

UNICESUMAR – CENTRO UNIVERSITÁRIO DE MARINGÁ
CENTRO DE CIÊNCIAS EXATAS TECNOLÓGICAS E AGRÁRIAS
CURSO DE GRADUAÇÃO EM ENGENHARIA DE *SOFTWARE*

***DEEP LEARNING* APLICADO AO RECONHECIMENTO DE IMAGEM PARA
PRESERVAÇÃO DE ANIMAIS SILVESTRES: UMA PESQUISA EXPLORATÓRIA**

BEATRIZ DE ARAUJO FERRAZ

MARINGÁ – PR

2022

Beatriz de Araujo Ferraz

***DEEP LEARNING* APLICADO AO RECONHECIMENTO DE IMAGEM PARA
PRESERVAÇÃO DE ANIMAIS SILVESTRES: UMA PESQUISA EXPLORATÓRIA**

Artigo apresentado ao Curso de Graduação em Engenharia de Software da UNICESUMAR – Centro Universitário de Maringá como requisito parcial para a obtenção do título de Bacharel(a) em Engenharia de *Software*, sob a orientação da Profa. Dra. Iara Carnevale de Almeida.

MARINGÁ – PR

2022

FOLHA DE APROVAÇÃO
BEATRIZ DE ARAUJO FERRAZ

***DEEP LEARNING* APLICADO AO RECONHECIMENTO DE IMAGEM PARA
PRESERVAÇÃO DE ANIMAIS SILVESTRES: UMA PESQUISA EXPLORATÓRIA**

Artigo apresentado ao Curso de Graduação em Engenharia de Software da UNICESUMAR –
Centro Universitário de Maringá como requisito parcial para a obtenção do título de
Bacharel(a) em Engenharia de *Software*, sob a orientação da Profa. Dra. Iara Carnevale de
Almeida.

Aprovado em: ____ de _____ de ____.

BANCA EXAMINADORA

Nome do professor – (Titulação, nome e Instituição)

Nome do professor – (Titulação, nome e Instituição)

Nome do professor – (Titulação, nome e Instituição)

DEEP LEARNING APLICADO AO RECONHECIMENTO DE IMAGEM PARA PRESERVAÇÃO DE ANIMAIS SILVESTRES: UMA PESQUISA EXPLORATÓRIA

Beatriz de Araujo Ferraz

RESUMO

A preservação da fauna silvestre é indispensável para manter o equilíbrio do ecossistema. O *Deep Learning* (DL) equivale a um recurso recente que tem sido utilizado para apoiar essa tarefa. Diante do exposto, esta pesquisa tem como objetivo geral apresentar os recursos necessários para efetuar o reconhecimento de imagens por meio do DL, de forma a identificar, catalogar e contabilizar animais silvestres. Para tanto, foi realizada uma pesquisa bibliográfica exploratória com abordagem qualitativa. O presente trabalho apresenta, portanto, estudos que indicam como o DL é empregado na identificação, na catalogação e na contabilização de animais. A partir disso, conclui-se que aplicar o reconhecimento de imagem por meio de DL para automatizar as tarefas mencionadas é vantajoso devido, principalmente, à redução de tempo e de esforço humano poupados.

Palavras-chave: Aprendizado profundo. Inteligência artificial.

DEEP LEARNING APPLIED IN IMAGE RECOGNITION FOR WILDLIFE CONSERVATION: AN EXPLORATORY RESEARCH

ABSTRACT

The preservation of wildlife is indispensable to maintain the balance of the ecosystem. Deep Learning (DL) is a recent resource that has been used to support this task. Given the above, this research has as a general objective to present the resources needed to perform image recognition through DL, in order to identify, catalog, and account for wild animals. To this end, an exploratory bibliographic research with a qualitative approach was carried out. The present work presents, therefore, studies that indicate how DL is used in the identification, cataloging, and accounting of animals. From this, it is concluded that applying image recognition through DL to automate the mentioned tasks is advantageous mainly due to the reduction of time and human effort saved.

Keywords: Deep learning. Artificial intelligence.

SUMÁRIO

1 INTRODUÇÃO	5
1.1 OBJETIVO GERAL	6
1.2 OBJETIVOS ESPECÍFICOS	6
2 REFERENCIAL TEÓRICO	6
2.1 RECONHECIMENTO DE IMAGEM	6
2.2 <i>DEEP LEARNING</i>	7
3 METODOLOGIA	8
4 DESENVOLVIMENTO	9
4.1 PRESERVAÇÃO ANIMAL APOIADA PELO <i>DEEP LEARNING</i>	13
5 CONCLUSÕES E TRABALHO FUTURO	22
REFERÊNCIAS	23

1 INTRODUÇÃO

Muitos animais estão ameaçados de extinção devido à ação humana no meio ambiente. Segundo Conde *et al.* (2019), a taxa de extinção de animais vertebrados no último século se apresenta cem vezes mais alta do que seria considerado natural. Diante disso, é importante monitorar os grupos de forma a verificar o crescimento ou declínio da população de uma determinada espécie no seu ambiente natural.

Recursos tecnológicos têm sido utilizados para monitorar os animais, como câmeras localizadas em pontos estratégicos nas florestas e savanas. Contudo, Norouzzadeh *et al.* (2018) indicam que o grande número de imagens geradas impossibilita que a análise completa dessas imagens possa ser realizada por seres humanos.

A Inteligência Artificial (em inglês, *Artificial Intelligence* – AI) pode ser entendida como a ciência de treinar máquinas para executar tarefas humanas de forma autônoma. De acordo com Goodfellow, Bengio e Courville (2016), desde o princípio da AI, os computadores são capazes de resolver problemas considerados complexos para os seres humanos, desde que esses problemas estejam descritos com exatidão, por exemplo, um conjunto de regras matemáticas.

Um dos desafios da AI é fazer com que um computador seja capaz de solucionar problemas que um ser humano solucionaria intuitivamente, como reconhecer rostos em uma imagem. Esse processo de reconhecimento é intuitivo (ou automático) e seus passos são difíceis para sistematização. Tal dificuldade sugere, então, que os sistemas de AI precisam ter a capacidade de aprender sozinhos (GOODFELLOW; BENGIO; COURVILLE, 2016).

Ademais, uma das subáreas da AI foca no Aprendizado de Máquina (em inglês, *Machine Learning* – ML). Trata-se de uma tecnologia que confere aos computadores a capacidade de aprender com apoio de especialistas humanos. Segundo Goodfellow, Bengio e Courville (2016), para que o aprendizado ocorra, o computador deve ser capaz de extrair informações de dados puros e, não obstante, identificar padrões nessas informações. Todavia, conforme postula Burkov (2019), a maioria das técnicas utilizadas no ML carece de supervisão humana para que sejam indicados os padrões iniciais. Essa necessidade se torna um empecilho para que o ML alcance níveis satisfatórios de autonomia.

O Aprendizado Profundo (em inglês, *Deep Learning* – DL) é uma subárea do *Machine Learning* que permite autonomia de aprendizado à máquina. Uma das soluções do DL é o reconhecimento de imagem que, conforme expõem Amirian *et al.* (2018), consiste na capacidade de um computador identificar corretamente elementos contidos em imagens.

Esta pesquisa busca, portanto, responder à seguinte questão-problema: “Como o reconhecimento de imagens por meio de *Deep Learning* é utilizado para identificar, catalogar e contabilizar animais selvagens na natureza?”.

1.1 OBJETIVO GERAL

O objetivo geral desta pesquisa é apresentar os recursos necessários para efetuar o reconhecimento de imagens por meio de *Deep Learning*, de forma a identificar, catalogar e contabilizar animais silvestres.

1.2 OBJETIVOS ESPECÍFICOS

Para alcançar o objetivo geral proposto, foram definidos os objetivos específicos evidenciados a seguir.

1. Compreender conceitos principais sobre reconhecimento de imagens e *Deep Learning*.
2. Compreender conceitos principais sobre como ocorre o reconhecimento de imagens utilizando *Deep Learning*.
3. Verificar estudos sobre aplicações que utilizem reconhecimento de imagens por meio de *Deep Learning*.
4. Determinar os recursos necessários para catalogar, identificar e contabilizar animais selvagens na natureza.

2 REFERENCIAL TEÓRICO

A presente seção pretende trazer os conceitos básicos para que este estudo pudesse ser realizado: o conceito sobre reconhecimento de imagem e, na sequência, sobre *Deep Learning*.

2.1 RECONHECIMENTO DE IMAGEM

Segundo Amirian *et al.* (2018), o reconhecimento de imagem tem um papel muito importante para a ciência e para a indústria. A capacidade de um computador compreender o

conteúdo de uma imagem é o que possibilita a existência de carros autônomos, geração automática de legendas de imagem, colorização automática de imagens e mais aspectos (AMIRIAN *et al.*, 2018).

De acordo com Sejnowski (2018), uma das primeiras abordagens utilizadas para o reconhecimento de imagem por meio de visão computacional consistia em extrair um *template* de um objeto exibido em uma imagem e, posteriormente, testar se o *template* correspondia aos *pixels* de outra imagem na qual o mesmo objeto estava contido. Sejnowski (2018) afirma que essa abordagem não obteve sucesso devido ao fato de que os *pixels* de duas imagens que contêm o mesmo objeto serão correspondentes somente se o objeto estiver exatamente na mesma posição nas duas imagens. Portanto, essa abordagem foi abandonada e, segundo Sejnowski (2018), a visão computacional alcançou progresso quando, ao invés de se comparar os *pixels* das imagens, passaram a ser analisadas e comparadas as características do objeto.

Conforme postula Burkov (2019), o *Machine Learning* (ML) requer que as características de um objeto sejam inicialmente identificadas e apontadas por um ser humano. De acordo com LeCun, Bengio e Hinton (2015), desenvolver algoritmos extratores de características exige muita habilidade e conhecimento, já que é um trabalho que pode ser evitado utilizando *Deep Learning* (DL). Logo, ao usar DL, o computador, após ser alimentado com um numeroso conjunto de dados – nesse caso, imagens –, é capaz de identificar as características de um objeto de forma autônoma (LECUN; BENGIO; HINTON, 2015).

2.2 DEEP LEARNING

Para Goodfellow, Bengio e Courville (2016), o DL é uma subárea do ML que permite que os computadores aprendam por experiência e entendam o mundo por meio de uma hierarquia de conceitos na qual os conceitos complexos são compreendidos por intermédio da interpretação de conceitos mais simples.

Conforme salienta Sejnowski (2018), o DL foi criado por meio de uma engenharia reversa do cérebro. Os algoritmos de aprendizado utilizados no DL – conhecidos como redes neurais artificiais – são inspirados na forma que os neurônios se comunicam. Essas redes se apresentam em um modelo distribuído por camadas e, atualmente, as redes neurais artificiais têm dezenas de camadas de profundidade (SEJNOWSKI, 2018).

Segundo Goodfellow, Bengio e Courville (2016), existem dois métodos para medir a profundidade de um modelo de aprendizagem. O primeiro considera a profundidade do grafo computacional que, por sua vez, baseia-se na quantidade de instruções sequenciais que o

modelo é capaz de executar – esse método é comparável a avaliar qual o caminho mais longo de um fluxograma. Já o segundo considera o grafo do modelo probabilístico que, por seu turno, pauta-se na representação da relação entre os conceitos.

Nesse ínterim, Goodfellow, Bengio e Courville (2016) enfatizam que não existe um consenso acerca de qual método de medição de profundidade é o mais adequado; assim como não existe consenso sobre qual é a quantidade mínima de camadas que um modelo deve ter para ser considerado um modelo de aprendizagem profunda.

Serão coletados, nesta pesquisa, mais estudos a respeito do DL com foco na maneira em que essa tecnologia é utilizada para o reconhecimento de imagem.

3 METODOLOGIA

Esta pesquisa é de natureza aplicada com abordagem qualitativa, por meio de pesquisa exploratória. Para contemplar os objetivos, será realizada uma pesquisa bibliográfica.

Por natureza aplicada, conforme destaca Gil (2002), entende-se uma pesquisa capaz de prover conhecimentos passíveis de aplicação prática. Concernente à abordagem qualitativa, Gil (2002) elucida que se refere a uma abordagem na qual o processo de análise dos dados envolve: redução, categorização e interpretação dos dados, assim como a redação do relatório. No que tange à pesquisa exploratória, ainda em consonância com os postulados de Gil (2002), trata-se de uma pesquisa cuja finalidade é promover maior familiaridade com o tema e a compreensão do problema.

Referente à pesquisa bibliográfica, Gil (2002) pontua se tratar de uma pesquisa desenvolvida com base em material já elaborado – no caso, artigos científicos e livros. A pesquisa bibliográfica será realizada seguindo o protocolo de Revisão Sistemática de Literatura (RSL), com base em Biolchini *et al.* (2005) e Kitchenham e Charters (2007), com as etapas evidenciadas a seguir.

1. Definição do problema de pesquisa: “Como o reconhecimento de imagens por meio de *Deep Learning* é utilizado para identificar, catalogar e contabilizar animais selvagens na natureza?”.
2. Definição das palavras-chave: partindo-se da problemática, foram escolhidas as seguintes palavras chaves na língua portuguesa: “reconhecimento de imagem”, “aprendizado profundo” ou “aprendizagem profunda”, “animais” e “preservação animal”; para a língua inglesa, as palavras: “*image recognition*”, “*deep learning*”, “*animals*” e “*animal conservation*”.

3. Seleção das fontes de dados: *The Proceedings of the National Academy of Sciences* (PNAS), IEEE, Scopus, *Science Direct*, *Nature* e Portal da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

4. Relevância, com critérios de inclusão e de exclusão: nessa etapa, as fontes serão verificadas e classificadas de acordo com a sua relevância para o tema da pesquisa, levando em consideração os seguintes critérios de:

a. inclusão: trabalhos publicados nos últimos 10 anos, ou seja, de 2012 até 2022; escritos em português ou inglês; devem estar relacionados ao tema proposto e conter, pelo menos, uma das palavras-chave definidas; estar disponíveis para leitura completa (consulta ou *download*); e

b. exclusão: não atender aos critérios de inclusão; trabalhos duplicados; cursos, tutoriais, *workshops*, palestras e afins.

5. Procedimentos de seleção: os artigos serão selecionados por meio da análise dos títulos e resumos, de modo que:

a. seja comprovada a relação com o tema da pesquisa; e

b. satisfaça os critérios de inclusão e exclusão.

6. Análise: nessa fase, serão analisadas as possíveis fontes de erro (*bias*) que podem comprometer a relevância deste estudo. O Google Acadêmico será utilizado para determinar a relevância dos resultados dos artigos separados para leitura completa, por meio da quantidade de citações.

7. Extração e apresentação dos resultados: será realizada uma análise qualitativa com enfoque descritivo sobre todos os estudos primários, visando a satisfazer o problema de pesquisa proposto.

O resultado esperado é um levantamento dos recursos necessários para efetuar o reconhecimento de imagens, por meio de *Deep Learning*, para identificar, catalogar e contabilizar a fauna silvestre.

4 DESENVOLVIMENTO

O Quadro 1 apresenta o resultado da aplicação das etapas (1)-(5) do Protocolo RSL, realizado no período de 06/2022 a 11/2022. Efetivaram-se buscas utilizando o Google Acadêmico e, ao considerar os critérios de inclusão e exclusão, foram selecionados trabalhos de pesquisa que atendiam aos critérios previamente estabelecidos. A maioria dos trabalhos

excluídos em primeira análise não tinha um conteúdo diretamente relacionado ao tema desta pesquisa; outros foram excluídos por aparecerem duplicados em diferentes bases. Em uma segunda análise, outros trabalhos – que, de início, tinham sido selecionados – foram descartados por já estarem publicados: em periódicos nacionais, com Qualis inferior a B3 ou sem Qualis; em periódicos internacionais sem fator de impacto.

Além disso, foram utilizadas palavras-chave previamente definidas e associadas entre si a fim de filtrar, mais eficazmente, as pesquisas realizadas, de forma que estas apresentem conteúdo similar ao tema desta pesquisa: “reconhecimento de imagem” e “aprendizagem profunda”; “reconhecimento de imagem” e “aprendizado profundo”; “reconhecimento de imagem” e “*deep learning*”; “reconhecimento de imagem”, “*deep learning*” e “animais”; “*deep learning*” e “preservação animal”; “*image recognition*” e “*deep learning*”; “*image recognition*”, “*deep learning*” e “*animals*”; “*deep learning*” e “*animal conservation*”. Importante ressaltar que o termo “*deep learning*” ocorre nas pesquisas em português, visto que esse termo é utilizado, também, em artigos da língua portuguesa.

Quadro 1 – Resultados obtidos até o passo 5, no Google Acadêmico

Palavras-chave	Resultados
“reconhecimento de imagem”	0 Artigos científicos 0 Teses 3 Dissertações 6 TCCs 0 Eventos científicos
“aprendizado profundo” ou “aprendizagem profunda”	0 Artigos científicos 0 Teses 3 Dissertações 6 TCCs 0 Eventos científicos
“preservação animal”	0 Artigos científicos 0 Teses 2 Dissertações 0 TCCs 0 Eventos científicos
“ <i>image recognition</i> ”	12 Artigos científicos 0 Teses 0 Dissertações 12 Eventos científicos

“ <i>deep learning</i> ”	12 Artigos científicos 0 Teses 0 Dissertações 12 Eventos científicos
“ <i>animal conservation</i> ”	7 Artigos científicos 0 Teses 0 Dissertações 9 Eventos científicos

Fonte: a autora (2022).

Ressalta-se que alguns dos artigos no Quadro 1 foram contabilizados em mais de uma busca, pois, por exemplo, o termo “*Deep Learning*” ocorre, também, em artigos sobre “Reconhecimento de imagem” (ou “*Image recognition*”). Além disso, uma das dificuldades encontradas na etapa 4 foi a escassez de material escrito em português e relevante para o tema desta pesquisa. Contudo, houve uma vasta quantidade de material pertinente escrito em inglês.

Na sequência, o Quadro 2 apresenta quais são os tipos de pesquisa encontrados nos idiomas Português/BR, Português/PT e Inglês, como também o fator de impacto das revistas internacionais. É importante ressaltar que o fator do impacto de cada revista foi consultado com o auxílio da ferramenta *on-line Scimago Journal & Country Rank*¹.

Quadro 2 – Resultados divididos por idioma e com fator de impacto indicado

Língua	Artigo científico – Qualis CAPES ou fator de impacto	Evento científico – Artigo completo	Tese	Dissertação	TCC
Português/BR	0	0	0	2	6
Português/PT	0	0	0	1	0
Inglês	12 - Q1	12 - Q1	0	0	0

Fonte: a autora (2022).

A seguir, o Quadro 3 apresenta os trabalhos de pesquisa que foram selecionados para leitura completa. Esses trabalhos foram organizados de acordo com o ano de publicação e o número de citações.

¹ Disponível em: <https://www.scimagojr.com/>. Acesso em: set. 2022.

Quadro 3 – Trabalhos selecionados para leitura completa

TÍTULO/AUTOR(ES)	ANO	CITAÇÕES
Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning (NOROUZZADEH, M. S. <i>et al.</i>).	2018	625
Detecting mammals in UAV images: Best practices to address a substantially imbalanced dataset with deep learning (KELLENBERGER, B.; MARCOS, D.; TUIA, D.).	2018	186
Deep learning object detection methods for ecological camera trap data (SCHNEIDER, S.; TAYLOR, G. W.; KREMER, S.).	2018	127
Recognition of endangered pantanal animal species using deep learning methods (ARRUDA, M. S. de <i>et al.</i>).	2018	10
Wildlife surveillance using deep learning methods (CHEN, R. <i>et al.</i>).	2019	27
Deep learning for inexpensive image classification of wildlife on the Raspberry Pi (CURTIN, B. H.; MATTHEWS, S. J.).	2019	14
Identification of animal individuals using deep learning: A case study of giant panda (HOU, J. <i>et al.</i>).	2020	33
Tracking hammerhead sharks with deep learning (PEÑA, A. <i>et al.</i>).	2020	3
Deep learning methods for multi-species animal re-identification and tracking—a survey (RAVOOR, P. C.; SUDARSHAN, T. S. B.).	2020	14
Counting sea lions and elephants from aerial photography using deep learning with density maps (PADUBIDRI, C. <i>et al.</i>).	2021	3

Fonte: a autora (2022).

Foram selecionados dez trabalhos que descrevem estudos práticos realizados utilizando o reconhecimento de imagem por meio de *Deep Learning* para executar, ao menos, uma das seguintes tarefas: detectar a presença de animais; identificar a espécie a qual os animais pertencem; contabilizar os animais presentes em uma determinada região; descrever o comportamento dos animais; e identificar animais individualmente. Na sequência, expõe-se o conteúdo desses estudos.

4.1 PRESERVAÇÃO ANIMAL APOIADA PELO *DEEP LEARNING*

O conteúdo desta seção foi construído a partir da leitura dos artigos selecionados, levando em consideração os questionamentos evidenciados no Quadro 4.

Quadro 4 – Perguntas elaboradas para conduzir a leitura dos artigos

PERGUNTAS
Onde o estudo foi realizado?
Quais são os animais silvestres focados?
Qual é o objetivo geral da pesquisa?
Quais são os recursos tecnológicos de <i>software</i> utilizados?
Quais são os recursos tecnológicos de <i>hardware</i> utilizados?
Quais foram as dificuldades e facilidades encontradas para a execução da pesquisa?
Qual é a acurácia/precisão na identificação dos animais utilizando a tecnologia empregada?
Quais sugestões de trabalho futuro o autor do artigo apresenta?

Fonte: a autora (2022).

Os artigos selecionados apresentam os seguintes locais onde cada estudo foi realizado: Norouzzadeh *et al.* (2018) indicam o Parque Nacional do Serengeti na Tanzânia, na África. Hou *et al.* (2020) apresentam a base “Pesquisa e Criação de Pandas Gigantes de Chengdu” e as bases de Dujiangyan e Wolong de “*China Conservation and Research Center for the Giant Panda*”, na China, na Ásia; esses autores também trazem a informação de que foram utilizadas fotos de 6 pandas de Chengdu, 9 de Dujiangyan e 10 de Wolong.

Arruda *et al.* (2018) realizaram o estudo no Pantanal, no Brasil. Kellenberger, Marcos e Tuia (2018) efetuaram suas pesquisas em *Kuzikus Wildlife Reserve*, na Namíbia, na África. Chen *et al.* (2019) indicam o Reino Unido como o local das imagens de um dos dois *datasets* utilizados na pesquisa. Padubidri *et al.* (2021) realizaram o estudo nas Ilhas Aleutas, nos EUA, bem como na África; para este último, não se especificou um país ou região do continente.

Schneider, Taylor e Kremer (2018) indicam que as imagens de um dos *datasets* utilizados provêm da Holanda e do Panamá, enquanto o outro contém imagens do Parque Nacional do Serengeti, na Tanzânia, na África. Ravoore e Sudarshan (2020) informam que os

estudos levantados foram realizados na China; Peña *et al.* (2020) apontam que o estudo ocorreu na Ilha de Galápagos, no Equador, e na Ilha do Cobo, na Costa Rica. O estudo de Curtin e Matthews (2019) não apresentou essa informação.

Também são apresentados quais animais silvestres tiveram foco em cada um dos estudos: Norouzzadeh *et al.* (2018) indicam que seu estudo trata de 48 espécies de animais distintas e ressaltam, entretanto, que a pesquisa trabalhou com a identificação de apenas uma espécie por imagem, descartando cenários em que animais de espécies diferentes apareciam em uma mesma imagem. Hou *et al.* (2020) focam exclusivamente em pandas-gigantes. Arruda *et al.* (2018) focam em: antas, araras-canindé, capivaras, caititus, onças pardas, papagaios-verdadeiros, quatis-de-cauda-anelada e tamanduás-bandeiras. Kellenberger, Marcos e Tuia (2018) focam em mamíferos de grande porte, no entanto não se explicitam quais características visuais são consideradas para distinguir um mamífero das demais categorias de animais.

Chen *et al.* (2019) tratam de texugos. O estudo de Padubidri *et al.* (2021) foca em leões-marinhos-de-steller e em elefantes-africanos. Ravor e Sudarshan (2020) analisaram três estudos sobre tigres-siberianos. Peña *et al.* (2020) focam em tubarões-martelo. Curtin e Matthews (2019) realizaram o estudo com leopardos-das-neves. A pesquisa de Schneider, Taylor e Kremer (2018) não especifica em quais animais o estudo foi focado. Entretanto, sabe-se que um dos *datasets* utilizados contém imagens da savana africana.

Naturalmente, é necessário descrever onde se pretende chegar com o estudo. O objetivo geral da pesquisa de Norouzzadeh *et al.* (2018) é identificar, contar e descrever o comportamento dos animais de espécies distintas. Nota-se que, por “descrever o comportamento”, entende-se identificar se o animal presente na foto está sentado, deitado, em pé, movimentando-se, alimentando-se etc. A pesquisa de Hou *et al.* (2020) tem como objetivo identificar pandas gigantes individualmente, isto é, saber diferenciar um indivíduo do outro. Arruda *et al.* (2018) desejam apresentar um método para detectar e reconhecer animais selvagens na natureza e mostrar que esse método garante maior precisão na identificação de animais em comparação direta ao método Fast R-CNN. Kellenberger, Marcos e Tuia (2018) pretendem detectar a presença de animais em imagens capturadas por meio de veículo aéreo não tripulado (VANT ou, mais popularmente, *drones*) utilizando Redes Neurais Convolucionais (em inglês, *Convolutional Neural Networks* – CNNs) e apresentar técnicas de como melhorar a precisão na detecção de animais em *datasets* desbalanceados – especialmente naqueles em que a maioria das imagens não contém animal algum.

Chen *et al.* (2019) se propõem a desenvolver um *framework* robusto capaz de identificar e classificar espécies de animais selvagens e seus comportamentos em imagens estáticas – e,

então, aplicar o mesmo algoritmo a vídeos. Padubidri *et al.* (2021) têm por objetivo contabilizar, com alta precisão, a população de leões-marinhos-stellar e de elefantes africanos utilizando visão computacional por meio de *Deep Learning*, em conjunto com aerofotografia e mapas de densidade demográfica. Schneider, Taylor e Kremer (2018) pretendem identificar, quantificar e localizar animais de diferentes espécies em imagens de armadilhas fotográficas; nessa pesquisa, a identificação de diferentes espécies busca ser feita, inclusive, quando se trata de espécies distintas em uma mesma imagem. Ravoor e Sudarshan (2020) desejam explorar a aplicabilidade de métodos existentes de reidentificação de animais para o monitoramento automatizado da vida selvagem; por reidentificação, entende-se saber se um animal presente na imagem B é o mesmo indivíduo que aparece na imagem A. Peña *et al.* (2020) propõem um método automatizado de detectar e monitorar tubarões-martelo, baseando-se em imagens de vídeo. Por fim, Curtin e Matthews (2019) têm como objetivo detectar leopardos-das-neves de maneira automatizada utilizando equipamentos de baixo custo.

Os artigos apresentam, também, quais são os recursos tecnológicos de *software* utilizados no desenvolvimento e/ou treinamento dos modelos de aprendizado: Norouzzadeh *et al.* (2018) indicam que os principais recursos de *software* utilizados são CNNs com arquitetura VGG. Hou *et al.* (2020) utilizam CNNs com arquitetura VGGNet e a ferramenta *Tensorflow* para treinamento de modelo. Arruda *et al.* (2018) usam algoritmo SLIC para identificar as áreas da imagem em que existem animais presentes e CNNs com arquitetura VGGNet para identificar as espécies dos animais; foi utilizada a ferramenta MATLAB com MatConvNet para a implementação dos CNNs. Kellenberger, Marcos e Tuia (2018) utilizam CNN implementada recorrendo ao *framework* PyTorch; para treinamento de modelo, foi usado o algoritmo *Adam optimizer*; o sistema operacional da máquina utilizada era Linux.

Chen *et al.* (2019) relatam ter treinado dois modelos CNN utilizando arquitetura AlexNet para um deles; a ferramenta MATLAB foi utilizada para avaliar as *performances* dos modelos. Padubidri. *et al.* (2021) utilizam CNNs com modelo U-Net básico para o primeiro dos dois modelos treinados (chamado de Model-K); para o segundo modelo treinado (chamado de Model-2): CNNs com modelo U-Net, com EfficientNet-B5 para extrair as características das imagens. Schneider, Taylor e Kremer (2018) treinaram dois modelos com abordagens diferentes: *Faster Region-Convolutional Neural Network* (Faster R-CNN) e *You-Only-Look-Once v2.0* (YOLO v2.0), ambos utilizando arquitetura ResNet-101.

A pesquisa exploratória de Ravoor e Sudarshan (2020) levantou três estudos, e dois deles utilizaram variações de ResNet para a extração de características e Triplet Loss para treinamento; ademais, Ravoor e Sudarshan (2020) informam que um dos estudos por eles

analisados (a título de exemplo: Liu, Zhang e Guo, datado de 2019) utilizou uma arquitetura chamada de *Part-Pose Guided Network (PPGNet)*; outro estudo (Shukla *et al.*, proveniente de 2019) usa o modelo DenseNet121 treinado por uma combinação de técnicas de entropia cruzada e divergência de Kullback-leibler. Peña *et al.* (2020) usam CNNs com YOLOv3 aprimorado (foram adicionadas 18 camadas à arquitetura convencional); toda a implementação foi feita em Python, com CNN Darknet, *framework* Pytorch e bibliotecas Numpy, OpenCV, scikit-learn e scikit-image. Curtin e Matthews (2019) utilizam CNNs com as bibliotecas TensorFlow e Keras; o sistema operacional da máquina usada para construir o modelo era Windows 10.

Além dos recursos tecnológicos de *software*, são apresentados os recursos tecnológicos de *hardware* utilizados na captura das imagens e/ou desenvolvimento e/ou treinamento dos modelos: Norouzzadeh *et al.* (2018) indicam que, para a captura de imagens, foram utilizadas armadilhas fotográficas. No entanto, os autores não trazem informações sobre os recursos de *hardware* usados para o desenvolvimento ou treinamento dos algoritmos. Hou *et al.* (2020) informam que foi utilizada uma GPU Nvidia Quadro P5000 (16GB) para o treinamento de modelo e, para a captura de imagens, foram usadas câmeras (inclusive, aquelas de aparelhos celulares). Arruda *et al.* (2018) indicam que foi usada uma GPU Nvidia Titan X para o experimento e que as imagens térmicas foram capturadas com uma câmera térmica FLIR SC640.

Kellenberger, Marcos e Tuia (2018) afirmam que, para o treinamento, usaram-se CPU Intel Xeon e uma placa gráfica Nvidia GeForce GTX 1080Ti; enquanto, para capturar as imagens, foi usado um *drone* SenseFly3 eBee, portando uma câmera Canon PowerShot S110 RGB; também foi usado um sensor térmico multiespectral. Chen *et al.* (2019) relatam o uso do processador Intel i7-7700K (4.2 GHz \times 4), 16 GB de RAM e um Nvidia GeForce RTX 2080. Padubidri *et al.* (2021) afirmam que, para o treinamento, foi usada uma GPU Nvidia GeForce RTX 2060; as imagens aéreas dos elefantes foram capturadas por uma câmera Canon 6D reflex monobjetiva acoplada a uma aeronave *SkyReach BushCat*. Peña *et al.* (2020) indicam o uso de um Nvidia DGX com 4 GPUs Tesla; os autores ressaltam, entretanto, que o modo multi-GPU não foi utilizado nos treinamentos.

Curtin e Matthews (2019) utilizaram um Raspberry-Pi 3B+ que, por sua vez, tem um processador 1.4GHz ARM A53, 1GB de RAM e *bluetooth* integrado; para construir o modelo, foi utilizado um *notebook* Dell Latitude 7350 com CPU 64-bit Intel Core M-5Y71 1.2 GHz, 8GB de RAM. As pesquisas de Schneider, Taylor e Kremer (2018) e Ravoor e Sudarshan (2020) não apresentam qualquer informação a respeito das tecnologias de *hardware* utilizadas.

O Quadro 5 expressa as dificuldades e facilidades encontradas pelos autores no que diz respeito às suas pesquisas. Salienta-se que o estudo de Ravor e Sudarshan (2020) realizou uma pesquisa bibliográfica exploratória, em que são evidenciadas as facilidades/dificuldades dos estudos por eles identificadas.

Quadro 5 – Facilidades e dificuldades encontradas

Norouzzadeh et al. (2018)
Facilidades: não constam.
Dificuldades: segundo os autores, as imagens obtidas por meio das chamadas armadilhas fotográficas (em inglês, <i>camera traps</i>) não costumam ser perfeitas. Muitas vezes, o(s) animal(is) presente(s) na imagem está/estão muito distante(s), perto demais ou é/são apenas parcialmente visível/visíveis. Além disso, as condições do tempo, sombra e iluminação influenciam no nível de dificuldade de se identificar o conteúdo de uma imagem. Os autores ainda apresentam mais um problema: algumas espécies de animais estão muito presentes no <i>dataset</i> utilizado, enquanto outras, raramente, são encontradas. Isso pode fazer com que o algoritmo se torne enviesado, ou seja, ao classificar um animal, o sistema tende a classificá-lo como um dos animais que ele viu mais vezes.
Hou et al. (2020)
Facilidades: o estudo foi realizado em animais em cativeiro; portanto, obter boas imagens não foi um problema.
Dificuldade: pandas-gigantes são muito semelhantes entre si; isso aumenta o nível de dificuldade em identificar indivíduos. Em outras espécies de animais, é possível distinguir indivíduos por diferenças no seu porte físico e até fisionomia com maior facilidade.
Arruda et al. (2018)
Facilidades: não constam.
Dificuldades: não constam.
Kellenberger, Marcos e Tuia (2018)
Facilidades: não constam.
Dificuldades: os autores apontam que muitos dos trabalhos relacionados à identificação de animais por reconhecimento de imagem são realizados em parques nacionais ou reservas bastante pequenas. No entanto, a reserva de Kuzikus é moderadamente grande, o que faz com que o avistamento de animais seja raro e que a tarefa de contabilizá-los seja árdua. Outros pontos de dificuldade são: (i) a savana é um bioma heterogêneo e isso significa que, ao percorrer uma ampla extensão, é possível notar que existe variação na aparência da paisagem. A propósito, isso não seria problema para um humano, mas uma máquina treinada, tendo apenas uma parte de determinado ambiente como “plano de fundo”, pode cometer erros em um ambiente de aparência diferente; (ii) cobrir uma área extensa também implica lidar com

um maior número de objetos que podem ser confundidos com animais em uma visão aérea (truncos de árvores, arbustos etc.).
Chen et al. (2019)
Facilidades: não constam.
Dificuldades: segundo os autores, quando se trata de uma gravação de vídeo, o movimento de um animal causa variações nos valores dos <i>pixels</i> e esse seria o primeiro indicativo da presença de um chamado “objeto de interesse” (no caso, qualquer animal). Porém, uma movimentação identificada em vídeo também pode ser causada pelo vento movendo a vegetação da região ou, então, a própria câmera. Solução encontrada: uma das medidas tomadas para evitar que qualquer movimento gerasse um falso positivo foi a de descartar movimentos que abrangessem áreas muito grandes, já que, de acordo com os autores, o movimento gerado por animais, comumente, abarca áreas pequenas.
Padubidri et al. (2021)
Facilidades: segundo os autores, o <i>dataset</i> referente aos leões-marinhos-de-steller apresentava melhor qualidade de imagem e isso fez com que fosse possível observar mais detalhes do que em relação aos elefantes africanos.
Dificuldades: uma dificuldade mencionada foi a de se identificar filhotes. Em ambas as espécies estudadas, a dificuldade se deve, principalmente, ao fato de que os filhotes estão sempre muito próximos da mãe e, por isso, pode ser difícil identificar a presença de mais um indivíduo nas imagens. No caso dos elefantes, esse fato, aliado à diferença de tamanho em relação aos elefantes adultos, faz com que os filhotes, frequentemente, apareçam parcial ou quase totalmente ocultos nas imagens. Além de serem ocultados pela mãe ou outros elefantes maiores, os filhotes são facilmente encobertos por árvores, por exemplo. Já no caso dos leões-marinhos-de-steller, os filhotes são muito parecidos com as pedras presentes em seu ambiente, criando um efeito de camuflagem.
Schneider, Taylor e Kremer (2018)
Facilidades: não constam.
Dificuldades: os autores apontam que o <i>dataset Snapshot Serengeti</i> – que apresenta a maior coleção pública de imagens de armadilhas fotográficas no mundo – é bastante desbalanceado no conteúdo de suas imagens. <i>Datasets</i> desbalanceados podem deixar o algoritmo de reconhecimento enviesado – conceito já explicado por Norouzzadeh et al. (2018). Os autores também destacam como dificuldade o fato de que, em muitas fotos, os animais aparecem bem distantes, parcialmente ocultos ou perto demais da câmera.
Ravor e Sudarshan (2020)
Facilidades: não constam.
Dificuldades: assim como para a identificação, os desafios da reidentificação incluem: variação de iluminação, condições climáticas, diferentes ângulos de captura de imagem, oclusão parcial ou completa do animal presente na imagem etc. Diante do exposto, destaca-se a dificuldade causada por diferentes ângulos de captura, pois alguns ângulos podem ocultar as principais características de identificação de um animal. Por exemplo, este artigo ressalta

que estudos indicaram que é mais fácil reidentificar um mesmo tigre, baseando-se em duas imagens capturadas do ângulo lateral, já que, assim, é possível fazer uma melhor comparação das listras; com uma imagem lateral e outra frontal, torna-se muito mais difícil saber se trata do mesmo indivíduo.

Peña *et al.* (2020)

Facilidades: o corpo do tubarão-martelo tem um formato único e isso facilita sua identificação.

Dificuldades: apesar da facilidade citada, a identificação ainda é uma tarefa difícil devido a condições do ambiente que não são possíveis de serem controladas, como: baixa iluminação, oclusão causada por outras espécies de animais (exemplo: pode ser difícil identificar um tubarão em meio a um cardume de outra espécie de peixe), reflexo da luz do sol na água, o tubarão estar muito longe da câmera etc. Solução encontrada: para tentar contornar e diminuir o impacto dessas dificuldades, foram adicionadas mais 18 camadas de profundidade à arquitetura de *software* utilizada nessa pesquisa, conforme mencionado anteriormente. O maior número de camadas aumenta a precisão na identificação.

Curtin e Matthews (2019)

Facilidades: não constam.

Dificuldades: os autores relatam as mesmas dificuldades já mencionadas por outros autores, como: influência da qualidade de imagem, iluminação, o animal estar muito distante da câmera ou parcialmente oculto. Os autores apontam que seu modelo tem dificuldade em realizar a detecção quando, especialmente, o animal não toma a maior parte da imagem ou quando a imagem contém muita folhagem verde (nesse caso, em uma classificação que foi dividida entre “leopardo-das-neves”, “humano” ou “outro”, a imagem tende a ser classificada como “outro”, ou seja, uma imagem que não contém nenhum objeto de interesse). Os autores também mencionam não exatamente uma dificuldade, mas uma limitação do estudo: a maioria das imagens dos leopardos-das-neves provém de zoológicos que os mantêm em ambientes com folhagens verdes. Entretanto, o habitat natural dos leopardos-das-neves é rochoso e, durante parte do ano, coberto de neve. É preciso reconhecer que isso dificultaria a reprodução dos resultados obtidos nesse estudo, caso ele viesse a ser aplicado em habitat natural.

Fonte: a autora (2022).

Uma das informações que se espera encontrar sobre os estudos citados é, certamente, se eles obtiveram sucesso em alcançar o objetivo estabelecido. A seguir, apresenta-se a acurácia/precisão registrada por cada um dos modelos de aprendizado em detectar/identificar/contabilizar animais selvagens: Norouzzadeh *et al.* (2018) afirmam que o modelo por eles desenvolvido apresenta 93.8% de acurácia geral e 99.3% se o sistema classificar uma imagem apenas quando tem alto nível de confiança na identificação. Hou *et al.* (2020) indicam que seu modelo apresenta 95% de acurácia geral; >90% em imagens de baixa qualidade; 78,7%, tratando-se de animais que não estavam presentes no *dataset* usado no

treinamento; 95% para imagens rotacionadas em até 20°; 91.1% em fotos rotacionadas de 20-45°. Arruda *et al.* (2018) indicam precisão média de 83.89%, o que faz do método apresentado ao menos 6% mais preciso que o método Fast R-CNN com o qual ele foi diretamente comparado. Chen *et al.* (2019) informam que o modelo CNN-1 apresentou 95.86% de acurácia na classificação binária e 83.07% na multiclassificação; já o modelo CNN-2 apresentou 98.05% de acurácia na classificação binária e 90.32% na multiclassificação.

Padubidri *et al.* (2021) indicam que foram treinados dois modelos e salientam que a acurácia foi medida por raiz do erro quadrático médio. Assim, o Model-K registrou acurácia de 2.17 para leões-marinhos e 0.81 para os elefantes; já o Model-2, por sua vez, registrou acurácia de 1.88 para os leões-marinhos-de-steller e 0.60 para os elefantes, apresentando a melhor *performance* entre os dois. Schneider, Taylor e Kremer (2018) apontam dois modelos treinados: Faster R-CNN e YOLO v2.0; o modelo Faster R-CNN apresentou acurácia de 93.0% no *dataset Reconyx Camera Trap* e 76,7% no *dataset Snapshot Serengeti*. O modelo YOLO v2.0 apresentou acurácia de, respectivamente, 73.0% e 40.3% nos mesmos *datasets* – *performance* considerada ruim pelos pesquisadores.

Para os três estudos analisados por Ravoor e Sudarshan (2020), foi avaliado o nível de acurácia na reidentificação em dois cenários, em que são comparadas imagens: (i) de uma única câmera; e (ii) de “câmeras cruzadas”, isto é, imagens de um mesmo indivíduo capturadas por câmeras diferentes. Ravoor e Sudarshan (2020) informam, então, que o estudo de Li *et al.*, de 2019, registrou acurácia de 89.4% e 77.1%, respectivamente, para os cenários descritos; o estudo de Liu, Zhang e Guo, também de 2019, apresentou acurácia de 97.7% e 93.6%; e o de Shukla *et al.*, assim como os demais, proveniente de 2019, registrou 92.7% e 84.5%. Peña *et al.* (2020) utilizaram a raiz do erro quadrático médio para medir a precisão e afirmam que o método proposto apresentou 0.99 de precisão em determinar a caixa delimitadora em volta dos tubarões e 0.93 de *recall* em detectá-los nas imagens, de fato; esses resultados foram comparados aos resultados obtidos com o YOLOv3 tradicional, o qual, por sua vez, apresentou, respectivamente, 0.95 e 0.83 de precisão e *recall* para as mesmas tarefas. Peña *et al.* (2020) concluem que aumentar o número de camadas da arquitetura melhorou, por certo, sua *performance*. Curtin e Matthews (2019) informam que o método proposto apresentou 97% de acurácia em imagens estáticas e 74% de acurácia em uma simulação de realidade na qual as imagens foram distanciadas da câmera. Kellenberger, Marcos e Tuia (2018) não apresentam a

acurácia, mas enfatizam >90% de *recall*² alcançados pelo modelo treinado utilizando todas as técnicas propostas para contornar os problemas causados por *datasets* desbalanceados.

Por fim, os autores apresentam quais trabalhos futuros eles pretendem desenvolver ou quais pontos eles acreditam que podem ser explorados por outros pesquisadores.

- Norouzzadeh *et al.* (2018) pretendem abordar uma limitação de seu modelo: identificar apenas uma espécie por imagem. Serão feitos estudos a fim de superar essa limitação. Além disso, os autores pretendem estudar as possíveis vantagens de sistemas híbridos, utilizando times de humanos voluntários para auxiliar na classificação das imagens. Diferentes sistemas serão estudados. Um exemplo citado envolve ter um time de humanos voluntários para classificar somente as imagens que o modelo desenvolvido evidencia como de baixo nível de confiança. Uma vez classificadas pelos humanos, essas imagens seriam usadas em um novo treinamento de modelo, melhorando o nível de precisão na classificação.

- Hou *et al.* (2020) indicam que seu modelo pode ser aplicado em estudos futuros sobre pandas e outras espécies, pois o modelo é adaptável. O modelo desenvolvido já está presente em um pacote do Matlab para que outros pesquisadores e ecologistas possam utilizá-lo.

- Arruda *et al.* (2018) desejam incluir um maior número de espécies nos treinamentos e pretendem comparar seu modelo a outros, como o Faster-RCNN e Yolo2.

- Kellenberger, Marcos e Tuia (2018) apenas declaram que seus métodos de treinamento podem ser aplicados em qualquer modelo de detecção de objeto baseado em *Deep Learning*, sugerindo, então, que outros pesquisadores o empreguem para os fins que desejarem.

- Chen *et al.* (2019) pontuam que o algoritmo desenvolvido não é capaz de identificar mais de uma espécie de animal diferente em uma mesma foto e, para isso ser possível, seriam necessárias adaptações. Outra melhoria pertinente seria deixar o *software* mais “*user friendly*” para que ele seja lançado como programa ou *Web Service* e disponibilizado para pesquisadores de forma acessível.

- Padubidri. *et al.* (2021) indicam que, em trabalhos futuros, o foco será direcionado a melhorar a precisão da detecção de animais. Será investigada a possibilidade de gerar dados sintéticos para os treinamentos, principalmente para gerar imagens de animais que estejam ocultos na imagem em algum nível e de animais de que são quase indistinguíveis do

² Por *recall*, compreende-se a capacidade de o modelo detectar resultados classificados como positivos (BURKOV, 2019). Por **precisão**, compreende-se a razão entre quantidade de positivos verdadeiros e a quantidade de resultados previstos como positivos (BURKOV, 2019). Por **acurácia**, compreende-se a razão entre a quantidade de acertos sobre o total de entradas (BURKOV, 2019).

cenário de fundo. Funções de Gauss serão utilizadas para aprimorar o alinhamento entre os mapas de densidade demográfica e a real posição dos animais.

- Schneider, Taylor e Kremer (2018) destacam melhorias e evoluções possíveis a partir de seu trabalho: o modelo Faster R-CNN poderia ser utilizado para obter informações mais detalhadas sobre o comportamento dos animais. Isso incluiria identificar em que época(s) do ano determinadas espécies se reproduzem, quantos filhotes costumam nascer etc.

- Schneider, Taylor e Kremer (2018) também apontam que a reidentificação de indivíduos, isto é, identificar um animal e, depois, conseguir reconhecê-lo como indivíduo já identificado anteriormente, é uma área de pesquisa ainda pouco avançada. Os autores destacam o quão importante seria avançar nessa área, pois isso permitiria uma maior automatização da estimativa populacional.

- Ravor e Sudarshan (2020) apenas dizem esperar que seu trabalho sirva de base para mais pesquisas na área de monitoramento de animais silvestres.

- Peña *et al.* (2020) declaram que, em trabalhos futuros, serão feitas comparações entre o método desenvolvido na presente pesquisa e outros métodos existentes; os autores também destacam que o método proposto não ofereceu uma solução satisfatória para o problema de oclusão, especificamente; dão a entender que é um ponto a ser explorado.

- Curtin e Matthews (2019) manifestam a intenção de aprimorar o modelo desenvolvido para que ele passe a detectar, com maior precisão, os animais que estão distantes ou parcialmente ocultos. Os autores também pretendem aplicar seu método no Raspberry Pi Zero, uma versão mais moderna e ainda mais barata do Raspberry Pi (de acordo com os autores, o Raspberry Pi Zero custa \$10.00; em comparação, o Raspberry Pi 3B+ custa \$35.00).

5. CONCLUSÕES E TRABALHO FUTURO

Este estudo permitiu compreender que o reconhecimento de imagem por meio de *Deep Learning* pode ser um grande aliado no trabalho da preservação da fauna. O número de imagens gerado pelas armadilhas fotográficas e demais métodos de captura de imagens da vida selvagem costuma ser bastante grande, e a automatização da análise dessas imagens agiliza o trabalho de identificação, catalogação e contabilização de espécies de animais buscado pelos biólogos pesquisadores.

Durante a elaboração desta pesquisa, percebeu-se que alguns estudos se complementam: Schneider, Taylor e Kremer (2018) destacam a importância da reidentificação de animais e os

problemas causados por não conseguirem fazê-la. Ravoor e Sudarshan (2020) realizam uma pesquisa exploratória especificamente sobre esse tema.

Alguns estudos apresentam dificuldades/problemas causados por *datasets* desbalanceados, pois estes deixam o algoritmo enviesado, o que diminui sua precisão. Norouzzadeh *et al.* (2018) advertem que o seu modelo de aprendizado é capaz de identificar somente uma espécie de animal por imagem e estabelecem que a capacidade de identificar diferentes espécies em uma mesma imagem deve ser buscada apenas em trabalhos futuros. Um dos temas tratados por Kellenberger, Marcos e Tuia (2018) é justamente apresentar métodos para lidar com *datasets* desbalanceados, minimizando os impactos negativos citados por Norouzzadeh *et al.* (2018). A pesquisa de Schneider, Taylor e Kremer (2018) recorre a essa capacidade e a apresenta como parte do objetivo geral da pesquisa. Isso demonstra o quão atentos os pesquisadores estão às faltas e limitações da tecnologia por eles estudada, quais são as implicações dessas limitações e o que se poderia alcançar caso essas limitações fossem superadas.

Como trabalho futuro, sugere-se aprofundar o conteúdo referente às Redes Neurais Convolucionais, uma vez que elas são utilizadas em quase todos os estudos apresentados nesta pesquisa. Além disso, salienta-se a ausência de estudos realizados na Amazônia e, da mesma forma, não existem estudos realizados no Parque Nacional do Iguaçu em Foz do Iguaçu, no estado do Paraná, no qual seria interessante explorar a possibilidade de reproduzir um estudo semelhante aos que aqui foram apresentados.

REFERÊNCIAS

AMIRIAN, S. *et al.* Dissection of Deep Learning with Applications in Image Recognition. In: INTERNATIONAL CONFERENCE ON COMPUTATIONAL SCIENCE AND COMPUTATIONAL INTELLIGENCE (CSCI), 2018, [S. I.]. **Anais...** [S. I.]: IEEE, 2018, p. 1142-1148.

ARRUDA, M. S. de *et al.* Recognition of endangered pantanal animal species using deep learning methods. In: INTERNATIONAL JOINT CONFERENCE ON NEURAL NETWORKS (IJCNN), 2018, [S. I.]. **Anais...** [S. I.]: IEEE, 2018, p. 1-8.

BIOLCHINI, J. *et al.* Systematic review in software engineering. **System Engineering and Computer Science Department COPPE/UFRJ**, Rio de Janeiro, v. 679, n. 5, p. 45, 2005.

BURKOV, A. **The Hundred-Page Machine Learning Book**. 2. ed. Quebec, CAN: Andriy Burkov, 2019.

- CHEN, R. *et al.* Wildlife surveillance using deep learning methods. **Ecology and evolution**, [S. I.], v. 9, n. 17, p. 9453-9466, 2019.
- CONDE, A. D. *et al.* Data gaps and opportunities for comparative and conservation biology. **PNAS**, [S. I.], v. 116, n. 19, p. 9658-9664, 2019.
- CURTIN, B. H.; MATTHEWS, S. J. Deep learning for inexpensive image classification of wildlife on the Raspberry Pi. In: ANNUAL UBIQUITOUS COMPUTING, ELECTRONICS & MOBILE COMMUNICATION CONFERENCE (UEMCON), 10., 2019, [S. I.]. **Anais...** [S. I.]: IEEE, 2019, p. 0082-0087.
- GIL, A. C. **Como elaborar projetos de pesquisa**. 4. ed. São Paulo: Atlas, 2002.
- GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep Learning**. 1. ed. Cambridge, USA: MIT Press, 2016.
- HOU, J. *et al.* Identification of animal individuals using deep learning: a case study of giant panda. **Elsevier, Biological Conservation**, [S. I.], v. 242, p. 108414, 2020.
- KELLENBERGER, B.; MARCOS, D.; TUIA, D. Detecting mammals in UAV images: best practices to address a substantially imbalanced dataset with deep learning. **Elsevier, Remote Sensing of Environment**, [S. I.], v. 216, p. 139-153, 2018.
- KITCHENHAM, B.; CHARTERS, S. **Guidelines for performing systematic literature reviews in software engineering**. Durham: Technical Report EBSE 2007-001, Keele University and Durham University Joint Report, 2007.
- LECUN, Y.; BENGIO, Y.; HINTON, G. Deep Learning. **Nature**, [S. I.], v. 521, n. 7553, p. 436-444, 2015.
- NOROUZZADEH, M. S. *et al.* Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning. **PNAS**, [S. I.], v. 115, n. 25, p. E5716-E5725, 2018.
- PADUBIDRI, C. *et al.* Counting sea lions and elephants from aerial photography using deep learning with density maps. **BMC, Animal Biotelemetry**, [S. I.], v. 9, n. 1, p. 1-10, 2021.
- PEÑA, A. *et al.* Tracking hammerhead sharks with deep learning. In: COLOMBIAN CONFERENCE ON APPLICATIONS OF COMPUTATIONAL INTELLIGENCE (IEEE CoIACI), 2020, [S. I.]. **Anais...** [S. I.]: IEEE, 2020, p. 1-6.
- RAVOOR, P. C.; SUDARSHAN, T. S. B. Deep learning methods for multi-species animal re-identification and tracking – a survey. **Elsevier, Computer Science Review**, [S. I.], v. 38, p. 100289, 2020.
- SCHNEIDER, S.; TAYLOR, G. W.; KREMER, S. Deep learning object detection methods for ecological camera trap data. In: CONFERENCE ON COMPUTER AND ROBOT VISION (CRV), 15., 2018, [S. I.]. **Anais...** [S. I.]: IEEE, 2018, p. 321-328.
- SEJNOWSKI, T. **Deep Learning Revolution**. 1. ed. Cambridge, USA: MIT Press, 2018.