

**ANEXO 09 - EDITAL 02/2022**

**FORMULÁRIO DE PROJETO DE PESQUISA**

**1 TÍTULO DO PROJETO**

Modelos de redes neurais para classificação de peixes recifais do Atlântico e Pacífico Oriental

**2 CAMPUS**

*Campus:* Ituiutaba

**3 INTRODUÇÃO** *(explicação de forma clara e objetiva da problemática a ser investigada e das possíveis hipóteses, caso tenha)*

A área de Inteligência Artificial se encontra em pleno desenvolvimento e é o foco de inúmeras pesquisas na área de computação. Sua utilização em diferentes tipos de sistemas é cada vez mais comum. Redes Neurais Artificiais é uma parte da inteligência computacional. Seu objetivo é simular o comportamento de uma rede de neurônios naturais utilizando neurônios artificiais.

Os peixes de recife, ou peixes recifais, representam um dos grupos de vertebrados mais diversos da Terra, com mais de 7000 espécies distribuídas ao redor do globo. Essa riqueza não é distribuída uniformemente geograficamente. O Atlântico (AT) e o Pacífico Oriental (Eastern Pacific, EP) abrangem 30% da fauna global de peixes. Essas áreas têm sido consideradas as mais isoladas do hotspot de biodiversidade marinha no Indo-Pacífico devido a barreiras físicas distintas, como o fechamento de Tethyan e a distância entre o Pacífico Oriental e o Pacífico Ocidental. Apesar de sua riqueza de espécies comparativamente menor, esses reinos hospedam assembleias de peixes únicas caracterizadas por uma proporção notável de endemismos regionais e espécies com grande tamanho corporal.

Neste projeto de pesquisa, utilizaremos o maior banco de dados de características de história de vida, aspectos biogeográficos e de conservação atualmente disponíveis para a fauna de peixes recifais dos reinos AT e EP. O banco de dados, elaborado por Quimbayo et. al. (2021) inclui 21 características distribuídas em comportamentais, morfológicas, e ecológica, bem como biogeográficas e aspectos de conservação.

Este banco de dados de peixes de recife oferece a oportunidade de explorar novas questões ecológicas e evolutivas em diferentes escalas e fornece ferramentas para a conservação de espécies com base nessas características. Além disso, esse banco de dados possui informações que podem ser utilizadas como base de treinamento para modelos de Inteligência Artificial para classificação de peixes recifais.

## 4 OBJETIVOS

### 4.1 Geral *(dimensão mais ampla pretendida com a pesquisa)*

Este projeto tem como objetivo geral o desenvolvimento de modelos de Inteligência Artificial baseado em Redes Neurais Artificiais para classificação de peixes recifais utilizando como base de dados para treinamento, o banco de dados elaborado por Quimbayo et. al. (2021).

### 4.2 Específicos *(dimensão mais detalhada dos procedimentos de busca do referido projeto)*

Este projeto tem como objetivos específicos:

- Coleta e tratamento de dados do banco de dados de peixes recifais elaborado por Quimbayo et. al. 2021;
- Extração das características principais para treinamento dos modelos de classificação;
- Concepção, design e desenvolvimento dos modelos de classificação baseado em redes neurais artificiais;
- Testes e avaliação dos modelos desenvolvidos.
- Divulgação científica dos resultados obtidos.

## 5 REFERENCIAL TEÓRICO *(apresenta de forma clara, concisa e direta, o embasamento teórico, os autores que servirão de base, conforme a trajetória definida no projeto)*

Monitorar o comportamento de diferentes espécies de peixes é de primordial importância para obter informações sobre um sistema ecológico marinho. A contagem e distribuição das várias espécies de peixes podem fornecer informações valiosas sobre a saúde do sistema ecológico e podem ser usadas como parâmetro para monitorar as mudanças ambientais. A classificação dos peixes também pode ajudar a traçar seus movimentos e dar padrões e tendências em suas atividades, proporcionando um conhecimento mais profundo sobre a espécie como um todo.

O trabalho de Malde et. al. (2019) mostra que nos últimos anos, houve avanços rápidos nos campos de inteligência artificial e aprendizado de máquina e, em particular, os chamados sistemas de aprendizado profundo, que agora são capazes de resolver tarefas complexas que anteriormente exigiam experiência humana. Os autores destacam que essa

tecnologia é diretamente aplicável a muitos problemas importantes de análise de dados e fornece as ferramentas necessárias para resolver muitos desafios complexos na ciência marinha e no gerenciamento de recursos. Neste trabalho, os autores fazem uma breve revisão dos desenvolvimentos recentes em aprendizado profundo e destacamos as muitas oportunidades e desafios para a adoção efetiva dessa tecnologia nas ciências marinhas.

O trabalho de Rathi et. al. (2017) propõe uma forma classificação automatizada de espécies de peixes. Essa classificação de peixes de alta precisão é necessária para maior compreensão do comportamento dos peixes em Ictiologia e por biólogos marinhos. Este trabalho de Rathi et. al. (2017) utiliza técnicas baseada em Redes Neurais Convolucionais, Deep Learning e Processamento de Imagens para alcançar uma precisão de 96,29%.

Quimbayo et. al. (2021) elaborou um banco de dados de peixes recifais do Oceano Atlântico e do Oceano Pacífico Oriental que inclui 21 características distribuídas em comportamentais, morfológicas, e ecológica, bem como biogeográficas e aspectos de conservação. Este banco de dados de peixes de recife oferece a oportunidade de explorar novas questões ecológicas e evolutivas em diferentes escalas e fornece ferramentas para a conservação de espécies com base nessas características. Além disso, esse banco de dados possui informações que podem ser utilizadas como base de treinamento para modelos de Inteligência Artificial para classificação de peixes recifais.

## 6 MATERIAIS E MÉTODOS *(detalhamento de como será realizada a presente pesquisa. Quais materiais e métodos que serão utilizados, como será e, em que momento. As informações devem ser detalhadas e suficientes para que outros pesquisadores possam repetir o experimento)*

Este projeto de pesquisa propõe a criação de modelos de redes neurais artificiais para classificação de peixes recifais do Atlântico e do Pacífico Oriental, com base nas informações contidas no banco de dados criado por Quimbayo et. al. (2021).

Este projeto de pesquisa **não** utilizará animais das espécies classificadas como Filo Chordata e subfilo vertebrata (Lei 11.794/2008), uma vez que o banco de dados com as informações dos peixes já está pronto e disponível em formato de arquivo de texto, e de forma online para utilização ([link](#)). Toda a execução do projeto será realizada em laboratório de Informática, utilizando apenas computadores, sendo assim, a sua execução **não** oferecerá risco à saúde humana, de outros animais e nem risco ambiental.

O banco de dados, elaborado por Quimbayo et. al. (2021) inclui 21 características distribuídas em comportamentais (área de origem, atividade do ração, tamanho do grupo,

nível na coluna d'água, três medidas de temperatura preferida), morfológicas (tamanho corporal máximo, classe de tamanho, formato do corpo, proporção, nadadeira caudal, boca posição) e ecológica (nível trófico, dieta, estratégia de desova, profundidade de ocorrência, duas constantes alométricas, duração e tempo de vida das larvas pelágicas), bem como biogeográficas (índice de distribuição geográfica, extensão de distribuição, distribuição de espécies em 20 províncias marinhas, Latitude Norte e Sul de ocorrência, número total de províncias onde as espécies ocorrem, ocorrência na AT e EP) e aspectos de conservação (estado da IUCN, vulnerabilidade e preço de mercado global). Compilamos esses dados por meio de uma revisão cuidadosa de 104 checklists locais publicados entre 1982 e 2020, repositórios online, relatórios locais, livros e monografias sobre famílias ou gêneros específicos. Limitamos nosso banco de dados a localidades situadas entre as latitudes 51°N e 45°S que incluem habitats biogênicos e/ou recifes rochosos rasos e mesofóticos superiores. Nosso banco de dados abrange 2.198 espécies pertencentes a 146 famílias e 655 gêneros de peixes recifais distribuídos em dois reinos marinhos (1.458 no AT, 829 no EP e 89 nos dois reinos) e 20 províncias marinhas.

Na primeira parte do projeto, serão coletados e preparados os dados de conjuntos de dados usando a biblioteca Python com Pandas. Na segunda parte do projeto, serão criados e configurados os modelos de redes neurais usando a plataforma Python e PyTorch.

Para este projeto, usaremos o Python como linguagem de programação. Para coletar e preparar o conjunto de dados, usaremos a biblioteca Pandas. Para os modelos de redes neurais (treinamento e teste), usaremos a plataforma PyTorch. Para preparar os dados a serem usados nos modelos PyTorch, seguiremos este roteiro:

- Selecionar e extrair as características essenciais para a construção dos modelos;
- Separar os dados para treinamento;
- Separar os dados para teste;
- Treinamento dos modelos;
- Teste e avaliação dos modelos;
- Refinamento dos modelos;

Com os dados já separados e preparados, poderemos criar os modelos de redes neurais. Os modelos de redes neurais serão treinados com os conjuntos de dados preparados previamente. O treinamento consiste em alimentar a rede com os dados estatísticos fomentando com informações sobre as classes desejadas. A partir dos modelos treinados, passaremos a fase de testes. Os testes dos modelos consistem em utilizar os modelos já treinados para classificação das espécies com base em dados coletados que a rede não conhece, ou seja, dados que não fizeram parte do conjunto de treinamento da rede.

Os resultados dos testes mostrarão se o grau de precisão da previsão realizada pela rede neural. Caso os resultados não sejam satisfatórios, será necessário realizar os treinamentos novamente, porém com parâmetros diferentes até que os resultados alcancem um grau de precisão adequado.

**7 JUSTIFICATIVAS** *(relevância da pesquisa: prática e intelectual. Quais as contribuições para compreensão ou solução do problema que poderá advir com a realização de tal pesquisa)*

Os volumes crescentes de dados, o aumento da complexidade dos dados e a redução da qualidade dos dados representam desafios para a disciplina de ciências marinhas, mas, ao mesmo tempo, os avanços recentes no aprendizado de máquina oferecem novas possibilidades de abordá-los. Os sistemas para análise automática de dados podem ser considerados em vários níveis, desde tornar o trabalho manual mais eficiente até novas análises de dados complexos e heterogêneos.

Monitorar o efeito de medidas preventivas e de recuperação requer a estimativa da biomassa e abundância de peixes por meio de amostragem regular de suas populações em corpos d'água como lagos, rios e oceanos.

Manter um registro do número de peixes por espécie e marcar as espécies ameaçadas em grandes e pequenos corpos d'água é exigido pelas instituições envolvidas. A maioria dos métodos disponíveis se concentra na classificação de peixes fora da água, pois a classificação subaquática apresenta desafios como ruídos de fundo, distorção de imagens, presença de outros corpos d'água nas imagens, qualidade da imagem e oclusão.

**8 CRONOGRAMA FÍSICO DE EXECUÇÃO** *(elencar principais atividades a serem realizadas durante o período de execução do projeto)*

Atividades	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov	Dez
Grupo de estudo e revisão de literatura.	X	X							
Coleta de dados para pesquisa		X	X						
Extração de características dos dados coletados para pesquisa			X	X					
Construção do algoritmo de previsão, com ajustes e refinamentos.				X	X	X			
Experimentos						X	X		
Divulgação de resultados							X	X	
Elaboração de relatório final de atividades								X	X

**9 PLANO DE TRABALHO DO BOLSISTA** *(Detalha a participação do bolsista em cada atividade elencada)*

Atividades	Ações
Revisão de literatura	Coletar material bibliográfico referente a reconhecimento de padrões em ciências marinhas e biologia marinha.
Coleta de dados de pesquisa	Acessar o repositório de dados dos peixes recifais do Atlântico e Pacífico Oriental e coletar os dados determinados para realização do estudo.
Extração de características dos dados coletados	Realizar a extração de características e padrões importantes dos dados coletados utilizando a linguagem de programação Python e suas bibliotecas.
Construção do algoritmo de previsão	Concepção, desenvolvimento e refinamento do algoritmo de classificação de peixes recifais.
Experimentos	Condução da fase experimental do projeto de pesquisa com a realização de testes e experimentos sobre os modelos desenvolvidos.
Divulgação de resultados	Publicação e divulgação dos resultados dos experimentos em periódicos e eventos científicos.
Elaboração de relatório final	Redigir o relatório final de atividades do projeto de pesquisa, apresentando os resultados da pesquisa, e os comprovantes de divulgação científica.

**10 REFERÊNCIAS**

QUIMBAYO, Juan Pablo et al. Life-history traits, geographical range, and conservation aspects of reef fishes from the Atlantic and Eastern Pacific. *Ecology*, v. 102, n. 5, p. e03298, 2021.

Rathi, Dhruv; Jain, Sushant; and S. Indu, "Underwater Fish Species Classification using Convolutional Neural Network and Deep Learning," 2017 Ninth International Conference on Advances in Pattern Recognition (ICAPR), 2017, pp. 1-6, doi: 10.1109/ICAPR.2017.8593044.

Malde, Ketil; Handegard, Nils Olav; Line Eikvil, Arnt-Børre Salberg, Machine intelligence and the data-driven future of marine science, *ICES Journal of Marine Science*, Volume 77, Issue 4, July-August 2020, Pages 1274–1285, <https://doi.org/10.1093/icesjms/fsz057>