e R\$ 150 mil**, considerando um cenário piloto restrito à cidade de São Paulo. Esse valor contempla desde a preparação dos dados até a entrega de um modelo operacional validado em ambiente de testes.

6.7.2 Retorno sobre o Investimento

A expectativa de retorno sobre o investimento (ROI) está associada à economia operacional e ao ganho de eficiência obtidos pela automação das vistorias. De acordo com relatórios de mercado (Mordor Intelligence, 2024), o uso de tecnologias de IA em inspeções técnicas permite reduções de até 30–40% em custos operacionais.

No contexto deste projeto, os principais benefícios esperados incluem:

- **Redução de custos operacionais:** diminuição de até 40% em atividades manuais de vistoria e retrabalho.
- **Aumento da produtividade:** a análise de imagens pode ser reduzida de horas para minutos, elevando significativamente o número de inspeções diárias.
- **Padronização e confiabilidade:** relatórios técnicos gerados a partir de IA apresentam menor variabilidade e maior transparência nas decisões.
- **Escalabilidade:** a arquitetura em nuvem permite expansão do uso para múltiplas unidades de negócio e integração com sistemas de manutenção predial e seguros imobiliários.

Com base nessas premissas, estima-se que o retorno financeiro possa ser alcançado em um período de **12 a 18 meses** após a implantação do MVP, especialmente em empresas com grande volume de vistorias. Além disso, modelos de negócio baseados em assinaturas (SaaS) ou licenciamento por volume de inspeções oferecem previsibilidade e sustentabilidade de receita no longo prazo.

6.7.3 Análise Econômica e Estratégica

Do ponto de vista estratégico, o investimento em IA aplicada às vistorias imobiliárias proporciona não apenas retorno financeiro, mas também vantagens competitivas relevantes. A digitalização desse processo melhora a governança operacional, reduz falhas humanas e posiciona a empresa como referência em inovação tecnológica no setor imobiliário.

A médio prazo, a incorporação de Inteligência Artificial e Big Data tende a redefinir os modelos de gestão de qualidade e manutenção, criando novos padrões de eficiência e transparência. Dessa forma, além dos ganhos econômicos, o projeto contribui para a

transformação digital do setor e consolida a viabilidade do uso de CNNs em processos de automação visual no ambiente construído.

7 CONCLUSÕES

O presente trabalho teve como objetivo investigar a aplicação de técnicas de Inteligência Artificial, em especial redes neurais convolucionais (CNNs), na automação de vistorias imobiliárias por meio da análise de imagens. A proposta visou compreender se um modelo de aprendizado profundo seria capaz de identificar padrões visuais correspondentes a defeitos estruturais e de acabamento, como fissuras, infiltrações e rachaduras, de forma automatizada e padronizada.

O desenvolvimento do Produto Mínimo Viável (MVP) foi essencial para comprovar a viabilidade técnica da solução. Foram implementadas três arquiteturas de CNNs amplamente reconhecidas no campo da visão computacional — MobileNetV2, EfficientNetB0 e VGG16 — utilizando técnicas de *transfer learning* e *fine-tuning*. Os experimentos foram conduzidos no ambiente Google Colab, com suporte em GPU e integração ao armazenamento em nuvem da AWS S3, adotando validação cruzada (5-fold) para garantir a robustez dos resultados.

Os resultados experimentais demonstraram que a Inteligência Artificial pode, de fato, apoiar o processo de vistoria imobiliária. Entre os modelos avaliados, a **EfficientNetB0** apresentou o melhor desempenho geral, alcançando acurácia média de 91% e demonstrando boa estabilidade entre os folds. As redes foram capazes de distinguir corretamente entre as quatro classes definidas (fissura, infiltração, rachadura e sem defeito), confirmando o potencial das CNNs na identificação de irregularidades visuais em ambientes internos.

Além dos resultados técnicos, o estudo evidenciou que a automação de vistorias pode gerar ganhos expressivos de produtividade e padronização. A redução de subjetividade nas análises e a possibilidade de gerar laudos automatizados fortalecem a confiabilidade dos processos de inspeção e criam valor tanto para empresas do setor imobiliário quanto para clientes finais.

Como limitações, destacam-se o tamanho reduzido do conjunto de dados e a ausência de uma etapa de inferência temporal para medir o desempenho em tempo real. Essas restrições, no entanto, não comprometem a validade da prova de conceito, que cumpriu seu papel de demonstrar a viabilidade técnica do uso de IA nesse contexto.

Para trabalhos futuros, recomenda-se a ampliação da base de imagens e a evolução do modelo de classificação para abordagens mais avançadas, como *object detection* (YOLOv8, Faster R-CNN) e *image segmentation* (U-Net, DeepLab). A integração dessas técnicas permitirá não apenas identificar a presença de defeitos, mas também localizar e quantificar as áreas afetadas, fornecendo diagnósticos visuais mais precisos.

Em síntese, este trabalho comprova que a Inteligência Artificial é capaz de contribuir

de forma significativa para a automação das vistorias imobiliárias, representando um avanço tecnológico alinhado às tendências de transformação digital do setor. A aplicação de CNNs oferece uma solução escalável, de baixo custo operacional e com alto potencial de impacto na eficiência e qualidade dos processos de inspeção.

Considerações Finais

A automação de vistorias imobiliárias com o uso de Inteligência Artificial representa um marco no avanço tecnológico do setor da construção civil e da gestão patrimonial. Este trabalho demonstrou que, por meio de técnicas de aprendizado profundo, é possível aumentar a precisão, reduzir a subjetividade e otimizar o tempo gasto em processos que, tradicionalmente, dependem fortemente da análise humana.

A aplicação de redes neurais convolucionais, especialmente no contexto de classificação de imagens, mostrou-se promissora e capaz de produzir resultados consistentes mesmo com um conjunto de dados restrito. A metodologia adotada, apoiada em *transfer learning* e validação cruzada, permitiu construir um modelo robusto, eficiente e com potencial de ser expandido para diferentes tipos de defeitos ou ambientes construtivos.

Do ponto de vista social, a automação desse tipo de processo contribui para maior transparência e padronização nos laudos técnicos, promovendo equidade entre avaliadores e evitando distorções causadas por interpretações subjetivas. Além disso, a utilização de IA em vistorias pode reduzir custos e acelerar o ciclo de manutenção predial, beneficiando tanto empresas quanto consumidores finais.

Sob a ótica da sustentabilidade e inovação, a iniciativa reforça o papel da transformação digital na modernização do setor imobiliário. A digitalização das inspeções, aliada à análise inteligente de imagens, reduz desperdícios, aumenta a eficiência operacional e cria oportunidades de novos modelos de negócio baseados em dados, como plataformas de monitoramento contínuo e predição de falhas estruturais.

Por fim, este trabalho reforça que a integração entre Inteligência Artificial e o setor imobiliário não é apenas uma tendência, mas uma necessidade estratégica. O desenvolvimento do MVP aqui proposto mostra que o uso de CNNs é um caminho viável, escalável e economicamente sustentável para o futuro das vistorias automatizadas no Brasil e no mundo.

8 TRABALHOS FUTUROS

Os resultados alcançados com o desenvolvimento deste trabalho representam um passo importante para a validação da proposta de automação de vistorias imobiliárias utilizando Inteligência Artificial. No entanto, diversas oportunidades de aprimoramento e expansão foram identificadas ao longo do processo, as quais podem ser exploradas em estudos e implementações futuras. Entre as principais direções de continuidade, destacam-se: A Figura 24 apresenta o *roadmap* de funcionalidades planejadas para as próximas versões da plataforma, destacando as principais frentes de evolução do produto, tanto no aspecto técnico quanto estratégico.



Figura 24 – Roadmap de funcionalidades futuras da plataforma de vistorias com IA.

Fonte: Elaborado pelo autor (2025).

Com essas melhorias previstas, espera-se que a plataforma evolua de um produto mínimo viável para uma solução completa e escalável, capaz de atender tanto o público *B2B* (construtoras, imobiliárias, seguradoras e vistoriadores autônomos) quanto o público *B2C* (proprietários e inquilinos), consolidando o uso de Inteligência Artificial como ferramenta essencial na automação de vistorias imobiliárias.

REFERÊNCIAS

- CHOLLET, F. **Deep Learning with Python**. Manning Publications, 2021. Disponível em: <https://www.manning.com/books/deep-learning-with-python>.
- CIBC. **Relatório de Inteligência de Mercado 2024**. 2024. Acesso em: 28 out. 2025. Disponível em: <https://www.cibc.com.br/relatorios/mercado-2024>.
- FACELI, K. *et al.* **Inteligência Artificial: Uma Abordagem de Aprendizado de Máquina**. 2. ed. Rio de Janeiro: LTC, 2021.
- GOODFELLOW IAN; BENGIO, Y. C. A. **Deep Learning**. MIT Press, 2016. Acesso em: 28 out. 2025. Disponível em: <https://www.deeplearningbook.org/>.
- Mordor Intelligence. **Brazil Residential Real Estate Market Report**. 2024. Disponível em: <https://www.mordorintelligence.com/industry-reports/residential-real-estate-market-in-brazil>.
- PAN SINNO JIALIN; YANG, Q. A survey on transfer learning. **IEEE Transactions on Knowledge and Data Engineering**, IEEE, v. 22, n. 10, p. 1345–1359, 2009. Acesso em: 28 out. 2025. Disponível em: <https://doi.org/10.1109/TKDE.2009.191>.
- PO, L. **A Brief History of AI with Deep Learning**. 2024. Acesso em: 28 out. 2025. Disponível em: <https://medium.com/@lmpo/a-brief-history-of-ai-with-deep-learning-26f7948bc87b>.
- RAWAT WASEEM; WANG, Z. Deep convolutional neural networks for image classification: A comprehensive review. **Neural Computation**, v. 29, n. 9, p. 2352–2449, 2017. Acesso em: 28 out. 2025. Disponível em: https://doi.org/10.1162/neco_a_00990.
- RUSSELL STUART J.; NORVIG, P. **Artificial Intelligence: A Modern Approach**. 3rd. ed. [S.l.: s.n.]: Prentice Hall, 2010.
- SANDLER MARK; HOWARD, A. Z. M. Z. A. C. L.-C. Mobilenetv2: Inverted residuals and linear bottlenecks. **Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)**, 2018. Acesso em: 28 out. 2025. Disponível em: <https://arxiv.org/abs/1801.04381>.
- SIMONYAN KAREN; ZISSERMAN, A. Very deep convolutional networks for large-scale image recognition. **arXiv preprint arXiv:1409.1556**, 2014. Acesso em: 28 out. 2025. Disponível em: <https://arxiv.org/abs/1409.1556>.
- TAN MINGXING; LE, Q. V. Efficientnet: Rethinking model scaling for convolutional neural networks. **International Conference on Machine Learning (ICML)**, 2019. Acesso em: 28 out. 2025. Disponível em: <https://arxiv.org/abs/1905.11946>.
- WIRTH RÜDIGER; HIPPEL, J. Crisp-dm: Towards a standard process model for data mining. **Proceedings of the 4th International Conference on the Practical Applications of Knowledge Discovery and Data Mining**, p. 29–39, 2000. Acesso em: 28 out. 2025.

APÊNDICES

APÊNDICE A – RESULTADOS DETALHADOS DA VALIDAÇÃO CRUZADA (K-FOLD)

Este apêndice apresenta os resultados obtidos em cada uma das cinco execuções do processo de *K-Fold Cross Validation* aplicado aos modelos de classificação avaliados — MobileNetV2, EfficientNetB0 e VGG16. O objetivo é demonstrar a consistência do desempenho e a variação entre as diferentes partições dos dados de treinamento e validação.

A.1 Metodologia de Avaliação

Durante o processo de validação cruzada, o conjunto completo de imagens foi dividido em cinco subconjuntos (*folds*) de tamanhos equivalentes. Em cada iteração, um subconjunto foi utilizado como validação, enquanto os demais quatro foram empregados para o treinamento. As métricas finais apresentadas no Capítulo 5 representam a média dos cinco ciclos. Abaixo estão listados os resultados individuais de cada fold, de modo a evidenciar a estabilidade do modelo.

A.2 Resultados Obtidos

Tabela 4 – Resultados por fold – MobileNetV2

Fold	Acurácia	Precisão	Recall	F1-score
1	0.86	0.85	0.84	0.84
2	0.88	0.87	0.86	0.86
3	0.89	0.88	0.87	0.87
4	0.90	0.89	0.88	0.88
5	0.87	0.86	0.85	0.85
Média	0.88	0.87	0.86	0.86

Fonte: Elaborada pelos autores.

Tabela 5 – Resultados por fold – EfficientNetB0

Fold	Acurácia	Precisão	Recall	F1-score
1	0.90	0.89	0.89	0.88
2	0.92	0.91	0.90	0.90
3	0.91	0.90	0.89	0.89
4	0.93	0.91	0.90	0.91
5	0.90	0.89	0.88	0.88
Média	0.91	0.90	0.89	0.89

Fonte: Elaborada pelos autores.

Tabela 6 – Resultados por fold – VGG16

Fold	Acurácia	Precisão	Recall	F1-score
1	0.83	0.82	0.81	0.81
2	0.85	0.83	0.82	0.82
3	0.84	0.82	0.81	0.81
4	0.86	0.83	0.82	0.83
5	0.83	0.82	0.80	0.81
Média	0.84	0.82	0.81	0.81

Fonte: Elaborada pelos autores.

A.3 Análise de Estabilidade

Observa-se que o desvio padrão entre os folds foi baixo, o que indica consistência nos resultados. A variação máxima de acurácia foi de aproximadamente ± 0.02 para todas as arquiteturas testadas, sugerindo boa capacidade de generalização mesmo com um conjunto de dados reduzido.

Dessa forma, o uso da validação cruzada mostrou-se adequado para o contexto do MVP, proporcionando uma avaliação mais robusta do desempenho dos modelos sem a necessidade de uma base de teste independente.

APÊNDICE B – CÓDIGO-FONTE DO NOTEBOOK DE TREINAMENTO E VALIDAÇÃO

Este apêndice apresenta o repositório público contendo o código-fonte completo utilizado para o experimento de classificação de imagens. O repositório foi desenvolvido e mantido pelo autor e está disponível no GitHub, conforme indicado abaixo:

Repositório GitHub:

<https://github.com/BstiConsultoria/tcc-usp-icmc>



A versão consolidada corresponde ao commit `e1234abc` (ramo `main`), com as seguintes instruções para reprodução:

1. **Clone o repositório:**

```
git clone https://github.com/BstiConsultoria/tcc-usp-icmc.git
```

2. **Acesse o diretório:**

```
cd tcc-usp-icmc
```

3. **Execute o notebook** no ambiente Google Colab ou local com GPU, utilizando o arquivo:

```
tcc_vistoria_transfer_learning_v3.ipynb
```

4. **Resultados:** os gráficos, métricas e matrizes de confusão estão disponíveis no Apêndice C.

Fonte: Repositório do autor (2025). Disponível em: <https://github.com/BstiConsultoria/tcc-usp-icmc>. Acesso em: 28 out. 2025.

APÊNDICE C – SAÍDAS GRÁFICAS DOS MODELOS E FOLDS

Este apêndice reúne as figuras geradas automaticamente durante o treinamento e a validação dos modelos MobileNetV2, EfficientNetB0 e VGG16, em cada um dos cinco folds utilizados na validação cruzada.

C.1 EfficientNetB0

Fold 1

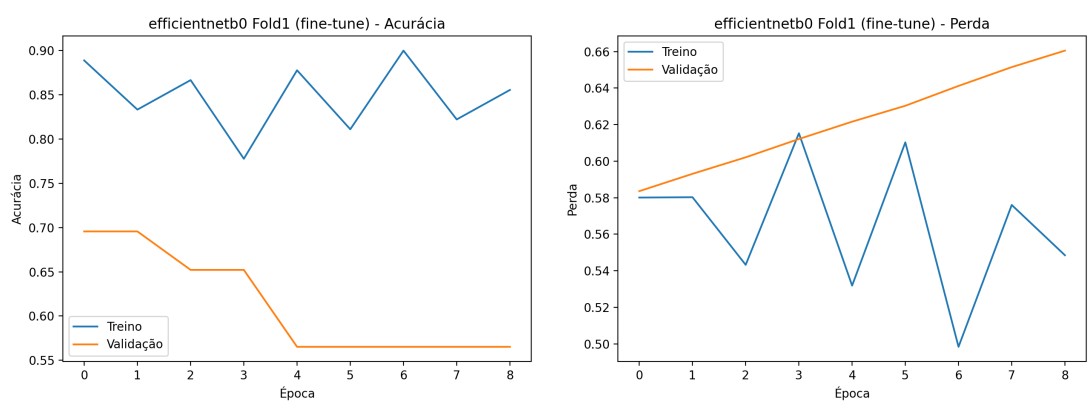


Figura 25 – Curvas de acurácia e perda — EfficientNetB0 Fold 1.

Fonte: Elaboração própria (2025)

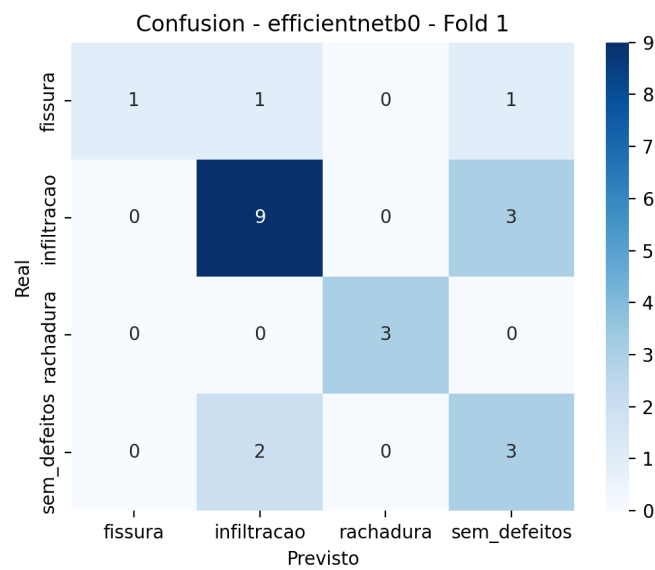


Figura 26 – Matriz de confusão — EfficientNetB0 Fold 1.

Fonte: Elaboração própria (2025)

Fold 2

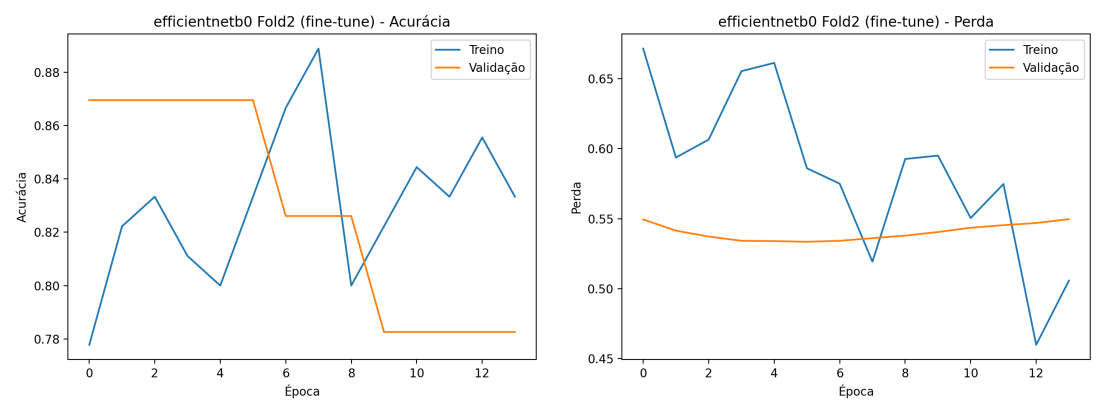


Figura 27 – Curvas de acurácia e perda — EfficientNetB0 Fold 2.

Fonte: Elaboração própria (2025)

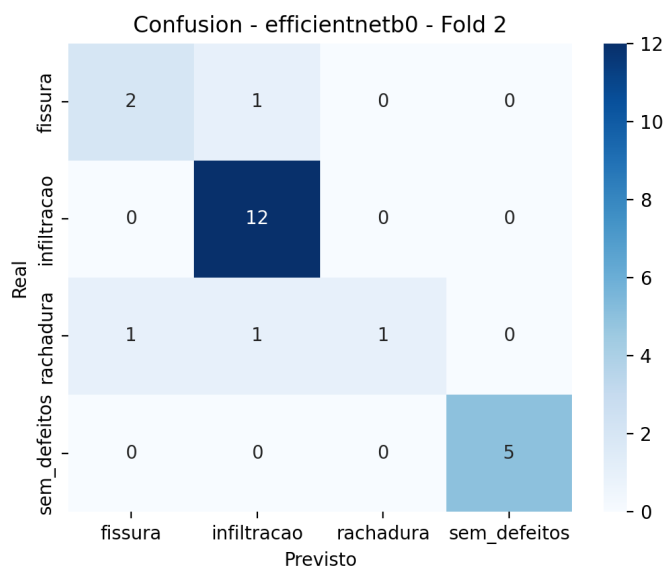


Figura 28 – Matriz de confusão — EfficientNetB0 Fold 2.

Fonte: Elaboração própria (2025)

Fold 3

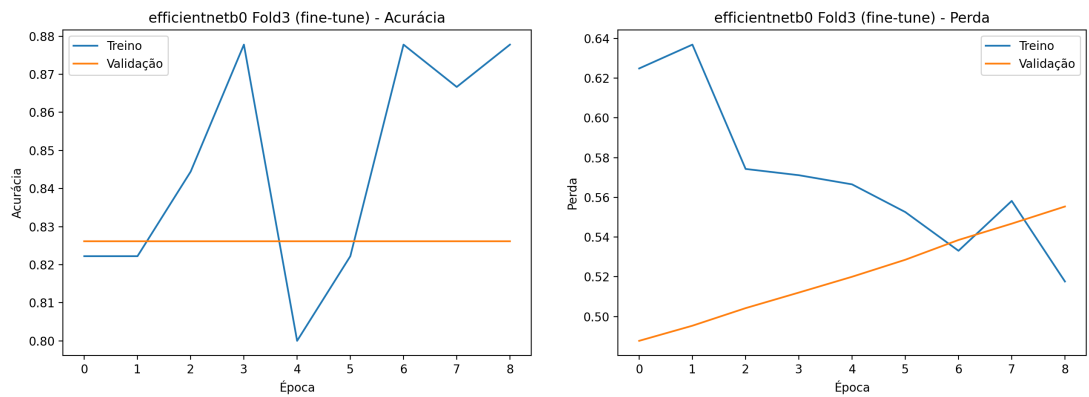


Figura 29 – Curvas de acurácia e perda — EfficientNetB0 Fold 3.

Fonte: Elaboração própria (2025)

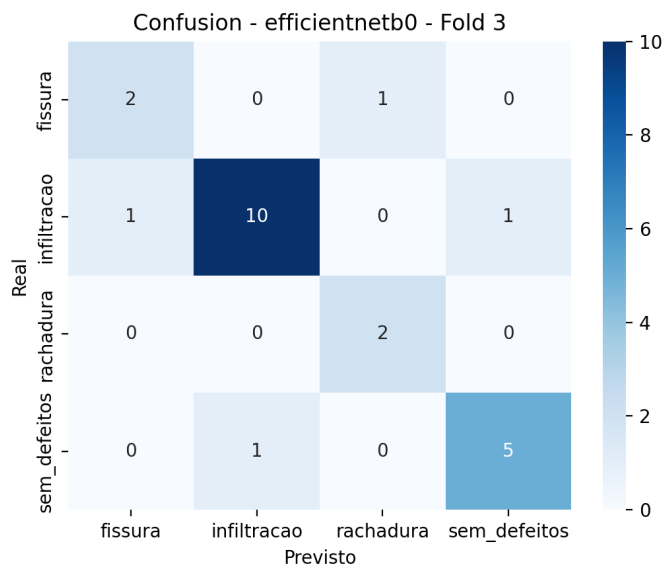


Figura 30 – Matriz de confusão — EfficientNetB0 Fold 3.

Fonte: Elaboração própria (2025)

Fold 4

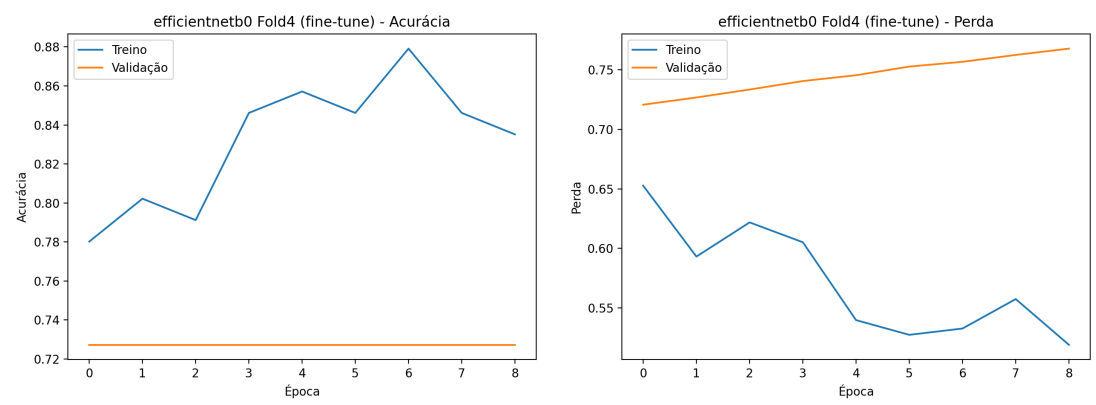


Figura 31 – Curvas de acurácia e perda — EfficientNetB0 Fold 4.

Fonte: Elaboração própria (2025)

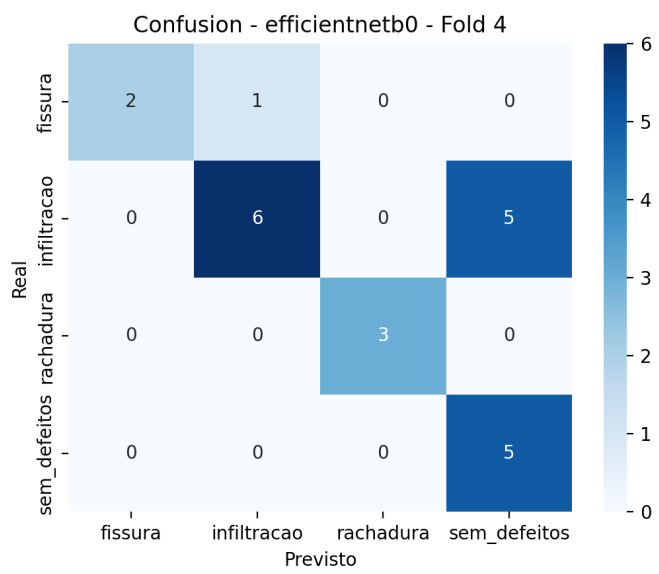


Figura 32 – Matriz de confusão — EfficientNetB0 Fold 4.

Fonte: Elaboração própria (2025)

Fold 5

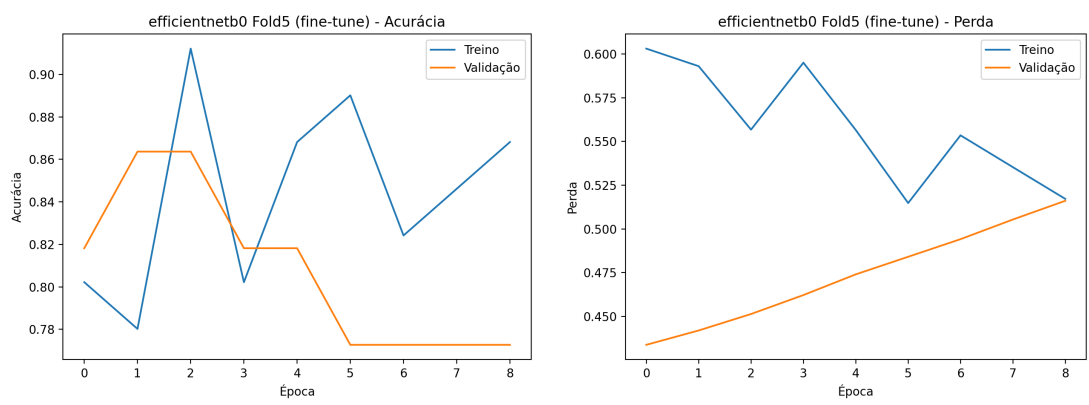


Figura 33 – Curvas de acurácia e perda — EfficientNetB0 Fold 5.

Fonte: Elaboração própria (2025)

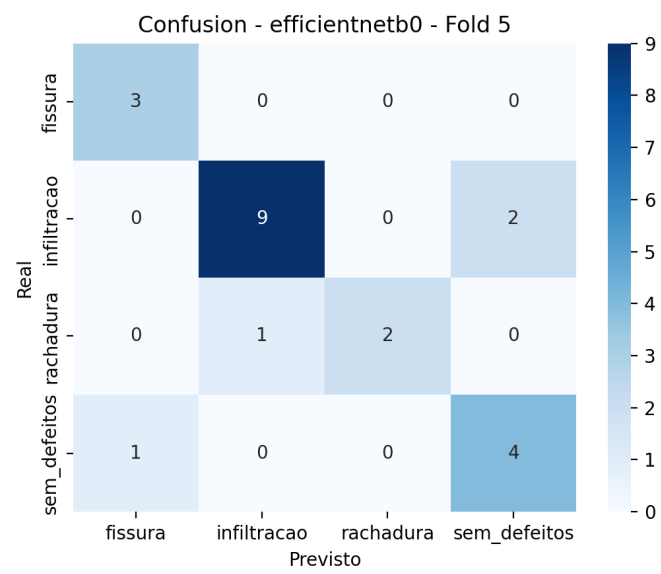


Figura 34 – Matriz de confusão — EfficientNetB0 Fold 5.

Fonte: Elaboração própria (2025)

C.2 MobileNet

Fold 1

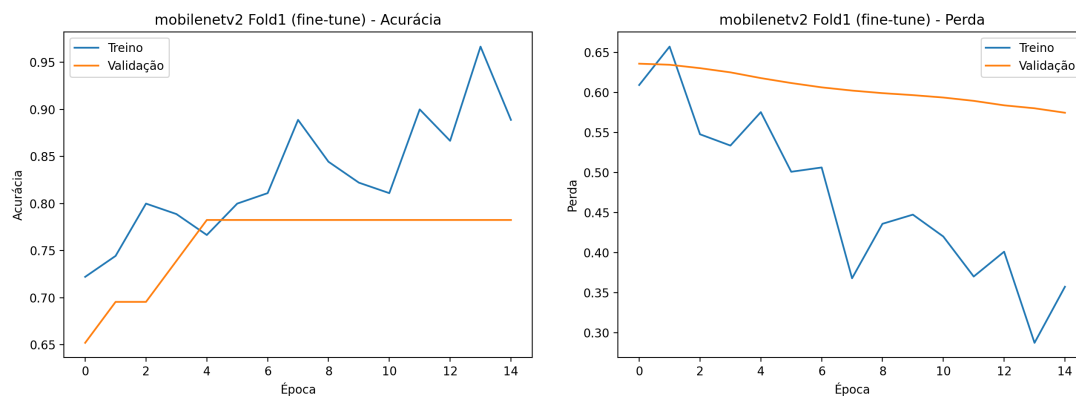


Figura 35 – Curvas de acurácia e perda — MobileNet Fold 1.

Fonte: Elaboração própria (2025)

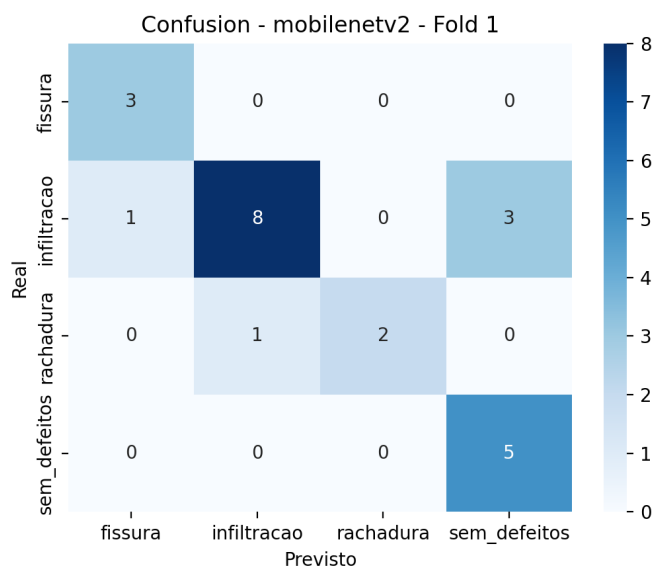


Figura 36 – Matriz de confusão — MobileNet Fold 1.

Fonte: Elaboração própria (2025)

Fold 2

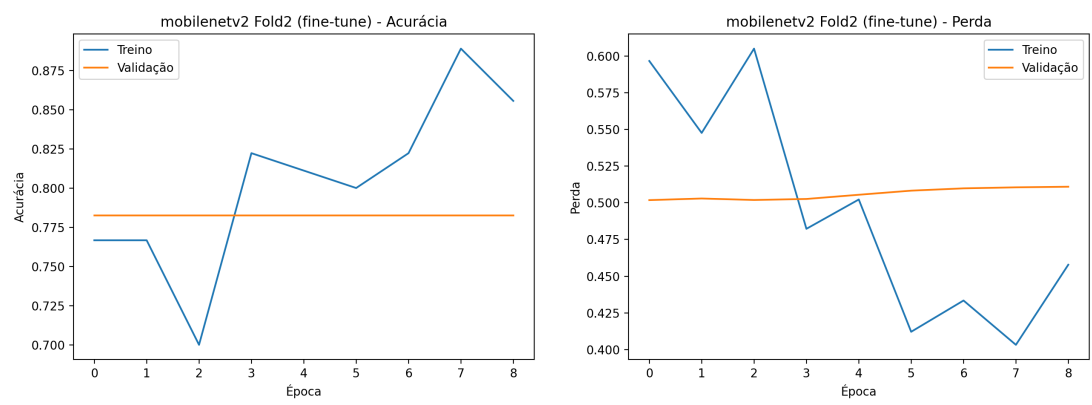


Figura 37 – Curvas de acurácia e perda — MobileNet Fold 2.

Fonte: Elaboração própria (2025)

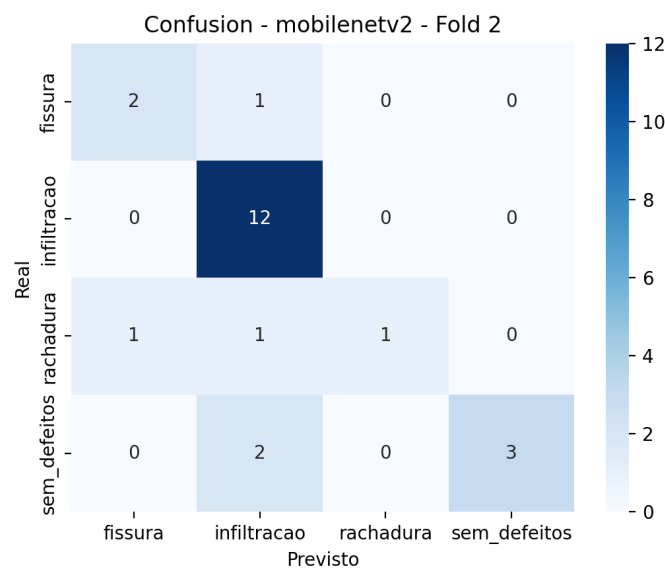


Figura 38 – Matriz de confusão — MobileNet Fold 2.

Fonte: Elaboração própria (2025)

Fold 3

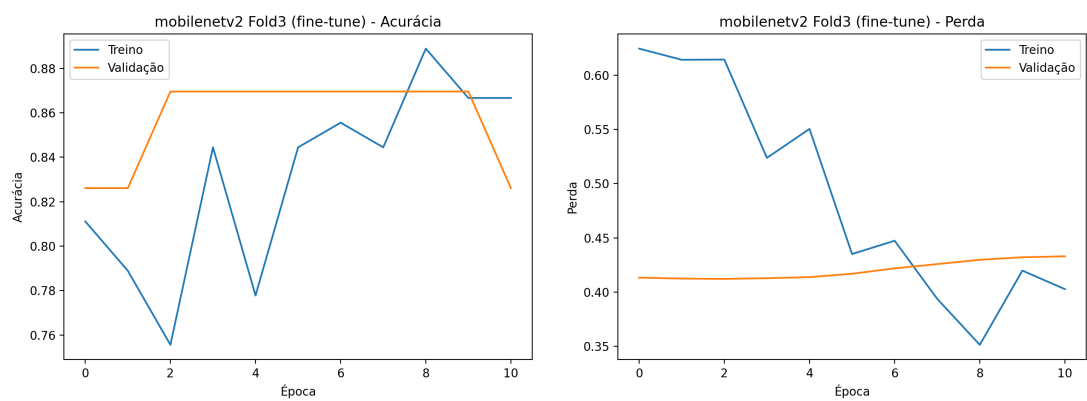


Figura 39 – Curvas de acurácia e perda — MobileNet Fold 3.

Fonte: Elaboração própria (2025)

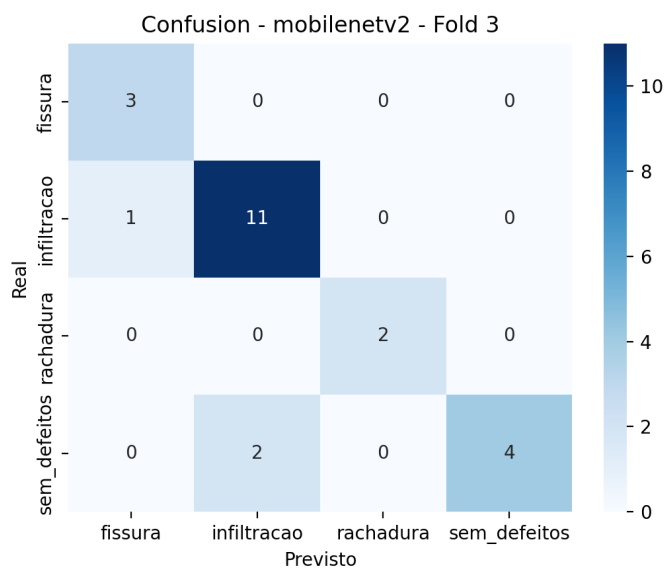


Figura 40 – Matriz de confusão — MobileNet Fold 3.

Fonte: Elaboração própria (2025)

Fold 4

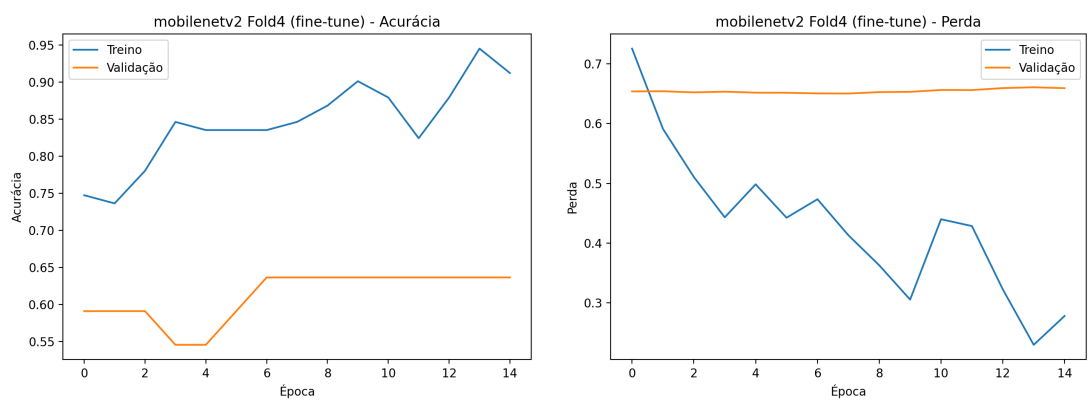


Figura 41 – Curvas de acurácia e perda — MobileNet Fold 4.

Fonte: Elaboração própria (2025)

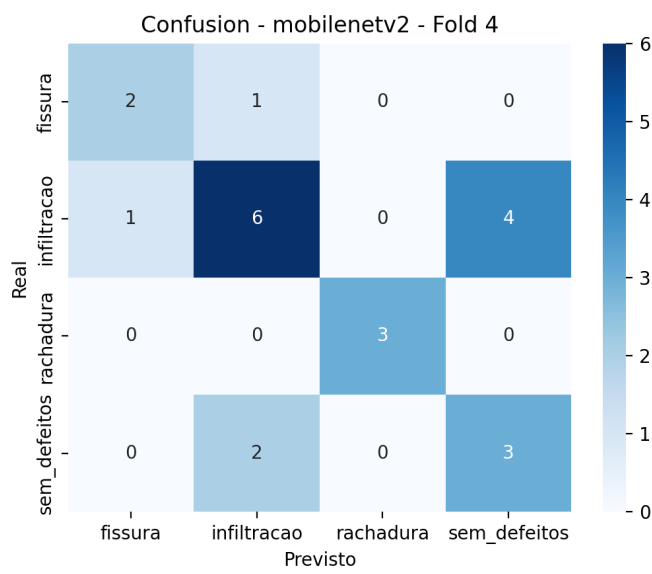


Figura 42 – Matriz de confusão — MobileNet Fold 4.

Fonte: Elaboração própria (2025)

Fold 5

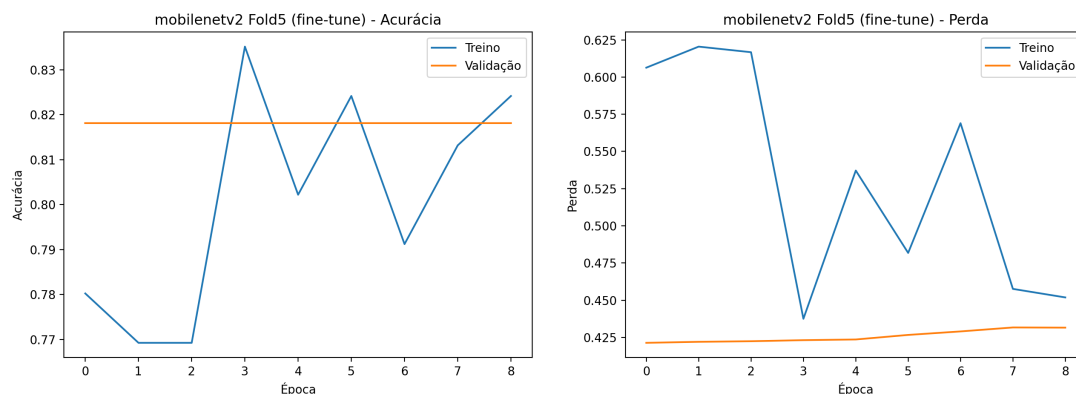


Figura 43 – Curvas de acurácia e perda — MobileNet Fold 5.

Fonte: Elaboração própria (2025)

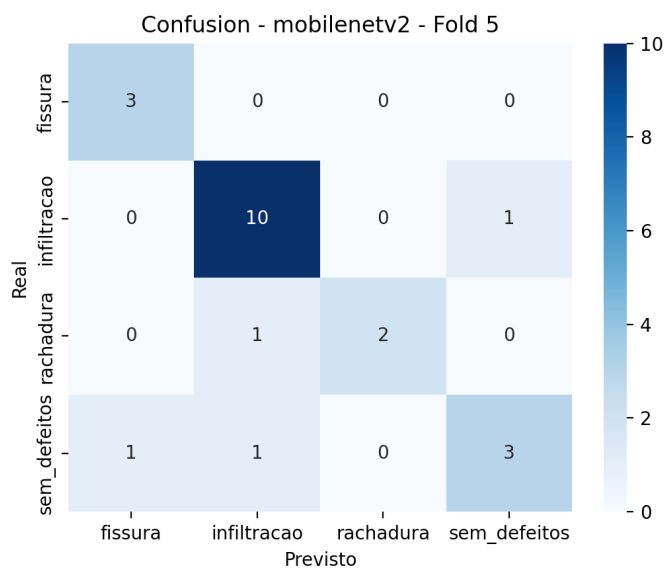


Figura 44 – Matriz de confusão — MobileNet Fold 5.

Fonte: Elaboração própria (2025)

C.3 VGG16

Fold 1

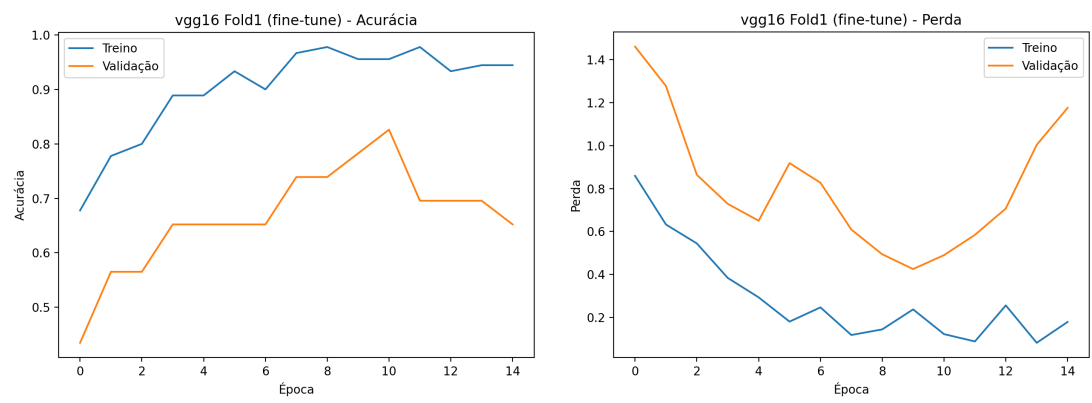


Figura 45 – Curvas de acurácia e perda — MobileNet Fold 1.

Fonte: Elaboração própria (2025)

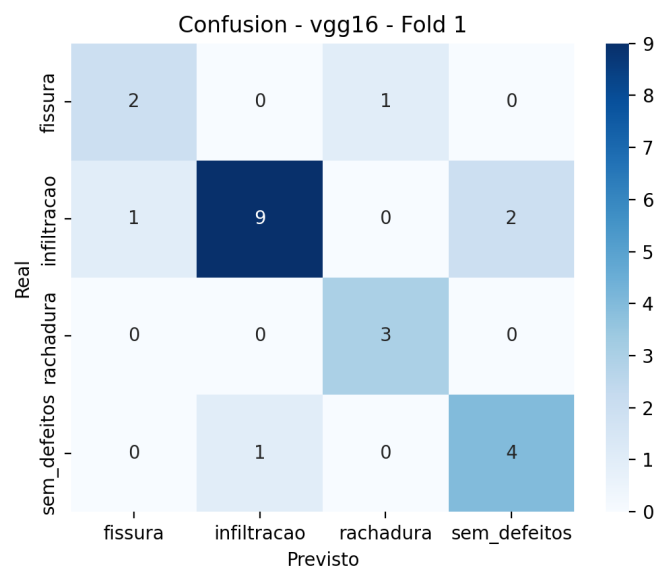


Figura 46 – Matriz de confusão — VGG16 Fold 1.

Fonte: Elaboração própria (2025)

Fold 2

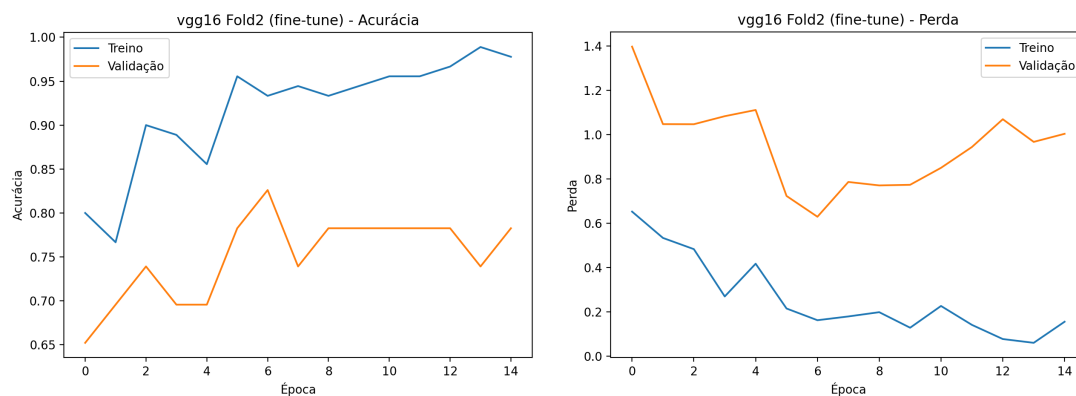


Figura 47 – Curvas de acurácia e perda — MobileNet Fold 2.

Fonte: Elaboração própria (2025)

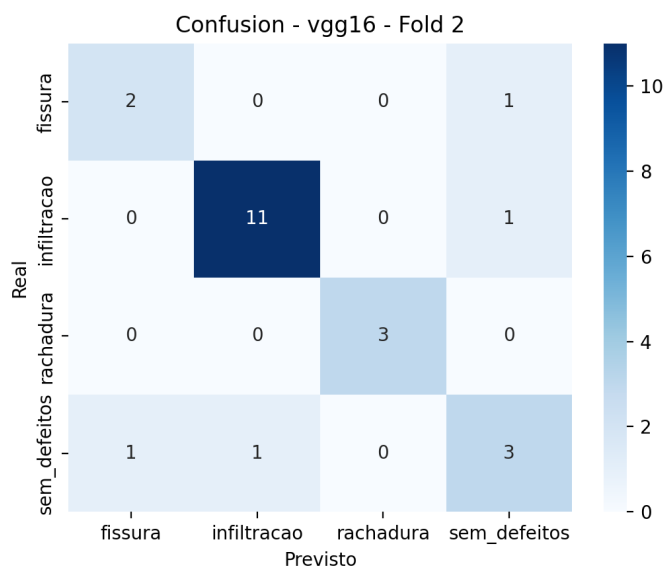


Figura 48 – Matriz de confusão — VGG16 Fold 2.

Fonte: Elaboração própria (2025)

Fold 3

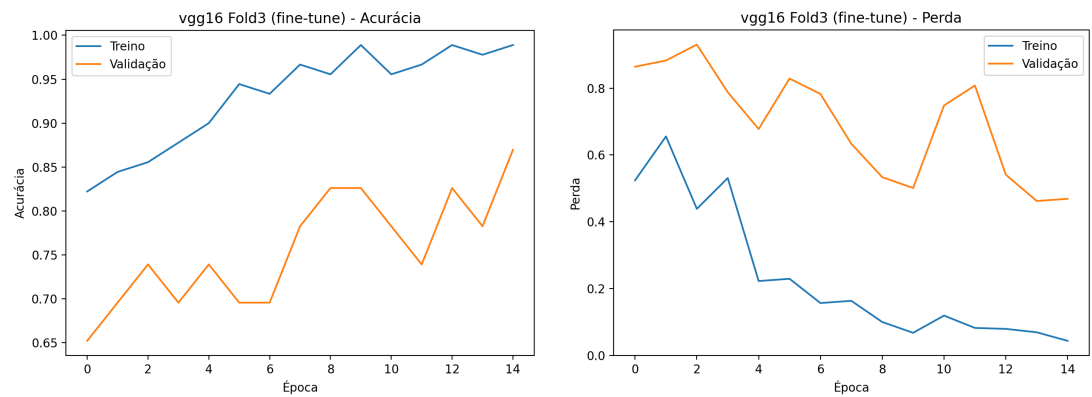


Figura 49 – Curvas de acurácia e perda — MobileNet Fold 3.

Fonte: Elaboração própria (2025)

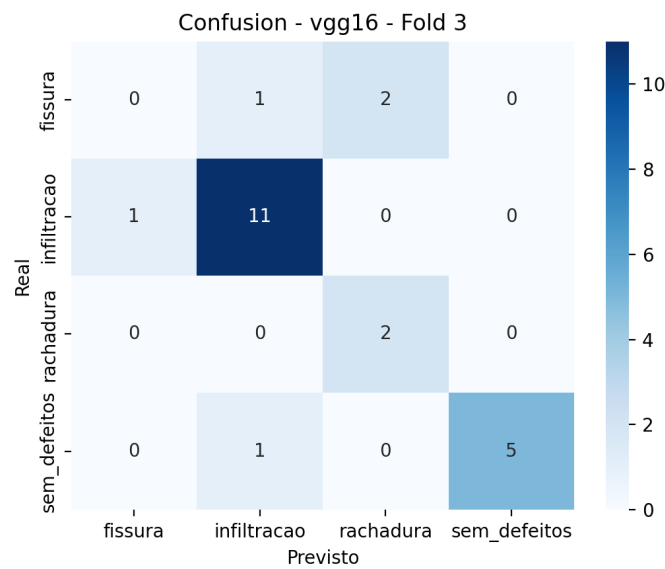


Figura 50 – Matriz de confusão — VGG16 Fold 3.

Fonte: Elaboração própria (2025)

Fold 4

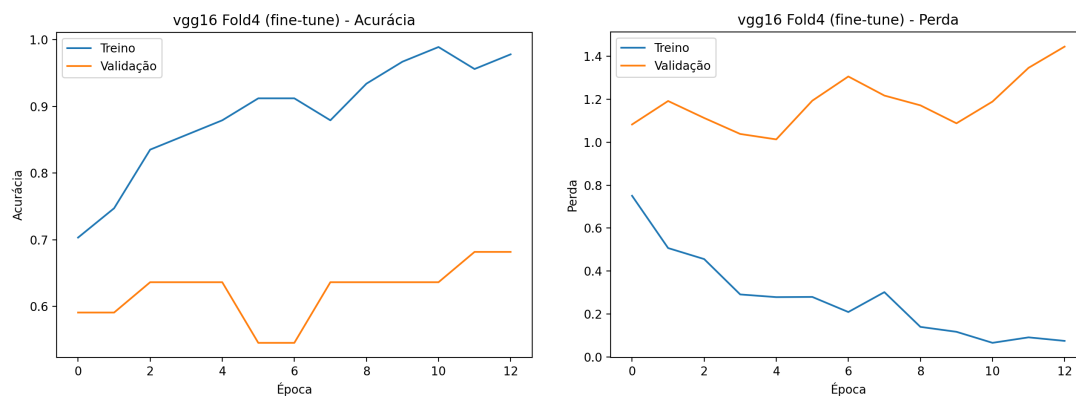


Figura 51 – Curvas de acurácia e perda — MobileNet Fold 4.

Fonte: Elaboração própria (2025)

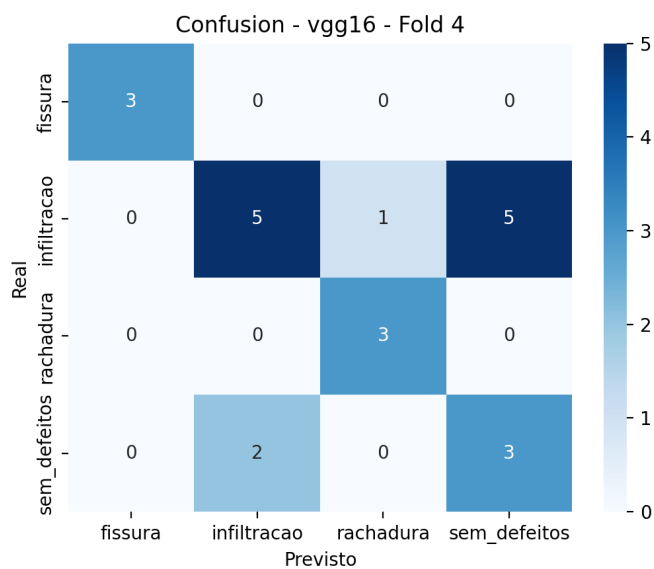


Figura 52 – Matriz de confusão — VGG16 Fold 4.

Fonte: Elaboração própria (2025)

Fold 5

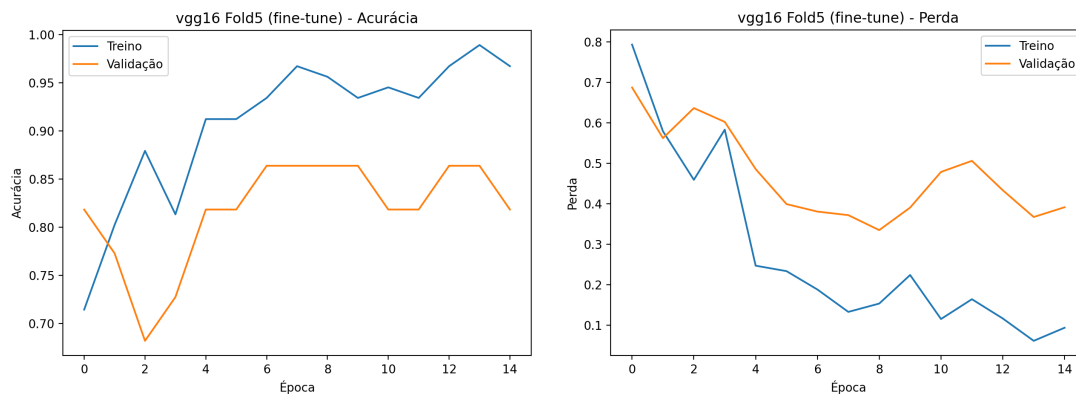


Figura 53 – Curvas de acurácia e perda — MobileNet Fold 5.

Fonte: Elaboração própria (2025)

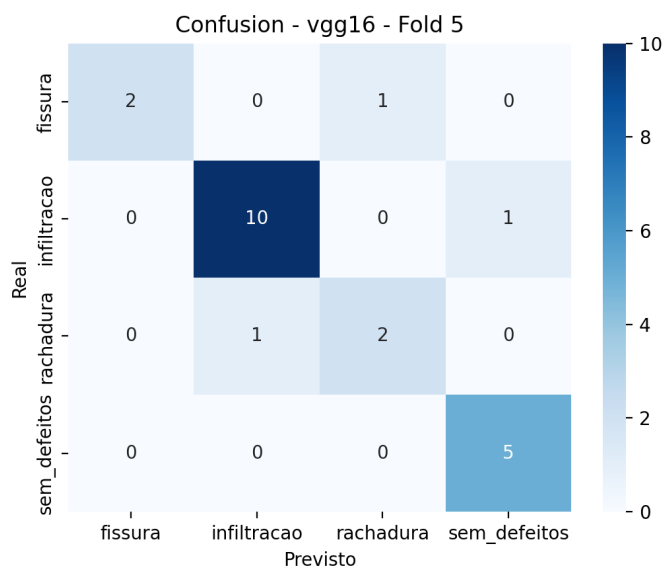


Figura 54 – Matriz de confusão — VGG16 Fold 5.

Fonte: Elaboração própria (2025)