

# IDENTIFICAÇÃO BIOMÉTRICA POR MEIO DE APRENDIZADO DE MÁQUINA APLICADO AO RECONHECIMENTO FACIAL DE CÃES

Misael Fernandes Córdova<sup>1</sup>, Merisandra Côrtes de Mattos<sup>2</sup>

**Resumo:** Um número expressivo de cães vive em situação de rua, sem lar ou tutor responsável, o que acarreta diversos problemas sociais, sanitários e de segurança. Animais domésticos abandonados podem representar riscos à saúde pública, ao bem-estar animal e à integridade das comunidades. A implementação de políticas efetivas de controle, rastreabilidade e cuidados depende, entre outros fatores, da capacidade de identificar individualmente esses animais. Diante desse cenário, esta pesquisa teve como objetivo avaliar e aplicar técnicas de aprendizado profundo para o reconhecimento facial de cães, utilizando duas abordagens distintas: o modelo *AlexNet* modificado, e outro baseado na arquitetura do *Resnet* com função *triplet loss*. O estudo foi estruturado em cinco fases principais: preparação da base de imagens, desenvolvimento das redes neurais, treinamento e testes dos modelos, avaliação das métricas de desempenho e análise comparativa dos resultados. Os experimentos demonstraram que o modelo *AlexNet* modificado apresentou desempenho superior, atingindo uma acurácia de até 97,8%, enquanto o modelo baseado na ResNet com triplet loss obteve acurácia máxima de 81,7%. Esses resultados reforçam a viabilidade do uso de redes neurais convolucionais adaptadas para fins de identificação automatizada de cães, podendo contribuir para sistemas de controle populacional, resgate e políticas públicas voltadas à proteção e bem-estar animal.

**Palavras-chave:** Inteligência artificial; Aprendizado profundo; Reconhecimento facial; Biometria Animal; *Resnet*; *AlexNet*.

---

<sup>1</sup>Grupo de Pesquisa em Inteligência Artificial Aplicada, Curso de Ciência da Computação, Universidade do Extremo Sul Catarinense (Unesc), misael@unesc.net

<sup>2</sup>Grupo de Pesquisa em Inteligência Artificial Aplicada, Curso de Ciência da Computação, Universidade do Extremo Sul Catarinense (Unesc), mem@unesc.net

**ABSTRACT:** A significant number of dogs live on the streets, without a home or responsible guardian, which causes various social, health, and safety problems. Abandoned domestic animals can pose risks to public health, animal welfare, and community integrity. The implementation of effective control, traceability, and care policies depends, among other factors, on the ability to individually identify these animals. Given this scenario, this research aimed to evaluate and apply deep learning techniques for facial recognition of dogs, using two different approaches: the modified AlexNet model and another based on the Resnet architecture with triplet loss function. The study was structured in five main phases: preparation of the image database, development of neural networks, training and testing of models, evaluation of performance metrics, and comparative analysis of results. The experiments demonstrated that the modified AlexNet model performed better, achieving an accuracy of up to 97.8%, while the ResNet-based model with triplet loss achieved a maximum accuracy of 81.7%. These results reinforce the feasibility of using convolutional neural networks adapted for automated dog identification, which can contribute to population control systems, rescue, and public policies aimed at animal protection and welfare.

**Keywords:** Artificial Intelligence; Deep Learning; Facial Recognition; ResNet; AlexNet

## 1 INTRODUÇÃO

O número de cães domésticos em todo o mundo foi estimado em 700 milhões, cerca de 75% dos cães não possui um lar ou dono (Smith et al., 2019). Do total de cães, estima-se que 52,2 milhões vivam no Brasil (Arruda; Garcia; Oliveira, 2020). Considerando cães e gatos, há aproximadamente 200 milhões de animais abandonados no mundo (Costa, 2022b), sendo 30 milhões no Brasil, dos quais 10 milhões são gatos e 20 milhões são cães. (Moda, 2023). Mundialmente, não somente a população de cães abandonados, mas a de animais não monitorados gera, principalmente em países em desenvolvimento, uma série de problemas, como de saúde, socioeconômico, de bem-estar, acidentes de trânsito e ataques desses animais a população (Mota-Rojas et al., 2021).

No controle e identificação são utilizados por vezes métodos tradicionais como marcadores corporais, tatuagens, entalhes na orelha e pintura, os quais são invasivos e ineficientes para uma população grande de animais (Kumar; Singh, 2014). O problema de reconhecimento e monitoramento de animais domésticos é uma questão importante nas cidades

inteligentes, pois pode ocorrer de animais de estimação serem perdidos, trocados ou roubados (Kumar; Singh, 2018).

As cidades inteligentes visam principalmente o conforto, a economia de recursos, a recuperação eficiente de informações, a comunicação rápida e a melhor segurança para os cidadãos e os animais domésticos (Kumar; Singh, 2018). Uma alternativa para alcançar esses objetivos são as tecnologias da internet das coisas, que fazem uso do sistema de posicionamento global em inglês *Global Position System* (GPS), redes *Bluetooth*, redes *Wireless* e Identificação por Rádio Frequência. No entanto, segundo Jang et al. (2020), tais tecnologias nem sempre são a escolha preferida dos proprietários já que o dispositivo é injetado na epiderme do animal, o que pode causar efeitos colaterais negativos em relação ao bem-estar do animal. Dispositivos externos, como coleiras e roupage m podem ser perdidos, além de ineficazes por conta de os tutores não serem rigorosos o bastante para garantir que o animal o use.

A biometria visual animal é um sistema alternativo a esse e as demais soluções, baseada em reconhecimento de padrões, usa as características fisiológicas ou a aparência fenotípica para o reconhecimento do animal (Kumar; Singh, 2014).

Os métodos de identificação biométrica animal são classificados em quatro categorias: pontos de focinho, padrão de íris, retina vascular e imagens faciais (Kumar; Singh, 2018). O reconhecimento de pontos do focinho funciona de modo parecido com as impressões digitais humanas, constituindo-se em um método eficiente e confiável. No entanto, para extrair as características do focinho do animal são necessárias imagens de alta resolução ou até mesmo equipamento de scanner, o que dificulta a coleta desses dados necessários para o treinamento dos algoritmos. As dificuldades também ocorrem para a coleta dos dados referentes ao padrão de íris e a identificação vascular. Entretanto, nos repositórios digitais é possível encontrar uma ampla variedade de imagens faciais dos cães.

O reconhecimento biométrico animal pode ser realizado por meio do reconhecimento de padrões usando aprendizado de máquina, que são métodos orientados por dados capazes de reproduzir, projetar e auxiliar no processamento das informações (Barber, 2012). Dentre esses métodos, destacam-se as redes neurais artificiais, algoritmos matemáticos inspirados no funcionamento do cérebro humano, especialmente nos neurônios e nas suas conexões (Costa, 2022a). Em especial, as Redes Neurais Convolucionais (RNCs) são inspiradas no córtex visual dos animais, atuando

como filtros sensíveis a pequenas sub-regiões dos dados de entrada, o que permite identificar correlações espaciais e aprender características visuais adaptáveis ao conjunto de dados (Albawi; Mohammed; Al-Zawi, 2017).

As RNCs já demonstraram resultados expressivos no reconhecimento facial humano nos estudos de Taigman et al. (2014); Schroff, Kalenichenko e Philbin (2015) e Deng et al. (2019a), bem como no reconhecimento facial de cães, com destaque para os modelos baseados em *ResNet* treinados com triplet loss (Mougeot; Li; Jia, 2019) e outro modelo popular como o *AlexNet*, que também obteve desempenho significativo em um sistema biométricos multimodal nos estudos de Yashavanth e Suresh (2023) e reconhecimento facial humano por aprendizado por transferência nos estudos de Singh et al. (2022).

Uma série de aspectos culturais podem alterar a influência dos cães sobre um ambiente e na saúde humana. Diferentes níveis de tolerância são observados quando se trata de responsabilidades associadas a posse desses animais. Em alguns países, principalmente os desenvolvidos, foram criadas políticas públicas com o intuito de minimizar os efeitos indesejáveis que animais domésticos podem causar a população, podendo-se elencar: danos a propriedade de outrem, desequilíbrio da fauna local, superpopulação de cães sem um lar e transmissão de doenças (Souza, 2017). A fim de minimizar os danos que uma população de animais não monitorada pode causar, são criadas campanhas educativas para a guarda responsável, auxílio a esterilização, vacinação e programas de adoção. No entanto, para que essas ações tenham êxito é necessário um mecanismo para a identificação do animal (Souza, 2017).

Um sistema de identificação animal desempenha um importante papel no controle, no gerenciamento de vacinação, na prevenção de doenças, na gestão da produção, na atribuição de propriedade e na rastreabilidade de cães (Souza, 2017). Além de visar garantir o bem-estar animal, esta pesquisa tem aderência direta aos Objetivos de Desenvolvimento Sustentável (ODS) propostos pela ONU, particularmente ao ODS 3, que se refere à saúde e bem-estar, ao prevenir surtos de doenças e facilitar o rastreamento de cães perdidos ou abandonados. Também está alinhado ao ODS 11, que busca tornar as cidades e comunidades mais sustentáveis, seguras e inclusivas, através de um monitoramento efetivo dos animais domésticos, contribuindo para reduzir acidentes e promover uma melhor convivência entre animais e seres humanos.

Neste contexto essa pesquisa busca aplicar um modelo modifi-

cado do *AlexNet* a fim de reconhecer cães e contrastá-lo com o modelo que é o estado atual da arte para o reconhecimento facial de cães que é o modelo baseado no *ResNet* que aplica a triplet loss. Assim, por meio da tecnologia da inteligência artificial é possível a criação de um sistema não invasivo, econômico e automático para o monitoramento dos animais de companhia.

## 2 TRABALHOS CORRELATOS

Na literatura, existem trabalhos internacionais relacionados à identificação facial de cães, dentre estes, pode-se citar o estudo de Kumar et al. (2017) que teve como objetivo o desenvolvimento de um sistema de reconhecimento de gado individualmente com base nos pontos do focinho. Para isso utilizaram a Projeção e Preservação de Localidade de Fisher em tempo real, foram aplicados métodos de distância e similaridade única do inglês *One-shot similarity (OSS)*, com os modelos de classificação *Support Vector Machine(SVM)* incremental. Para o pré-processamento das imagens foi aplicada a técnica de Equalização de Histograma Adaptativo Limitado por Contraste, do inglês *Contrast Limited Adaptive Histogram Equalization (CLAHE)*, para melhorar o contraste das imagens. Posteriormente foram aplicadas técnicas de segmentação para identificar a área de interesse e aplicado o agrupamento de cores pelo algoritmo *K-means* e algoritmos de segmentação baseados em textura. A validação das precisões foi feita de forma cruzada cinco vezes. A Projeção de Preservação Linear de Fisher do inglês *Fisher Locality Preserving Projections* é comparada com algoritmos descritores como recursos robustos acelerados do inglês *Speeded Up Robust Feature (SURF)*, padrão binário local do inglês *Local Binary Pattern*, o algoritmo Transformação de característica invariante de escala do inglês *Scale Invariant Feature Transform (SIFT)* e algoritmos de reconhecimento facial baseados na aparência Análise de Componentes Principais do inglês *Principal Component Analysis (PCA)*, Análise discriminante linear do inglês *Linear Discriminant Analysis*, Análise de Componentes Independentes do inglês *Independent Component Analysis (ICA)* e suas variações. Dentre os experimentos realizados a técnica o que obteve o melhor desempenho foi o método FLPP proposto atingindo uma porcentagem de acurácia de 96,87%.

Moreira et al. (2017) buscaram o reconhecimento facial para cães, com métodos usados para reconhecer humanos e avaliaram a sua eficácia nesse domínio de aplicação. Além de realizar também a comparação

de duas redes neurais para o reconhecimento facial de cães. Utilizaram como linhas de base algoritmos LBPH, *EingFace* e *FisherFace* e um método *Sparse*. O treinamento dos modelos foi realizado criando-se duas bases de dados de imagens que chamaram de *Flickr-dog* e *Snoopybook*. Primeiramente, propuseram rede neurais convolucionais superficiais (com até 3 camadas), Otimizaram a arquitetura da rede e empregaram pesos aleatórios o que chamaram de *Best Architecture to Retrieve K9 (BARK)*, a rede era composta por camadas empilhadas de: filtros lineares de convolução, função de ativação e normalização, a rede resultante extrai os recursos que são alimentados em um SVM Linear. A segunda proposta usa redes neurais profundas e pré-treinadas com o modelo Overfeat <sup>3</sup>. A rede foi melhorada e compreendeu seis camadas convolucionais e três totalmente conectadas, chamaram essa rede de *Wields Off-the-shelf OverFeat Feature (WOOF)*. Os autores relatam que redes convolucionais superficiais e profundas podem alcançar melhores resultados do que os reconhecedores faciais humanos quando utilizado em cães, além de que uma rede profunda pré-treinada demonstra melhores resultados.

Yoon, So e Rhee (2021) propõem o treinamento de um modelo em dois estágios para a utilização de um abrangente espaço vetorial. O modelo escolhido para treinamento foi o *Resnet-like* que foi construído com uma camada de *dropout* e usa camadas residuais para não haver *overfitting*. A metodologia do trabalho se deu em remover uma camada de normalização L2 para utilizar o aprendizado baseado na similaridade de cosseno no modelo *Resnet-like*. Além de treinar o modelo com a perda de tripleto, foi proposta uma função de perda que ajusta o comprimento do vetor de incorporação visando reduzir a diferença no comprimento dos mesmos entre a mesma classe. Foi possível verificar que a metodologia proposta superou o método tradicional de perda em termos de precisão e perda. Foi possível também alcançar os melhores resultados em outros modelos, como *FaceNet* <sup>4</sup>, *CosFace* <sup>5</sup> e *ArcFace* <sup>6</sup>. De modo geral, a metodologia aplicada alcançou uma melhora de 4% no desempenho do modelo de identificação

<sup>3</sup>Rede utilizada nos estudos de Sermanet et al. (2013), semelhante a proposta por Krizhevsky, Sutskever e Hinton (2012), no entanto, com melhor capacidade de inferência (Moreira et al., 2017)

<sup>4</sup>RNC profunda proposta por Schroff, Kalenichenko e Philbin (2015), utiliza-se de vetores de incorporação e a perda de tripleto para o treinamento.

<sup>5</sup>RNC profunda proposta Wang et al. (2018), utiliza-se a perda de cosseno de grande margem para o treinamento.

<sup>6</sup>RNC profunda proposta por Deng et al. (2019b) que utiliza-se a perda de margem angular aditiva para o treinamento.

da face de cachorros em termos de precisão e perda.

Tabela 1: Trabalhos Correlatos

Trabalho	Modelos	Resultados	Métricas
Kumar et al. (2017)	SURF	SURF: 94,57%	Acurácia
	LPB	LPB: 92,95%	
	PCA	PCA: 75,86%	
	LDA	LDA: 79,95%	
	ICA	ICA: 87,95%	
	FLPP	Modelo proposto: 96,87%	
Moreira et al. (2017)		Flickr-dog	Precisão
	EigenFaces	EigenFaces: 33,9%	
	FisherFaces	FisherFaces: 22,7%	
	LBPH	LBPH: 43,2%	
	Sparse	Sparse: 39,9%	
	BARK	BARKflickr: 67,6%	
WOOF	BARKsnoopy: 49,1%		
Yoon, So e Rhee (2021)		Resnet-like: 65,00%	Acurácia
	FaceNet	Triplet: 93,33%	
	CosFace	ArcFace: 84,00%	
	ArcFace	Triplet + VL stage 1: 89,00%	
	Resnet-like	Triplet + VL stage 2: 97,33%	
		ArcFace + VL stage 1: 84,33%	
	ArcFace + VL stage 2: 96,33%		
Ramlee et al. (2020)	AlexNet	Georgia Tech: 96.67% CASIA Face: 95.60% Estudantes University Teknikal Malaysia Melaka: 100,00%	Acurácia

Fonte: Elaborado pelo autor

Ramlee et al. (2020) buscaram obter alta precisão no reconhecimento facial humano utilizando a técnica de aprendizado por transferência com a rede neural *AlexNet* para a extração de características. Para isso, realizaram ajustes nas últimas camadas da rede, adaptando-as a novas tarefas de classificação, com o objetivo de melhorar o desempenho do reconhecimento mesmo diante de um número reduzido de imagens e limitações de hardware. Durante o pré-processamento, empregaram técnicas de aumento de imagens (*data augmentation*) para reduzir problemas de *overfitting*. Os testes realizados nos conjuntos de dados *Georgia Tech Face Database*, *CASIA Face Image Database* e um banco de imagens de estudantes da *University Teknikal Malaysia Melaka* demonstraram uma precisão

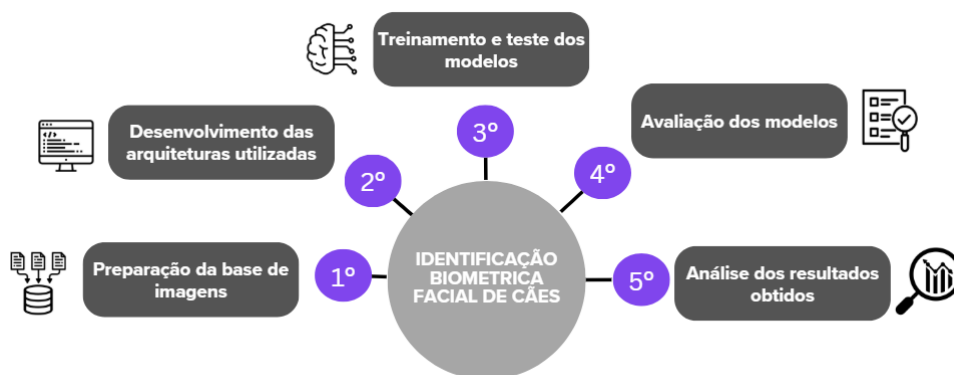
de 96.67%, 95.60% e 100% respectivamente, além de um menor tempo de reconhecimento. A tabela 1 descreve os trabalhos correlatos a esta pesquisa, mostrando os modelos utilizados, as métricas e o resultados.

### 3 MATERIAIS E MÉTODOS

Esta pesquisa consiste em um estudo aplicado no campo da visão computacional, com ênfase no reconhecimento facial de cães. Conforme referido anteriormente, o objetivo principal é a análise da performance e eficácia de métodos de aprendizado profundo, por meio de um modelo convolucional modificado a partir da arquitetura do *AlexNet* e do modelo de referência baseado na arquitetura *Resnet*, treinada com a função de perda *triplet loss*, que se constitui no estado da arte para a realização desta tarefa. Metodologicamente, o escopo desta pesquisa é de abordagem quantitativa, visto que analisa métricas de desempenho computacional em tarefas de classificação como acurácia, precisão e *recall*.

Quanto à natureza é aplicada e de base tecnológica, já que gera conhecimento relacionado à análise do uso de redes neurais convolucionais aplicadas ao reconhecimento facial de animais. Em relação aos objetivos da pesquisa, é caracterizada como pesquisa descritiva, pois coleta e organiza dados sobre uma determinada realidade sem explicá-los, tendo a delimitação de técnicas, métodos, modelos e teorias para a coleta e compreensão dos dados (Triviños, 2011). Quanto aos procedimentos, essa pesquisa envolve pesquisa bibliográfica e experimental por conta de compreender a aplicação e avaliação prática de algoritmos de inteligência artificial para a identificação biométrica de cães, bem como a manipulação de variáveis e a verificação dos resultados sobre o objeto de estudo (Wazlawick, 2021).

Figura 1 - Etapas realizadas na pesquisa



Fonte: Elaborador pelo autor

Esta pesquisa foi arquitetada em cinco etapas (figura 1) que visam atingir o objetivo proposto, iniciando com a preparação da base de imagens, seguida pelo desenvolvimento das redes neurais utilizadas, e posteriormente pelo treinamento e testes dos modelos e sua otimização. Por último, a realização da análise dos resultados obtidos a fim de identificar a eficácia dos modelos.

### 3.1 BASE DE IMAGENS

A base de imagens utilizada para o treinamento de as duas arquiteturas utilizadas nesta pesquisa foi coletada por Mougeot, Li e Jia (2019), algumas imagens foram tiradas por meio de câmeras de fotografia e outras encontradas na internet, se encontra disponível na nota de rodapé<sup>7</sup> sob a licença *Creative Commons Attribution 4.0 International*. Boa parte das imagens pode ser encontrada em sites de adoção sem fins lucrativos, como por exemplo: Streunerhilfe<sup>8</sup>, Tiko<sup>9</sup>, Pfotenhilfe<sup>10</sup>, LaSPA<sup>11</sup>, Tieronlinee<sup>12</sup>, Animal-happyend6<sup>13</sup>. O conjunto de imagens possui 1393 classes de cães e totaliza 8363 imagens das quais possuem região de fundo reduzida com o recorte somente da área de interesse (figura 2).

Observa-se que a base de imagens é composta por diversas raças de cães como: Rottweiler, Boxer, Shar-pei, Utonagan, Pastor McNab, cães sem raça definida e muitas outras.

#### 3.1.1 Pré-processamento

A fim de mitigar o *overfitting* e aprimorar a capacidade de generalização dos modelos, foram empregadas algumas técnicas de pré-processamento e aumento de dados (*data augmentation*), conforme preconizado por (Shorten; Khoshgoftar, 2019). Para isso, foram aplicadas transformações geométricas e fotométricas, como rotação angular, espelhamento horizontal, zoom e deslocamento de canais. Além disso, para a realização de determinados testes, as imagens originalmente com resolução de 104x104 *pixels* foram redimensionadas para 224x224 *pixels*. No que se refere as transformações específicas, realizou-se aumento do zoom da imagem na faixa de 0,1, rotação angular leve de 8°, deslocamento de canais com vari-

<sup>7</sup><https://zenodo.org/records/12578449>

<sup>8</sup><https://www.streunerhilfe-bulgarien.de/>.

<sup>9</sup><https://www.tiko.or.at/de/tiere/>.

<sup>10</sup><https://www.pfotenhilfe.org/tiervermittlung/>.

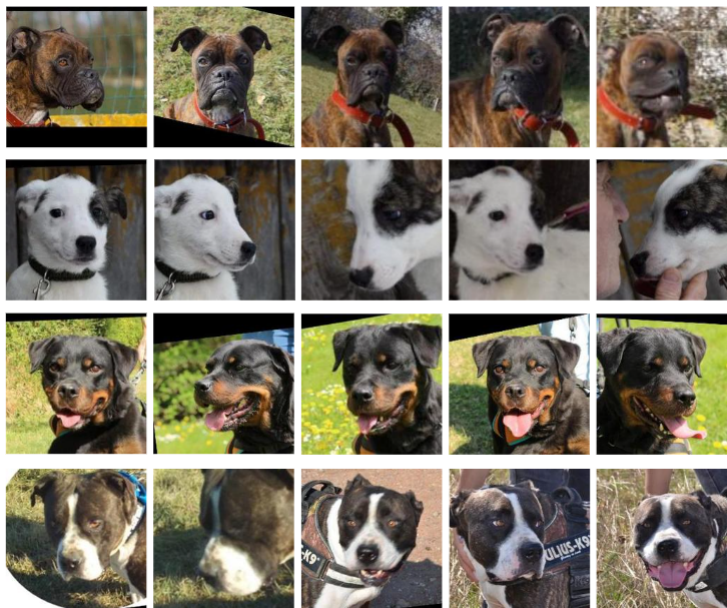
<sup>11</sup><https://www.la-spa.fr/>.

<sup>12</sup><https://www.tieronline.ch/>.

<sup>13</sup><https://www.animal-happyend.ch/>.

ação de 0,1 e espelhamento horizontal (somente para o modelo *AlexNet*). Essas transformações foram realizadas a fim de aumentar a diversidade do conjunto de treinamento por meio da geração de variações das imagens originais.

Figura 2 - Imagens de exemplo do conjunto



Fonte: Elaborado pelo autor

Para o modelo baseado na arquitetura *ResNet*, as imagens foram organizadas em trios, totalizando 10 mil grupos para o treinamento, o que correspondeu a 30 mil imagens. Desse conjunto, 1 mil trios (3 mil imagens) foram compostos por imagens aumentadas para diversificação do conjunto de dados.

No treinamento do modelo da arquitetura *AlexNet* modificada, realizados o aumento em até 4 vezes das imagens disponíveis, somando-se 33448.

### 3.2 DESENVOLVIMENTO DAS ARQUITETURAS UTILIZADAS

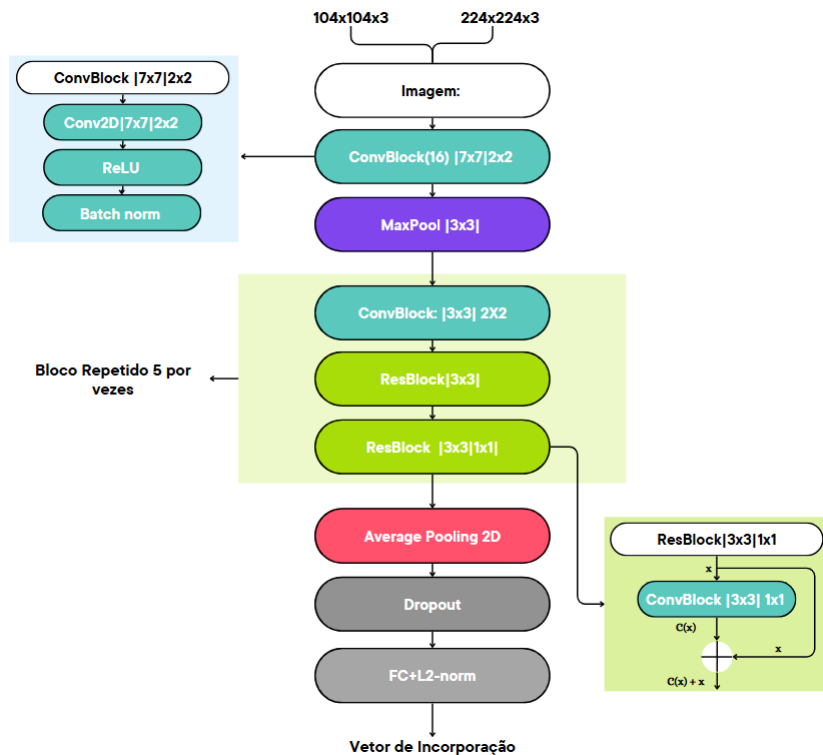
Para o desenvolvimento da arquitetura baseada no modelo *Resnet* como também do *Alexnet* modificado utilizou-se a linguagem de programação Python versão 3.11, no ambiente de desenvolvimento *Google Colaboratory Pro+*. A construção e o treinamento das redes neurais convolucionais foram realizados com o auxílio de bibliotecas especializadas em aprendizado profundo, citadas a seguir, que ofereceram suporte às operações de modelagem, treinamento, validação e avaliação dos modelos.

### 3.2.1 Arquitetura baseada no ResNet

O modelo utilizado nesta pesquisa foi inspirado na arquitetura *Resnet* proposta por Mougeot, Li e Jia (2019), que tem o intuito de evitar o overfitting em conjunto de imagens faciais de cães. A sua arquitetura é formada por camadas de *dropout* e camadas residuais que possibilitam a construção de redes mais profundas e eficazes na propagação do gradiente durante o processo de retropropagação, evitando assim o problema de desaparecimento do gradiente (figura 3).

A arquitetura implementada possui 63 camadas, distribuídas em um bloco central repetido sequencialmente 5 vezes, totalizando aproximadamente 5,8 milhões de parâmetros. A implementação do modelo foi realizada utilizando classes da biblioteca *keras* disponibilizada pelo framework *TensorFlow*. As principais classes empregadas foram *Input*, *Conv2D*, *Max-Pooling2D*, *Add*, *GlobalAveragePooling2D*, *Activation*, *Dropout*, *Flatten*, *Dense*, *Lambda* e *BatchNormalization*.

Figura 3 - Arquitetura baseada no *ResNet*



Fonte: Elaborado pelo autor adaptado de Yoon, So e Rhee (2021)

Para o treinamento do modelo com a função de perda tripla, do inglês *triplet loss*, é necessário definir três tipos de imagem: uma imagem âncora ( $x_a$ ), selecionada aleatoriamente a partir de uma classe específica do banco de imagens, uma imagem positiva ( $x_p$ ), pertencente à mesma

classe da âncora( $x_a$ ); e uma imagem negativa ( $x_n$ ), que é de uma classe diferente da classe da imagem âncora( $x_a$ ). O objetivo da perda tripla é minimizar a distância euclidiana entre os vetores de incorporação que a rede gerará da imagem âncora( $x_a$ ) e da imagem positiva( $x_p$ ), enquanto maximiza a distância entre os vetores da âncora e da imagem negativa ( $x_n$ ).

Adicionalmente, uma margem( $\alpha$ ) é utilizada como parâmetro de controle, assegurando que a distância entre os vetores da âncora e da imagem negativa seja, no mínimo( $\alpha$ ) unidades maior do que a distância entre os vetores da âncora e da imagem positiva. Como demonstra a seguinte fórmula:

$$\|f(x_a) - f(x_p)\|^2 + \alpha < \|f(x_a) - f(x_n)\|^2$$

A função de perda *loss* é ativada quando a soma da margem( $\alpha$ ) com a distância euclidiana entre os vetores de incorporação da imagem âncora( $x_a$ ) e da imagem positiva ( $x_p$ ) é maior do que a distância entre os vetores de incorporação da âncora( $x_a$ ) e da imagem negativa ( $x_n$ ).

Para o treinamento desta rede foi adotada a estratégia conhecida como *hard triplet*, descrita por Schroff, Kalenichenko e Philbin (2015) como uma abordagem que consiste em selecionar, para cada âncora, o vetor positivo( $x_p$ ) mais "difícil", ou seja, aquele cuja distância euclidiana em relação à âncora( $x_a$ ) é mais próxima da distância entre a âncora e a imagem negativa( $x_n$ ), o que se soma a maior distância euclidiana entre o vetor de incorporação da imagem âncora( $x_a$ ) e o vetor negativo( $x_n$ ) que está mais próximo à âncora ( $x_a$ ). O vetor positivo pode ser representado pela seguinte fórmula:

$$x_p \operatorname{argmax} \|f(x_a) - f(x_p)\|^2$$

E o negativo pela fórmula:

$$x_n \operatorname{argmin} \|f(x_a) - f(x_n)\|^2$$

A estratégia *hard triplet* foi aplicada dinamicamente durante o treinamento, sendo os triplets gerados a partir dos vetores de incorporação das imagens aumentadas, atualizados a cada 3 épocas. A margem( $\alpha$ ) na função de perda tripla foi definida em 0.3, conforme definido em Yoon, So e Rhee (2021) a fim de garantir uma separação significativa entre as amostras positivas e negativas. O modelo foi compilado utilizando o otimizador Adam, disponibilizado pela classe *optimizers* da biblioteca Keras. A taxa de

aprendizagem foi configurada de forma dinâmica ao longo do treinamento, de acordo com os valores apresentados na tabela 2.

Tabela 2 – Programação da taxa de aprendizado modelo baseado no *ResNet*

Épocas	Taxa de aprendizagem
39	0,001
12	0,0005
12	0,0003
6	0,0001
<b>Total</b>	<b>69</b>

Fonte: Elaborado pelo autor

### 3.2.2 Arquitetura *AlexNet* modificada

A arquitetura *AlexNet*, proposta por Krizhevsky, Sutskever e Hinton (2012), possui cinco camadas convolucionais, seguidas por três camadas completamente conectadas, sendo a última camada conectada a uma função de ativação softmax, que é responsável pela classificação nas diferentes classes do conjunto de imagens.

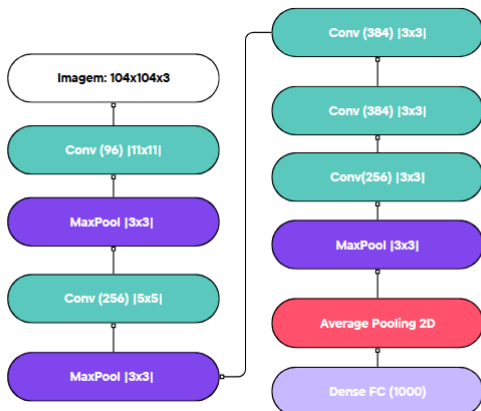
Entre os diferenciais desta arquitetura, tem-se a adoção da função de ativação de unidades lineares retificadas, Rectified Linear Unit (ReLUs), que acelera a convergência do treinamento; e, a utilização de camadas de e o pooling sobreposto que contribuem para a redução da dimensionalidade (Krizhevsky; Sutskever; Hinton, 2012).

A arquitetura modificada utilizada neste trabalho, com base no estudo Peruchi, Machado e Mattos (2024), possui uma camada a menos em relação à *AlexNet* original (figura 4). As duas últimas camadas totalmente conectadas foram substituídas por uma única camada do tipo *GlobalAveragePooling2D*, seguida por uma camada final completamente conectada, ajustada para corresponder ao número de classes do conjunto de imagens utilizado.

As camadas convolucionais originais são conectadas de forma a preservar os traços individuais extraídos nas camadas anteriores, possibilitando a combinação dessas características na etapa final de classificação. A camada *GlobalAveragePooling2D* realiza uma redução espacial no mapa de características e calcula a média para cada mapa de ativação. Dessa

forma, melhora capacidade de generalização do modelo, visto que considera uma imagem completa.

Figura 4 - Arquitetura *AlexNet* modificada



Fonte: Elaborado pelo autor adaptado de Peruchi, Machado e Mattos (2024)

Por fim, para o treinamento da rede foi definido o otimizador *Stochastic Gradient Descent (SGD)*, disponibilizado pela classe *optimizer* da biblioteca *Keras*, com taxa de aprendizado de 0.01. A função de perda utilizada foi a *categorical\_crossentropy*, que é apropriada para tarefas de classificação multiclasse.

### 3.3 TREINAMENTO E TESTES

O treinamento dos modelos foi realizado utilizando diferentes resoluções de imagens de entrada. Para ambas as arquiteturas, a baseada no *ResNet* e *AlexNet* modificada, foram utilizados conjuntos de imagens com resolução de 104x104 *pixels*. No modelo baseado no *ResNet* foram realizados experimentos com imagens redimensionadas para 224x224 *pixels*.

Os experimentos foram executados em máquinas virtuais disponibilizadas pelo Google Colaboratory, sendo que os realizados com imagens de 104x104 *pixels*, a máquina possuía a seguinte configuração: processador Intel(R) Xeon(R) CPU @ 2.20GHz, placa de vídeo NVIDIA A100-SXM4-40GB e 83.5GB de memória RAM. Já para os experimentos com imagens de 224x224 *pixels*, foi utilizada uma máquina virtual equipada com o mesmo processador, porém com 334.6GB de memória RAM e sem acelerador gráfico (GPU).

A base de imagens foi dividida na proporção de 80/20, sendo 80% das imagens para o treinamento e 20% para o teste, seguindo, conforme Geron (2019), o particionamento comum em experimentos supervisionados. No treinamento do modelo baseado no *ResNet* foram utilizadas

30.000 imagens e 3.000 para teste, e para o treinamento do modelo AlexNet foram utilizadas 33.448 imagens para treinamento e 8.363 imagens para teste. Todos os vetores das imagens, de acordo com Goodfellow, Bengio e Courville (2016), foram normalizados para o intervalo [0,1], a fim de padronizar a entrada da rede e acelerar a convergência do treinamento.

No treinamento do modelo AlexNet foi aplicado a técnica one-hot encoding que consiste em transformar cada rótulo de classe em um vetor binário, em que apenas um item é marcado com valor 1, indicando a classe correta, e os demais recebem 0. Essa técnica foi aplicada tanto para os rótulos de treinamento, quanto para os de teste, conforme Chollet (2021), esta é uma prática padrão em tarefas de classificação multiclasse, especialmente em redes neurais que utilizam a função de ativação softmax na saída. O teste do respectivo modelo foi realizado variando o número de épocas e o tamanho do lote, como mostra a tabela 4.

O modelo baseado no *ResNet* não realiza a classificação direta por classes, mas sim gera vetores de incorporação, de modo que a verificação é feita por meio da comparação entre pares de imagens. No processo de inferência, o modelo calcula a distância euclidiana entre os vetores gerados para o par de imagens e as classifica como pertencentes ao mesmo cão ou não, para isso utiliza um limiar fixo de 0.3, conforme a média de limiar utilizado por Yoon, So e Rhee (2021). Assim, se a distância entre os vetores for inferior a esse limiar, as imagens são consideradas do mesmo cão; caso contrário, de animais diferentes. Foram realizados testes variando a resolução da imagem de entrada e o tamanho dos lotes dos vetores de incorporação na saída (tabela 3).

Tabela 3 – Treinamento e testes com o modelo baseado no *ResNet*

<b>Resolução imagem de entrada</b>	<b>Tamanho dos lotes</b>
104x104px	32
104x104px	64
224x224px	64

Fonte: Elaborado pelo autor

A avaliação do desempenho dos modelos foi realizada por meio das métricas de avaliação quantitativa para modelos de classificação, as quais são: acurácia, precisão, *recall* e *F1-Score*.

A acurácia indica a proporção de itens classificados corretamente em relação ao número total de amostras avaliadas. A precisão verifica a

Tabela 4 – Treinamento e testes com o modelo *AlexNet* modificado

Tamanho dos lotes	Número de épocas
32	80
64	80
32	100
64	100

Fonte: Elaborado pelo autor

proporção de verdadeiros positivos entre todas as amostras classificadas como positivas pelo modelo. O *recall* avalia a capacidade do modelo de identificar corretamente os itens classificados como positivos. O *F1-Score* representa a média harmônica entre a precisão e o *recall*, sendo uma medida útil em casos de classes desbalanceadas (Powers, 2020).

Além disso, os testes foram realizados utilizando um subconjunto do conjunto de imagens, contendo imagens que não foram apresentadas ao modelo durante o treinamento, a fim de se avaliar a capacidade de generalização de cada uma das arquiteturas empregadas. Os procedimentos metodológicos adotados nesta pesquisa foram organizados de forma a permitir uma análise dos modelos gerados pelas arquiteturas *AlexNet* modificada e baseada no *ResNet* com *triplet loss* na tarefa de reconhecimento facial de cães.

A seleção da base de imagens, a definição das técnicas de pré-processamento, a configuração dos experimentos e a escolha das métricas de avaliação foram fundamentadas na literatura especializada da área. Os resultados experimentais e análise do desempenho de cada arquitetura são apresentados na sequência.

#### 4 RESULTADOS E DISCUSSÃO

Para a avaliação dos resultados experimentais obtidos pelos modelos na tarefa de reconhecimento facial de cães, foram conduzidos experimentos para medir as métricas citadas na sessão anterior, a fim de verificar o desempenho geral das arquiteturas na correta classificação das imagens. As tabelas 5 e 6 mostram os resultados obtidos pelos modelos baseados no *ResNet* e *AlexNet* modificado, respectivamente.

A partir dos dados apresentados na tabela 5, observa-se que o modelo baseado no *ResNet* apresentou pouca variação na acurácia entre os diferentes cenários testados, mesmo com a mudança da resolução da

imagem de entrada. O modelo variou a sua acurácia entre 0,806 e 0,817, apontando para a robustez do modelo frente as entradas de menor qualidade.

Tabela 5 – Resultados do modelo baseado no *ResNet*

Resolução da imagem de entrada	Tamanho dos lotes	Acurácia	Precisão	Recall	F1-Score
104x104px	32	0,806	0,9458	0,6573	0,7756
104x104px	64	0,817	0,9458	0,7018	0,7978
224x224px	64	0,8156	0,9246	0,6852	0,7871

Fonte: Elaborado pelo autor

Pode-se observar também que o modelo obteve valores de precisão variando de 0,9242 a 0,9458, sendo o melhor desempenho identificado nas imagens com resolução de 104x104 *pixels* e lote de 32 amostras. No referido cenário, o modelo foi mais eficaz em evitar falsos positivos, classificando corretamente as imagens previstas como positivas.

Os valores de recall desse modelo variaram de 0,6573 a 0,7018, como o melhor resultado para imagens com resolução de 104x104 *pixels* e lote de 64 amostras. Logo, percebe-se que o desempenho do modelo baseado no *ResNet* não melhora proporcionalmente ao aumento da resolução das imagens, contrariando a expectativa de que imagens maiores poderiam proporcionar uma melhor extração de características.

Os valores de *F1-Score* obtidos nos experimentos do modelo da referida arquitetura, variaram entre 0,7756 e 0,7978, o que denota pouca variação e um desempenho razoável, indicando consistência no comportamento do sistema. No entanto, não atinge níveis de desempenho compatíveis com aplicações que necessitam de elevada sensibilidade e precisão simultaneamente.

Tabela 6 – Resultados do modelo *AlexNet* modificado

Tamanho dos lotes	Número de épocas	Acurácia	Precisão	Recall	F1-Score
32	80	0,9777	0,9806	0,9777	0,9777
64	80	0,9782	0,9812	0,9782	0,9781
32	100	0,9728	0,9770	0,9728	0,9725
64	100	0,9770	0,9801	0,9770	0,9769

Fonte: Elaborado pelo autor

Os resultados apresentados na tabela 6, indicam que o modelo *AlexNet* modificado obteve acurácias superiores em comparação ao modelo baseado no *ResNet*, variando entre 0,9728 e 0,9782. Apesar das alterações no tamanho dos lotes e no número de épocas de treinamento, a

acurácia manteve-se, indicando que a arquitetura proposta apresenta capacidade de generalização. Assim, o modelo se apresentou menos sensível as configurações de parâmetros testadas.

Percebe-se também que este modelo obteve valores elevados de precisão, variando de 0,9770 a 0,9812, superando os resultados obtidos pelo modelo baseado no *ResNet* (tabela 5). A melhor precisão aconteceu no cenário de teste com lote de 64 imagens e 80 épocas de treinamento. Isso mostra que a arquitetura *AlexNet* modificada foi mais eficaz em minimizar falsos positivos, classificando corretamente os exemplos previstos como pertencentes à classe positiva.

Pode-se perceber que o modelo no *AlexNet* modificado obteve valores superiores de recall em comparação à arquitetura baseada no *ResNet*. Os resultados variaram de 0,9728 a 0,9782, mesmo diante da variação do número de épocas de treinamento e do tamanho dos lotes. Assim, pode-se dizer que o modelo apresenta alta sensibilidade, sendo capaz de encontrar corretamente as ocorrências positivas no conjunto de imagens. Este comportamento é fundamental em aplicações de reconhecimento facial, pois a omissão de verdadeiros positivos (falsos negativos) pode comprometer a confiabilidade.

Além disso, os resultados apresentados na tabela 6 para a métrica F1-Score demonstram que o modelo *AlexNet* modificado obteve valores elevados variando entre 0,9725 e 0,9781. Assim, indica que o modelo foi capaz de identificar corretamente as classes, minimizando os falsos positivos e negativos.

A partir da análise das métricas utilizadas para a avaliação dos modelos baseado na arquitetura *ResNet* com *triplet loss* e no *AlexNet* modificado, observa-se em todos os experimentos realizados e cenários, um desempenho superior do *AlexNet* modificado.

Tabela 7 – Melhores Resultados

Métrica	Modelo baseado <i>ResNet</i>	Modelo <i>AlexNet</i> modificado
Acurácia	0,817	0,9782
Precisão	0,9458	0,9812
<i>Recall</i>	0,7018	0,9782
<i>F1-Score</i>	0,7871	0,9781

Fonte: Elaborado pelo autor

A referida arquitetura apresentou uma acurácia mais alta, indi-

cando uma maior taxa de classificações corretas no geral, tendo-se os valores de precisão e de *recall* também superiores, evidenciando a capacidade do modelo em evitar falsos positivos e identificar corretamente as ocorrências positivas, respectivamente. No que se refere a métrica *F1-Score*, os valores obtidos pelo modelo superiores a 0,97 reforçam o desempenho adequado do modelo na tarefa de reconhecimento de padrões. A tabela 7 mostra os melhores resultados para ambos modelos.

Os resultados da pesquisa apontam na direção que para o reconhecimento facial de cães, com o conjunto de imagens empregadas e as condições experimentais definidas no presente estudo, a arquitetura *AlexNet* modificada foi mais eficiente, precisa e confiável, contribuindo para um melhor com menor risco de *overfitting* e eficaz na generalização dos padrões.

Considerando os resultados obtidos nesta pesquisa, na avaliação dos modelos de redes neurais convolucionais, a arquitetura do modelo *AlexNet* modificado obteve um melhor desempenho se comparado com a maioria dos trabalhos correlatos, sendo inferior somente ao modelo *Alexnet* utilizando o aprendizado por transferência para o reconhecimento facial de humanos no banco de dados dos Estudantes da *University Teknikal Malaysia Melaka* proposto por Ramlee et al. (2020). Por outro lado a arquitetura baseada no *ResNet* empregada nessa pesquisa, obteve um desempenho superior somente aos modelos propostos por Moreira et al. (2017).

## 5 CONCLUSÃO

A identificação individual de cães representa um desafio significativo para a implementação de políticas eficientes relacionadas ao controle populacional, saúde pública e bem-estar animal, especialmente em contextos urbanos, onde métodos tradicionais de identificação, como coleiras e microchips, apresentam limitações. Diante disso, nesta pesquisa foram implementadas e avaliadas duas abordagens baseadas em redes neurais convolucionais para o reconhecimento facial de cães: uma arquitetura baseada no modelo *ResNet* e outra utilizando uma versão modificada do *AlexNet*.

No desenvolvimento do estudo, as arquiteturas foram submetidas a treinamentos utilizando um banco de imagens faciais de cães com ampla diversidade de classes e características visuais. Após essa etapa, realizou-se a avaliação dos modelos por meio de uma série de experimentos com diferentes configurações de entrada, número de épocas e tamanho

dos lotes. Contudo, uma das dificuldades enfrentadas foi quanto a resolução das imagens, houve a necessidade de trabalhar com imagens de baixa resolução, devido à restrita capacidade computacional disponível para processar um volume grande de imagens em alta qualidade.

As avaliações foram realizadas por meio de métricas amplamente reconhecidas na literatura, incluindo acurácia, precisão, *recall* e *F1-score*. Os resultados obtidos demonstraram consistentemente a superioridade do modelo com arquitetura *AlexNet* modificada em relação ao modelo baseado na *ResNet*. Especificamente, a arquitetura *AlexNet* modificada obteve em seu melhor cenário uma acurácia de 97%, enquanto a arquitetura baseada na *ResNet* alcançou uma acurácia máxima de 81%.

Como futuros direcionamentos desta pesquisa, sugere-se a exploração de novas arquiteturas, particularmente aquelas que tenham demonstrado alto desempenho na classificação de imagens, investigando sua aplicabilidade em conjunto com técnicas de aprendizado por transferência para a identificação facial de cães. Adicionalmente, propõe-se o desenvolvimento de uma aplicação prática que permita a identificação efetiva dos cães, visando implementar medidas eficazes para controle e rastreabilidade.

## REFERÊNCIAS

ALBAWI, S.; MOHAMMED, T. A.; AL-ZAWI, S. Understanding of a convolutional neural network. In: **2017 International Conference on Engineering and Technology (ICET)**. [S.l.: s.n.], 2017. p. 1–6.

ARRUDA, E. C.; GARCIA, R. C. M.; OLIVEIRA, S. T. **Bem-estar dos cães de abrigos municipais no estado do paran , brasil, segundo o protocolo shelter quality**. Arquivo Brasileiro de Medicina Veterin ria e Zootecnia, 2020, v. 72, n. 2, p. 346–354.

BARBER, D. **Bayesian Reasoning and Machine Learning**. Cambridge ; New York: Cambridge University Press, 2012.

CHOLLET, F. **Deep Learning com Python**. 2. ed. S o Paulo: Novatec, 2021.

COSTA, J. R. T. **Aplic o de Deep Learning para an lise de sentimentos a partir de imagens**. 55 p. Monografia (Bacharel em Engenharia de Software) — Universidade de Bras lia, Bras lia, DF, maio 2022. Dispon vel em: <[https://bdm.unb.br/bitstream/10483/34021/1/2022\\_JoberthRogersTavaresCosta.pdf](https://bdm.unb.br/bitstream/10483/34021/1/2022_JoberthRogersTavaresCosta.pdf)>.

COSTA, M. **“Tem protetor passando fome com os animais”, desabafa dona de abrigo**. 2022. Dispon vel em: <<https://www.correiobraziliense.c>

[om.br/brasil/2022/04/4998353-tem-muito-protetor-passando-fome-com-os-animais-desabafa-dona-de-abrigo.html](https://om.br/brasil/2022/04/4998353-tem-muito-protetor-passando-fome-com-os-animais-desabafa-dona-de-abrigo.html).

DENG, J. et al. Arcface: Additive angular margin loss for deep face recognition. In: **2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)**. [S.l.: s.n.], 2019. p. 4685–4694.

DENG, J. et al. Arcface: Additive angular margin loss for deep face recognition. In: **2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)**. [S.l.: s.n.], 2019. p. 4685–4694.

GERON, A. **Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems**. 2. ed. Sebastopol: O'Reilly Media, 2019.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep Learning**. Cambridge, MA: MIT Press, 2016.

JANG, D.-H. et al. **Dog identification method based on muzzle pattern image**. Applied Sciences (Basel, Switzerland), 2020, v. 10, n. 24, p. 8994.

KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. **ImageNet Classification with Deep Convolutional Neural Networks**. 2012. Disponível em: <http://code.google.com/p/cuda-convnet/>.

KUMAR, S.; SINGH, S. K. **Biometric recognition for pet animal**. Journal of Software Engineering and Applications, 2014, v. 7, n. 5, p. 470–482.

KUMAR, S.; SINGH, S. K. **Monitoring of pet animal in smart cities using animal biometrics**. Future Generation Computer Systems, 2018, Elsevier B.V., v. 83, p. 553–563. ISSN 0167739X.

KUMAR, S. et al. **Real-time recognition of cattle using animal biometrics**. Journal of Real-Time Image Processing, 2017, Springer Verlag, v. 13, p. 505–526. ISSN 18618200.

MODA, A. B. **Pets abandonados: após adoção aumentar no início da pandemia, cães e gatos são deixados com flexibilização: são acolhidos por abrigos quando deixam de servir como suporte emocional para os isolados**. O GLOBO, 2023. Disponível em: <https://oglobo.globo.com/brasil/pets-abandonados-apos-adocao-aumentar-no-inicio-da-pandemia-caes-gatos-sao-deixados-com-flexibilizacao-25341144>.

MOREIRA, T. P. et al. **Where is my puppy? retrieving lost dogs by facial features**. Multimedia Tools and Applications, 2017, Springer New York LLC, v. 76, p. 15325–15340. ISSN 15737721.

MOTA-ROJAS, D. et al. **Abandonment of dogs in latin america: Strategies and ideas**. Veterinary World, 2021, p. 2371–2379.

MOUGEOT, G.; LI, D.; JIA, S. A deep learning approach for dog face verification and recognition. In: NAYAK, A. C.; SHARMA, A. (Ed.). **PRICAI 2019: Trends in Artificial Intelligence**. Cham: Springer International Publishing, 2019. p. 418–430. ISBN 978-3-030-29894-4.

PERUCHI, M. S.; MACHADO, R.; MATTOS, M. C. D. Trabalho de conclusão de curso (Bacharel no Curso de Ciência da Computação da Universidade), **Reconhecimento e classificação de espécies de peixes utilizando redes neurais convolucionais**. Criciúma: [s.n.], 2024. Disponível em: <<http://repositorio.unesc.net/bitstream/1/10984/1/Mateus%20Silveira%20Peruchi.pdf>>.

POWERS, D. M. W. **Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation**. 2020.

RAMLEE, R. A. B. et al. **Transfer learning using alexnet convolutional neural network for face recognition**. International Journal of Innovative Technology and Exploring Engineering, 2020, v. 9, p. 285–294. ISSN 22783075. Disponível em: <<https://www.ijtee.org/portfolio-item/K77760991120/>>.

SCHROFF, F.; KALENICHENKO, D.; PHILBIN, J. **FaceNet: A unified embedding for face recognition and clustering**. IEEE, 2015. Disponível em: <<https://doi.org/10.1109%2Fcvpr.2015.7298682>>.

SERMANET, P. et al. **Overfeat: Integrated recognition, localization and detection using convolutional networks**. 2013. Disponível em: <<http://arxiv.org/abs/1312.6229>>.

SHORTEN, C.; KHOSHGOFTAR, T. M. **A survey on image data augmentation for deep learning**. Journal of Big Data, 2019, v. 6, n. 1, p. 1–48. Accessed: 2025-06-03. Disponível em: <<https://doi.org/10.1186/s40537-019-0197-0>>.

SINGH, S. et al. Transfer learning: Convolutional neural network-alexnet achieving face recognition. In: **2022 2nd Asian Conference on Innovation in Technology (ASIANCON)**. [S.l.: s.n.], 2022. p. 1–4.

SMITH, L. M. et al. **The effectiveness of dog population management: A systematic review**. Animals, 2019, v. 9, n. 12, p. 1020.

SOUZA, C. T. B. D. M. F. J. D. L. **Registro geral de animais (rga): um sistema para o registro e identificação de animais de companhia**. 15 p. Dissertação (Trabalho de conclusão de curso (Bacharel em Tecnologia de informação e comunicação)) — Universidade Federal de Santa Catarina, Araranguá, 2017.

TAIGMAN, Y. et al. Deepface: Closing the gap to human-level performance in face verification. In: **2014 IEEE Conference on Computer Vision and Pattern Recognition**. [S.l.: s.n.], 2014. p. 1701–1708.

TRIVIÑOS, A. N. S. **Introdução à pesquisa em ciências sociais: a pesquisa qualitativa em educação**. São Paulo: Atlas, 2011.

WANG, H. et al. **Cosface: Large margin cosine loss for deep face recognition**. 2018. Disponível em: <<http://arxiv.org/abs/1801.09414>>.

WAZLAWICK, R. S. **Metodologia de pesquisa para ciência da computação**. Rio de Janeiro: LTC, 2021.

YASHAVANTH, T. R.; SURESH, M. Multimodal biometric system using alex net model. In: **2023 International Conference on Smart Systems for applications in Electrical Sciences (ICSSSES)**. [S.l.: s.n.], 2023. p. 1–4.

YOON, B.; SO, H.; RHEE, J. **A methodology for utilizing vector space to improve the performance of a dog face identification model**. Applied Sciences, 2021, v. 11, n. 5. ISSN 2076-3417. Disponível em: <<https://www.mdpi.com/2076-3417/11/5/2074>>.