

Detecção e Contagem do Pirarucu utilizando Técnicas de Visão Computacional

¹Wandelany F. de Abreu,¹Deam J. A. da Silva,¹Roberto P. do Nascimento,¹Josivan R. dos Reis,¹Márcio J. M. da Ponte

¹Instituto de Engenharia e Geociência

¹Programa Computação

¹Universidade Federal do Oeste do Pará

Santarém, Pará

lanyfialho@hotmail.com {deam.silva, roberto.nascimento, josivan.reis, marcio.ponte}@ufopa.edu.br

Resumo—Na atividade de manejo do pirarucu (*Arapaima gigas*) é imprescindível a vigilância constante e a realização de contagem para determinar o quantitativo de indivíduos no ambiente manejado por comunidades, com intuito de informar ao Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis (IBAMA) a existência de quantidade suficiente que autoriza a captura de uma quota sem ocasionar sobrepesca do estoque pesqueiro. Neste contexto, este trabalho objetiva detectar e contar automaticamente o pirarucu, através das imagens de vídeos e utilizando técnicas de visão computacional fundamentado no modelo de mistura gaussiana (GMM) e segmentação baseada na análise de movimentos (SBAM), resultando na detecção e contagem do pirarucu, embasado no deslocamento que realiza ao emergir no lago a cada vinte minutos para respirar.

Abstract—In the management activity of pirarucu (*Arapaima gigas*), constant vigilance and counting are essential to determine the quantity of individuals in the community-managed environment, in order to inform the Brazilian Institute of Environment and Renewable Natural Resources (IBAMA) of existence of a sufficient quantity to authorize the capture of a quota without overfishing the fishing stock. In this context, this work aims to automatically detect and count the pirarucu through video images and using computer vision techniques based on the Gaussian Mixing Model (GMM) and motion analysis based segmentation (SBAM), resulting in the detection and counting of the pirarucu, based on its displacement as it emerges from the lake every twenty minutes to breathe.

I. INTRODUÇÃO

Os estoques naturais de pirarucu¹ vem demonstrando uma diminuição na abundância desse recurso no ambiente natural, o que se reflete na minimização no volume e no tamanho médio das mantas desembarcadas em alguns portos amazônicos, levando muitos administradores da pesca a afirmar que o recurso enfrenta situação de sobrepesca, tendo em vista que a gestão dessa espécie é difícil, pois a regulamentação sobre o ordenamento da pesca não é seguida em muitas regiões da Amazônia [1].

Mediante as determinações tomadas pelos órgãos ambientais com o intuito de minimizar a diminuição da abundância

¹pirarucu (*A. gigas*) é um dos maiores peixes de água doce do mundo, capaz de atingir três metros de comprimento e até 200 kg, vive em águas quentes (24° a 31°C) de diversos rios e lagos em países como o Brasil, Peru, Colômbia e Guiana, detectado principalmente em ambientes lênticos, como lagos e canais de áreas de várzea.

de pirarucus em toda bacia amazônica temos: Em 1989, o Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis (IBAMA) estipula o tamanho mínimo de captura para a espécie (150 cm), o IBAMA proíbe qualquer tipo de captura e venda de pirarucu, sendo a pesca permitida apenas em áreas manejadas ou provenientes de cultivo. Estas medidas foram tomadas devido a diminuição das quantidades desembarcadas nos diferentes portos da Bacia Amazônica [2].

Os censos de pirarucus (contagens) têm grande importância para o manejo, pois fornecem a base para estabelecer quotas de pesca. O manejador interage diretamente com o recurso quando fornece a quantidade de pirarucus na sua área e utiliza esta informação para indicar o quanto pode ser pescado sem prejudicar a população [2].

O curso pode ser ministrado por técnicos munidos de referencial teórico específico e com auxílio de pescadores experientes certificados no método de contagem [3].

A contagem de pirarucu é feita para que tenha a estimativa do número de pirarucus maiores que 100 cm presentes no lago, por meio da observação de uma unidade de área do lago de no máximo dois hectares em 20 minutos, reposicionando-se quantas vezes forem necessárias até atingir a extensão do lago. Existe a necessidade de promover atividades constantes direcionados à verificação dos estoques desta espécie [3].

Neste contexto, este trabalho apresenta uma nova proposta, com objetivo de detectar o pirarucu e realizar a contagem automática, mediante técnicas de visão computacional, utilizando imagens de vídeo, visando a obtenção uma estimativa da quantidade da espécie, baseado no modelo de mistura gaussiana e na segmentação baseada em análise de movimento. Além disso, se pretendeu avaliar a performance dessas técnicas quanto ao tempo de execução e utilização dos recursos computacional.

Este artigo está estruturado em seções: I introdução, II metodologia, III os resultados da proposta, IV avaliação de desempenho e por fim, na seção V, a conclusão.

II. MATERIAIS E MÉTODOS

Para implementação dos métodos utilizou-se a IDE Matlab versão R2018b *student use toolbox*, que possui bibliotecas com funções específicas de visão computacional, como o modelo

de mistura gaussiana e segmentação baseado na análise de movimento. Para o processo de medição do desempenho da técnica mais eficiente de detecção, utilizou-se algoritmos e da matriz de confusão. Na aquisição de imagens foram utilizados quatro vídeos retirados do documentário realizado pelo Projeto PróVarzea (Subprograma Unidades de Conservação e Manejo dos Recursos Naturais executado pelo IBAMA, que constituiu uma ação conjunta do Governo do Brasil e da Cooperação Internacional, que apoia estudos e análises estratégicas sobre os temas que afetam a conservação e o uso racional da várzea, em especial, os recursos pesqueiros) na região do baixo Amazonas em diferentes cenários e locais, sendo o único material encontrado em ferramentas de busca da web. Após a divisão dos vídeos em *frames* (imagens), vídeo1 apresentou uma carga de trabalho de 201 imagens, vídeo2 214 imagens, vídeo3 393 imagens e vídeo4 321 imagens todos com formato 360x640x3 de 24 bits. Em relação as métricas de desempenho, foi analisado o tempo de execução, através do *rund and time* (executor de tempo) do programa Matlab que exibe o desempenho de cada método, além disso, avaliará o consumo em percentual dos recursos computacional de cada técnica, por intermédio do gerenciador de tarefas do computador. A execução dos testes foi realizada em um notebook com as seguintes configurações: um sistema operacional Windows 10 Pro X64; Processador Intel® Core™ i3 – 5005U CPU @2.00GHz; 4GB de memória RAM e 1T de HD.

A. Modelo de Mistura Gaussiana (GMM)

O GMM é uma mistura de k distribuições gaussianas que indica a mudança de estado dos pixels correspondentes de um quadro para outro. O algoritmo desenvolvido aplica misturas gaussianas a cada quadro e transforma imagens coloridas em imagens binárias. Para os pixels correspondentes que não sofrem alterações de estado, o valor 0 (preto) é atribuído e para pixels que sofrem alterações drásticas no estado, o valor 1 (branco) é atribuído. Dessa forma, é possível gerar os locais de todos os pirarucus em movimento no vídeo [4].

B. Segmentação

A segmentação é um dos mais impressionantes aspectos da visão: a capacidade de separar regiões de interesse do restante de uma imagem. O procedimento consiste apenas em estabelecer um limiar (*threshold*) e converter os pixels maiores do que esse limiar para 1 e o restante dos pixels para 0 [5].

C. Detecção utilizando o Modelo de Mistura Gaussiana

A princípio, cada vídeo foi dividido em *frames* (quadros) que são sequência de imagens, com a finalidade de avaliar o total de *frames* e determinar o número de imagens destinadas ao treinamento.

1) *Aquisição da imagem*: foi determinado 12,59% do total de *frames* do vídeo para treinamento.

2) *Subtração de fundo*: a definição do plano de fundo transformação das imagens em *rgb* com valores de pixels de 0 a 255 em imagem binária, em que o valor 0 (preto) é atribuído para o *background* e 1 para o *foreground* que é a região de interesse.

3) *Erosão*: comparando as operações morfológicas abertura composta pela operação erosão (reduz o elemento estruturante) seguida da dilatação (expande) utiliza o mesmo elemento estruturante, em contrapartida, a fechamento não é o contrário da abertura, porém, uma não é inversa da outra. Analisando individualmente a erosão mostrou ser mais eficiente na minimização dos ruídos e definição da região de interesse.

4) *Detecção*: encontrada a região de interesse são inseridos retângulos de cor vermelha que representa um pirarucu encontrado. Conforme a figura 1.



Figura 1. Fluxograma do Modelo de Mistura Gaussiana

D. Modelo de Mistura Gaussiana utilizando filtro mediana 2-D (GMMMED2-D)

Na etapa de treinamento são utilizados 22% dos *frames* de vídeo. A imagem de entrada no formato RGB é convertida para grayscale que assume 256 valores inteiros, onde 0 (cor preta) e 255 (cor branca). Posteriormente, o filtro mediana 2-D ordena os valores da imagem, seleciona o valor que corresponde à posição central, realiza a troca do valor de cada pixel pela mediana dos pixels vizinhos. Definida a subtração de fundo, a operação de erosão é aplicada para eliminar ruídos e determinar a região de interesse detectada.

E. Modelo de Mistura Gaussiana utilizando filtro mediana 3-D (GMMMED3-D)

Foram utilizados 10% dos *frames* para treinamento. Inicialmente um filtro mediana 3-D de tamanho 3x3x3 é aplicado na imagem em RGB replicando os valores de maneira espelhada nas bordas. Posteriormente, a operação de erosão é processada e a região de interesse é detectada.

F. Modelo de Mistura Gaussiana utilizando filtro convolução

Foram estabelecidos 10% do total de *frames* para treinamento. De início a imagem é convertida em *grayscale*, seguidamente uma máscara de tamanho 3x3x3 contendo valores iguais a 0.1 é aplicada em toda a imagem determinando o *foreground*. Porém apresenta ruídos e com isso a operação erosão é aplicada e a região de interesse detectada.

G. Modelo de Mistura Gaussiana utilizando binarização

A princípio, na etapa de treinamento é utilizado 10% dos *frames* e a imagem adquirida na etapa de aquisição é

transformada em *grayscale*, em seguida um limiar (*threshold*) de 0.5 é estabelecido, mediante a função 1, que converte os pixels menores que esse limiar em 0, após, o *foreground* é determinado e aplicada a operação erosão em que a região de interesse é detectada.

$$f(x,y) = \begin{cases} 0, & f(x,y) < l \\ 1, & \text{outro caso} \end{cases} \quad (1)$$

Em que $f(x,y)$ é a função que representa a imagem e l é o limiar.

H. Segmentação baseada na análise de movimentos (SBAM)

A segmentação baseada na análise de movimentos, realiza um treinamento mediante um *frame* representativo do vídeo, em seguida é aplicado ao processamento de todos os quadros no vídeo. Subsequentemente, o *frame* representativo em RGB é transformado em níveis de cinza (*grayscale*) assumindo 256 valores inteiros entre 0 (cor preta) e 255 (cor branca). Após as transformações os valores de pixels abaixo de 10 tornam-se plano de fundo. A maioria dos objetos de cor escura são removidos, mas outros estranhos permanecem, principalmente a vegetação, canoa e pescadores. No intuito de remover esses ruídos, a operação abertura morfológica foi implementada com objetivo de remover objetos pequenos da imagem, preservando os de estruturas grandes atingindo a região de interesse.

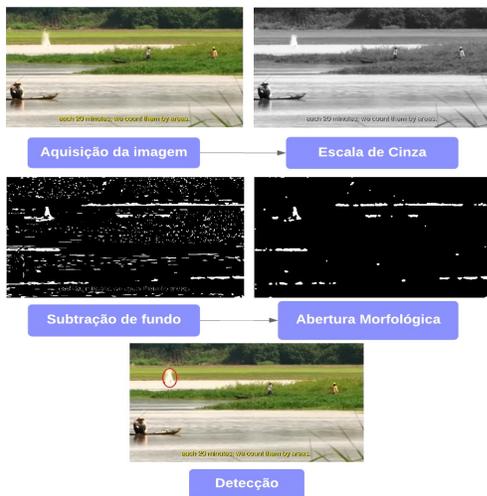


Figura 2.. Fluxograma da Segmentação baseada na análise de movimentos

III. RESULTADOS

Na avaliação dos resultados obtidos com os diferentes métodos, foram utilizadas quatro medidas, simplificadas na matriz de confusão da Tabela 1. A matriz de confusão é uma matriz de erros e acertos, que permite analisar a capacidade do método em reconhecer diferentes classes de instâncias. Verdadeiro positivo (VP) e Verdadeiro negativo (VN) indicam o quanto o classificador está acertando nas predições, enquanto falso positivo (FP) e falso negativo (FN) indicam o quanto o classificador está errando nas predições [6].

A medida de acurácia (A) e precisão (P) são:

$$A = \frac{(VP+VN)*100}{(VP+VN+FP+FN)} \quad P = \frac{(VP*100)}{(VP+FP)}$$

Tabela 1. Modelo de Classificação da Matrix de Confusão

		Pirarucu	
		Positivo	Negativo
Pirarucu	Positivo	Verdadeiro Positivo (VP)	Falso Positivo (FP)
	Negativo	Falso Negativo (FN)	Verdadeiro Negativo (VN)

Os resultados baseados na matriz de confusão e Acurácia estão presentes na tabela 2.

Tabela 2. Resultado da Acurácia em relação aos métodos

Dados	Taxa de Acerto (%)					SBAM
	GMM	GMM MED 2 - D	GMM MED 3 - D	GMM Conv.	GMM Bin.	
Vídeo1	9,00	17,39	55,55	50,98	50,00	60,00
Vídeo2	66,66	60,00	72,72	60,00	66,66	100,00
Vídeo3	15,00	15,38	46,51	88,23	46,87	71,42
Vídeo4	37,50	45,45	50,00	52,63	50,00	75,00

No pior caso, o modelo de mistura gaussiana para a entrada vídeo 1 exibe 9% e o vídeo 3, 15% de predições corretas. Porém a segmentação baseada na análise de movimentos apresenta o melhor caso, com 100% de acerto, conforme tabela 2.

Os resultados baseados na matriz de confusão e Precisão estão presentes na tabela 3

Tabela 3. Resultado da Precisão em relação aos métodos

Dados	Precisão (%)					SBAM
	GMM	GMM MED 2 - D	GMM MED 3 - D	GMM Conv.	GMM Bin.	
Vídeo1	6,25	3,22	13,63	10,71	7,31	33,33
Vídeo2	50,00	50,00	50,00	50,00	50,00	100
Vídeo3	10,53	5,55	8,33	66,66	10,52	100
Vídeo4	16,67	28,57	14,28	18,18	28,57	66,66

Resultados da medida de precisão da classe verdadeiros positivos, apresenta no pior caso o modelo de mistura gaussiana quando a entrada são os vídeos 1, 2, 3 que exibem respectivamente 6,25%, 10,53% e 16,67% de precisão. Porém a segmentação para a entrada vídeo 2 e 3, mostram no melhor caso 100% do total de detecções verdadeiro positivo. Conforme, tabela 3

IV. AVALIAÇÃO DE DESEMPENHO

1) *Tempo de execução dos métodos:* Avaliando o tempo de execução em relação aos métodos, o modelo de mistura gaussiana utilizando binarização apresenta o maior tempo de execução apresentando 138s, como pode ser observado na figura 3.

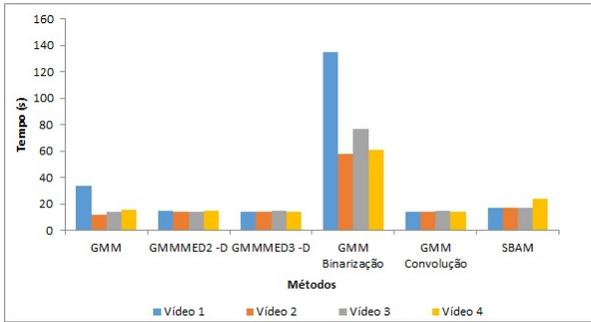


Figura 3. Tempo de execução em relação aos métodos

2) *Consumo em percentual de ocupação da unidade central de processamento em relação aos métodos:* Analisando o consumo de CPU em relação aos métodos o modelo de mistura gaussiana utilizando filtro mediana 2-D apresenta no geral o maior consumo, respectivamente 46%, 75%, 56% e 72%, como pode ser observado na figura 4..

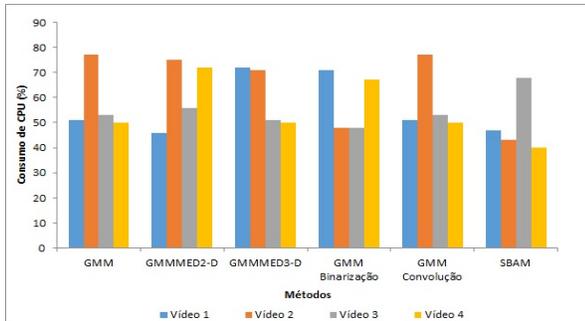


Figura 4. Consumo de CPU em relação aos métodos

3) *Consumo em percentual de ocupação de memória em relação aos métodos:* Verificando o consumo de memória em relação aos métodos em geral a segmentação baseada na análise de movimentos requer maior percentual de memória, apresentando respectivamente 83%, 85%, 90% e 88%, conforme a figura 5.

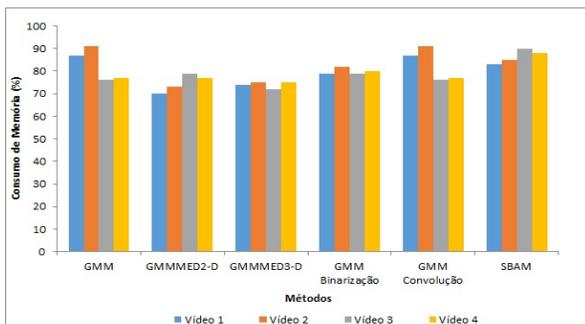


Figura 5. Consumo de Memória em relação aos métodos

V. CONCLUSÃO

Detectar o pirarucu em um lago é uma tarefa complexa, pois existe diversos fatores que dificultam a identificação do movimento no ambiente. Porém possível de ser identificado utilizando técnicas em visão computacional. Dessa forma, a segmentação baseada na análise de movimento mostrou ser

a mais eficiente para detecção do *Arapaima*. Além disso, fez-se a avaliação de desempenho do tempo de execução e dos percentuais de ocupação dos recursos computacionais em relação à execução dos métodos, considerando um fator importante para escolha de uma técnica, pois podem ser críticos e essenciais no desenvolvimento da implementação dos métodos e solução do problema.

Por fim, como produções futuras serão desenvolvidos *in loco*, a aquisição de imagens durante a contagem manual realizada pelos pescadores em uma determinada comunidade, com a finalidade de criar uma base de dados e avaliar a precisão da contagem automática utilizando os métodos proposto em relação a realizada pelos pescadores.

REFERÊNCIAS

- [1] N. M. N. Silva, A. A. Silva, T. M. P. Braga, and C. H. F. Júnior, "Diagnóstico do comércio de pirarucu nos mercados e feiras de santarém, Pará." *Biota Amazônia*, vol. 6, no. 4, pp. 49–53, 2016.
- [2] H. Q. A. and W. Sardinha, *A preservação e o uso sustentado dos pirarucus em Mamirauá*, H. Queiroz and W. Crampton, Eds. Sociedade Civil Mamirauá MCT-CNPq, 1999.
- [3] R. Silva, A. Gonçalves, and J. Marinho, *Contagem e Censo Populacional de Pirarucu*, E. Venturi, Ed. Tefé, AM: IDSM, 2013.
- [4] D. Santosh, P. P., and R. L. K. N., "Tracking multiple moving objects using gaussian mixture model," *International Journal of Soft Computing and Engineering (IJSCE)*, vol. 3, no. 2, 2013.
- [5] A. Backes and J. Junior, *Introdução à Visão Computacional usando Matlab*. Rio de Janeiro: Alta Books, 2016.
- [6] C. Souza and P. Santos, "Probabilistic logic reasoning about traffic scenes," *International Journal of Soft Computing and Engineering (IJSCE)*, vol. 3, no. 2, 2011.