**APLICAÇÃO DE DATA MINING PARA PREVISÃO DE RAÇAS DE CACHORROS**

DIANE S. REIS, ELENINHA A. SANTOS, BRUNO F. L. SOUSA, VITOR F. PEREIRA, GABRIEL S. EZEQUIEL

Faculdade Impacta de Tecnologia

São Paulo, São Paulo, Brazil

Email: {diane.reis,eleninha.santos,bruno.sousa,vitor.fernandes, gabriel.ezequiel}@aluno.faculdadeimpacta.com.br

**INTRODUÇÃO**

O ser humano vem moldando os canídeos a mais de 36 mil anos, gerando diferentes variações e raças para auxiliar os homens [1]. Como reflexo dessas modificações, observamos uma infinidade de raças caninas com as mais variáveis finalidade como companhia, boiadeiros, caça, guarda entre outras centenas de objetivos.

E todos esses anos de domesticação, levaram a crescente popularidade dos animais de estimação, em especial os cães. Isso têm impulsionado um próspero mercado voltado para pets nas últimas décadas. [2] Estima-se que existem um total de 149,6 milhões de pets no Brasil, sendo 58,1 milhões de cães e 27,1 milhões de gatos.Tendo em vista esse grande número de animais de estimação, o segmento pet mostra crescente evolução e grandes chances de seguir se consolidando como um ramo forte. Seguindo nessa mesma tendência, a diversidade de raças caninas disponíveis para compra, aumentou significativamente.

Com esse aumento, a identificação precisa das diferentes raças tornou-se um desafio, para os aspirantes a proprietários. Esta dificuldade de reconhecimento e distinção de raças caninas pode levar a escolhas equivocadas de aquisição, que não consideram as características comportamentais, físicas e de cuidados específicos de cada raça.

A capacidade de identificar com precisão a raça de um cão a partir de imagens pode auxiliar significativamente na promoção de compras responsáveis, alinhando as características do animal com o estilo de vida, expectativas e necessidades dos futuros proprietários. Além disso, essa abordagem pode aprimorar a eficiência dos canis ao informar as carências de cada raça para as famílias interessadas, podendo assim diminuir o número de abandono, devoluções e aumentar o bem-estar animal.

É aqui que o reconhecimento de raças caninas por meio de técnicas avançadas de processamento de imagens emerge como uma ferramenta de valor inestimável. Ao permitir a identificação precisa das raças de cães com base em características visuais.

[1]<https://www.livrosabertos.sibi.usp.br/portaldelivrosUSP/catalog/view/326/287/1212>

[2]<https://sebrae.com.br/sites/PortalSebrae/ufs/al/artigos/crescimento-do-mercado-pet-e-oportunidade-de-negocio,021731b7fe057810VgnVCM1000001b00320aRCRD>

**CARACTERIZAÇÃO DO PROBLEMA**

Com o crescimento de números de pets no país [3] e desse mercado como um todo, surgem oportunidades econômicas substanciais para empresas que atendem às necessidades dos proprietários de animais de estimação.

Tabela1. Instituto Pet Brasil - População de animais de estimação no Brasil - 2022 - Em milhões

| **Ano** | **Cães** | **Gatos** | **Aves** | **Peixes** | **Rep. Peq.Mam.** | **Total** |
| --- | --- | --- | --- | --- | --- | --- |
| **2018** | 54,2 | 23,9 | 39,8 | 19,1 | 2,3 | 139,3 |
| **2019** | 55,1 | 24,7 | 40 | 19,4 | 2,4 | 141,6 |
| **2020** | 55,9 | 25,6 | 40,4 | 19,9 | 2,5 | 144,31 |
| **2021** | 58,1 | 27,1 | 41 | 20,8 | 2,5 | 149,53 |

Produtos e serviços que vão desde alimentos, acessórios, serviços de saúde, hospedagem, até cuidados com o comportamento animal têm testemunhado um aumento na demanda. Isso cria oportunidades para inovações, como o nosso projeto propõe.

No entanto, o crescimento exponencial do mercado pet também trouxe à tona desafios críticos, especialmente no que diz respeito à seleção responsável de animais de estimação.

Muitos adotantes acabam escolhendo seus pets pelo aspecto físico, porte ou moda, m vez de considerar as necessidades específicas da raça em relação ao seu estilo de vida. E essa análise superficial preliminar pode acabar gerando problemas na convivência do animal com seu dono, as raças podem apresentar beleza mas precisam de cuidados específicos que raramente são levados em consideração, como por exemplo características comportamentais, físicas e de cuidados particulares de cada raça.

Esta dificuldade de reconhecimento e distinção de raças caninas pode levar a frustrações tanto para o dono quanto para o animal, que pode acabar afetando seu comportamento e causando problemas sérios. Nesse caso, a possibilidade de devolução do animal vira uma opção, gerando perda no valor investido pelo dono e desapontamento de ambas as partes[4].

Nesse contexto, a capacidade de identificar com precisão a raça de um cão a partir de imagens se destaca como uma solução valiosa. Ao fornecer aos futuros proprietários informações sobre as características comportamentais, físicas e de cuidados específicos de cada raça, o reconhecimento de raças caninas ajuda a alinhar de forma mais precisa as expectativas dos proprietários com as necessidades do animal. Isso não só melhora a convivência, mas também reduz o número de abandono de animais e devoluções.

[3]<https://www.gov.br/agricultura/pt-br/assuntos/camaras-setoriais-tematicas/documentos/camaras-setoriais/animais-e-estimacao/2019/26deg-ro-1/14-10-ipb_mercado_pet_resultados_2018_draft2.pdf>

<https://www.gov.br/agricultura/pt-br/assuntos/camaras-setoriais-tematicas/documentos/camaras-tematicas/insumos-agropecuarios/anos-anteriores/ibge-populacao-de-animais-de-estimacao-no-brasil-2013-abinpet-79.pdf>

[4]<https://jornal.usp.br/atualidades/modismo-com-racas-de-cachorro-gera-abandono-de-animais-e-inconveniencias-as-familias/>

**DESCRIÇÃO DA BASE DE DADOS**

A base de dados em questão abrange uma vasta variedade de raças, compreendendo um total de 120 diferentes tipos, englobando tanto aquelas que são consideradas as mais populares e comuns, quanto as mais incomuns e exóticas, totalizando assim 20.575 imagens de cães. Essas imagens foram coletadas a partir de múltiplas fontes, o que assegura uma representação amplamente diversificada de ambientes, condições de iluminação, ângulos de captura e contextos. Isso proporciona uma riqueza de informações e um conjunto de dados verdadeiramente abrangente para qualquer tipo de análise ou aplicação relacionada a cães. Possui também características notáveis: imagens de alta resolução, capturando detalhes cruciais que variam de características visuais a texturas de pelos.

Não foi necessário a manipulação de variáveis, pois a base trata-se somente de imagens, sendo assim, analisamos a imagem diretamente, sem passar por qualquer alteração, por fim, essa base encontra-se hospedada no site Kaggle[5].

[5]<https://www.kaggle.com/datasets/jessicali9530/stanford-dogs-dataset>

**APLICAÇÃO DAS TÉCNICAS**

Neste artigo, vamos explorar a criação de um sistema capaz de identificar a raça de um cão em uma imagem. Ao longo do nosso projeto, utilizamos a biblioteca TensorFlow em conjunto com a interface Keras para desenvolver, treinar e avaliar o nosso modelo de reconhecimento de raças caninas em fotografias.

Com o auxílio do Keras, preparamos cuidadosamente nosso conjunto de treinamento, o que inclui a adaptação das imagens para um formato compreensível para redes neurais, além da aplicação de transformações, como rotação, zoom e inversão horizontal. Essas transformações desempenham um papel fundamental na criação de um conjunto de dados diversificado e coeso, essencial para o treinamento eficaz do nosso modelo.

A estrela do nosso projeto é a rede neural convolucional, conhecida como CNN, uma poderosa ferramenta para análise de imagens. As CNNs, são um tipo de rede neural profunda especialmente projetado para resolver desafios em visão computacional. Essas redes são altamente eficazes no reconhecimento de imagens, segmentação de objetos e detecção de características visuais.

O treinamento é uma fase crítica do projeto. Utilizamos um conjunto de imagens representando diversas raças de cães para ensinar o modelo a reconhecê-las. Esse treinamento ocorre em várias etapas, permitindo ao modelo aprender progressivamente mais sobre as distintas raças caninas.

Finalmente, demonstramos como utilizar o modelo treinado para identificar a raça de um cão em uma nova imagem. Esse processo envolve a preparação da imagem, a execução da previsão pelo modelo e, por fim, a identificação da raça do cão representado na imagem.

Este projeto exemplifica de forma prática como é possível desenvolver um sistema de reconhecimento de imagens utilizando técnicas de aprendizado profundo. Desde a organização dos dados até o treinamento e avaliação do modelo, cada etapa desempenha um papel fundamental na construção de um sistema preciso e confiável. Em muitos casos, é necessário ajustar e aprimorar o sistema para otimizá-lo para aplicações do mundo real. As técnicas empregadas neste projeto podem ser aplicadas a uma variedade de problemas de reconhecimento de imagens.

**RESULTADOS**

Os resultados fornecidos nesta análise foram obtidos a partir de um conjunto de dados composto por 302 imagens, distribuídas entre duas raças distintas: Chihuahua e Doberman. Para avaliar o desempenho do modelo, utilizamos um gráfico representativo, que nos proporcionou uma visão ampla das tendências ao longo do treinamento. A **Figura 1** ilustra a acurácia alcançada em cada época (epoch), atingindo um notável patamar de 86% já na décima época.

Esses resultados destacam não apenas a eficácia do modelo proposto, mas também ressaltam a rapidez com que ele conseguiu aprender e se adaptar aos dados fornecidos para este cenário.

Gráfico, Gráfico de linhas

Descrição gerada automaticamente

Figura 1. Gráfico Acurácia inicial

Os resultados derivados de uma extensa coleção de 20.580 imagens, abrangendo 120 raças de cães diferentes, revelaram um desempenho aquém do esperado para o modelo em questão. A **Figura 2** fornece uma representação visual das taxas de acurácia observadas ao longo de cada época (epoch), revelando um desafio significativo no processo de aprendizado. O modelo atingiu apenas 19% de acurácia na última época, indicando uma dificuldade notável em distinguir entre as numerosas raças de cães presentes no conjunto de dados.

Gráfico, Gráfico de linhas

Descrição gerada automaticamente

Figura 2. Gráfico Acurácia final

**CONSIDERAÇÕES FINAIS**

Apesar dos desafios encontrados, este estudo proporciona insights valiosos sobre as limitações do modelo, incentivando futuras investigações para melhorar a capacidade de redes neurais em reconhecimento de imagens em conjuntos de dados extremamente heterogêneos. Estes resultados, embora modestos, são um passo importante na compreensão das complexidades associadas à classificação de raças de cães em grandes conjuntos de dados e destacam a necessidade contínua de avanços em técnicas de aprendizado de máquina para enfrentar tais desafios.

**REFERÊNCIAS BIBLIOGRÁFICAS**