

Acerca de Técnicas de Aumento de Dados para a Detecção Automática de Focos de Mosquito

Bettina D. Barros^{1,2}, Wesley L. Passos¹, Rafael Padilla¹, Sergio L. Netto¹, Eduardo A. B. da Silva¹, Gabriel M. Araujo³

¹{wesley.passos, rafael.padilla}@smt.ufrj.br, PEE/COPPE, Universidade Federal do Rio de Janeiro.

²bettina.barros@ntnu.no, IES, Universidade Norueguesa de Ciência e Tecnologia, Trondheim, Noruega.

³gmatos1983@gmail.com, Centro Federal de Educação Tecnológica Celso Suckow da Fonseca, Nova Iguaçu.

Resumo— Este trabalho discute técnicas de aumento de dados para detectar potenciais focos de mosquito usando vídeos gravados por um drone. Um sistema de visão computacional é proposto para identificar objetos associados a criadouros. Uma base de dados composta por seis vídeos aéreos contendo objetos como caixas d'água, pneus e garrafas é desenvolvida. Devido a dificuldade de obter uma base extensiva de cenários reais, técnicas de aumento artificial de dados são apresentadas e três métodos para inserir imagens de objetos em vídeos são considerados. Uma rede neural convolucional é utilizada para avaliar essas técnicas, indicando que o aumento artificial de dados reduz o sobreajuste, melhorando o desempenho geral da detecção.

Palavras-Chave— aumento de dados, detecção de objetos, focos de mosquito

Abstract— This work discusses data augmentation techniques for detecting mosquito breeding grounds using videos recorded by a drone. A computer-vision system is proposed for automatically detecting mosquito-breeding related objects. A database composed of six aerial videos containing breeding-related objects, such as water tanks, tires, and bottles, is devised. Due to the difficulty of obtaining extensive records of real scenarios, artificial data augmentation techniques are presented, and three methods for inserting objects into videos are considered. A convolutional neural network detector is used to evaluate these techniques, indicating that artificial data augmentation reduces overfitting, improving the overall detection performance.

Keywords— data augmentation, object detection, mosquito breeding grounds

I. INTRODUÇÃO

Segundo a Organização Mundial da Saúde (OMS), os mosquitos disseminam doenças que afetam 700 milhões de pessoas e causam cerca de um milhão de mortes a cada ano [1]. Dengue, chikungunya e zika juntas afetam milhões de pessoas por ano e são consideradas uma ameaça para pessoas que vivem em diversos países tropicais. O principal vetor dessas doenças é o mosquito *Aedes aegypti*, que se reproduz principalmente em água limpa acumulada em recipientes artificiais. As diretrizes da OMS alertam sobre os riscos de objetos com água acumulada, onde, para fins epidêmicos, os objetos críticos são reduzidos a caixas d'água, garrafas, baldes, pneus descartados e escoadouros externos [2]. Na ausência de uma vacina eficaz como método de controle do mosquito, os agentes de saúde realizam buscas por criadouros diariamente. Para auxiliar esses agentes, este trabalho propõe um sistema para detecção automática de objetos associados a criadouros.

Usando um drone para mapear uma área e técnicas de visão computacional para detectar os objetos críticos, o sistema torna o trabalho dos agentes mais rápido e eficaz.

As contribuições deste trabalho são: (i) investigação do problema prático e uma proposta completa de solução; (ii) desenvolvimento de uma versão inicial da base de dados de vídeos contendo objetos associados a alto potencial criadouro de mosquitos, como pneus, garrafas e caixas d'água; (iii) construção de uma base de dados de imagens de pneus para aumento artificial da base de dados de vídeos gravada; (iv) proposta de métodos de inserção para gerar amostras adicionais de vídeo de treinamento; (v) uso de uma rede neural para realizar a detecção de pneus usando a base de dados construída em suas versões original e estendida.

Para introduzir tais contribuições, este trabalho está organizado da seguinte forma: A Seção II descreve um sistema para detectar automaticamente objetos relacionados à reprodução de mosquitos. Na Seção III, uma base de dados contendo potenciais focos de mosquitos é desenvolvida utilizando um drone para gravação dos vídeos. Devido à dificuldade de obter registros extensivos de cenários reais, técnicas de aumento artificial de dados são apresentadas na Seção IV. A Seção V descreve uma rede neural para a detecção automática de objetos críticos, enquanto a Seção VI inclui experimentos que avaliam a eficácia das técnicas de aumento de dados. Por fim, a Seção VII apresenta as conclusões do artigo.

II. PROPOSTA DE SISTEMA

O sistema de detecção de objetos almejado deve ser capaz de, através da varredura de uma área pré-definida com um drone, gerar um mapa que indique as posições de possíveis criadouros mosquitos. O objetivo é fornecer um sistema de apoio à decisão aos agentes de saúde que possa acelerar as ações preventivas. Ao contrário dos trabalhos relacionados na literatura, o sistema proposto utiliza imagens aéreas, onde os criadouros geralmente aparecem em baixa resolução. Além disso, o sistema proposto se concentra na detecção de objetos associados a doenças transmitidas por mosquitos e não apenas em poças de água. A Figura 1 mostra um diagrama de blocos para o sistema proposto, que é dividido em três etapas: preparação de dados, desenvolvimento do sistema e uso do sistema.

Recebendo coordenadas GPS de uma área como entrada, o primeiro passo da preparação de dados é capturar vídeos sobre-

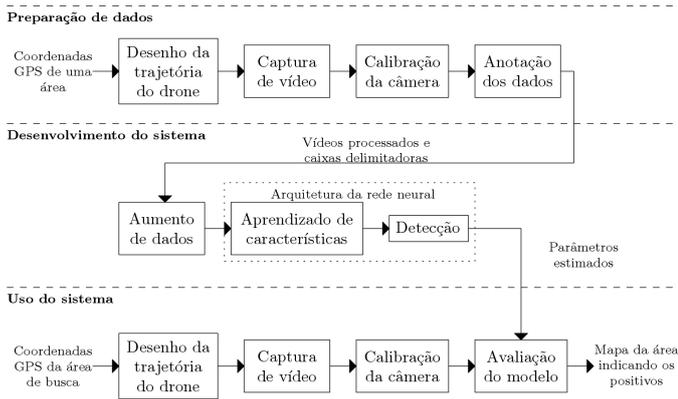


Fig. 1: Diagrama de blocos do sistema de detecção de objetos proposto.

voando a área definida. Em seguida, a calibração da câmera é realizada para remover distorções nos quadros de vídeo e, por fim, as caixas delimitadoras são anotadas, indicando os locais com objetos críticos. A etapa de preparação de dados deve ser executada em várias áreas com características diferentes para resultar em um banco de dados diversificado e representativo do problema.

A etapa de desenvolvimento do sistema recebe os vídeos processados e as caixas delimitadoras da etapa de preparação de dados e executa o aumento artificial de dados, o aprendizado de características e a detecção de objetos. O aumento de dados é uma técnica que busca reduzir o erro de generalização ao ampliar o conjunto de treinamento. O aprendizado de características e a detecção de objetos são realizados, por exemplo, por uma rede neural convolucional.

Por fim, o uso do sistema mostra como localizar os potenciais criadouros de mosquitos em uma área de busca. Os mesmos primeiros passos da preparação de dados são executados: desenho da trajetória do drone, captura de vídeo e calibração da câmera. Depois disso, usando os parâmetros estimados no desenvolvimento do sistema, o modelo é avaliado e um mapa de calor deverá ser construído (em versões futuras do sistema) contendo os locais com potenciais criadouros.

III. BASE DE DADOS DE VÍDEOS

Esta seção apresenta uma descrição da versão inicial atual do banco de dados “Mosquito Breeding Grounds”(MBG) criado para detectar focos de mosquitos em vídeos aéreos capturados por um drone [3]. Existem diversas bases de dados para trabalhar com detecção de água estagnada, como em [4]–[6]. No entanto, até onde sabemos, [4] é o único que trabalha com a detecção de objetos associados a doenças transmitidas por mosquitos. A base de dados de vídeo MBG contribui com uma extensão a diversos tipos objetos e inclui as caixas delimitadoras desses objetos e a telemetria do drone, permitindo a localização de cada objeto de interesse. Dessa forma, uma base de dados adequada para o problema de aprendizagem supervisionada para detectar e classificar objetos em vídeos aéreos deve conter: (i) controle do número máximo de parâmetros ou ao menos sua medida; (ii) expressivo número de amostras e representatividade de cada classe de objeto; (iii)

variabilidade do terreno, objetos, luminosidade e altura; (iv) correção das distorções de câmera; e (v) anotação confiável.

Como drones têm muitas restrições de voo, a principal dificuldade para construir tal base de dados é obter acesso a diversos tipos de cenários reais. Mesmo que o acesso ao terreno esteja disponível para vários locais, também é necessário saber onde estão os objetos relacionados a criadouros. Desta forma, a solução adotada neste trabalho é alterar manualmente os locais onde o acesso a pé está disponível, inserindo objetos representativos no cenário.

A câmera do drone foi calibrada usando o método de Zhang [7] para remover a maioria das distorções da lente. Para permitir uma abordagem supervisionada, os objetos de interesse nos vídeos gravados foram anotados quadro a quadro com caixas delimitadoras. Para acelerar esse processo, foi utilizado um sistema que interpola linearmente as caixas delimitadoras sobre o intervalo de 10 quadros. Cada vídeo na base de dados MBG possui um arquivo de anotação, contendo a caixa delimitadora dos objetos de interesse para cada quadro, e um arquivo de telemetria, que armazena os valores de vários parâmetros do drone durante a gravação. A primeira versão da base de dados compreende um total de aproximadamente 14 minutos de vídeo dividido em 6 partes, conforme ilustrado na Figura 2 e detalhado na Tabela I.

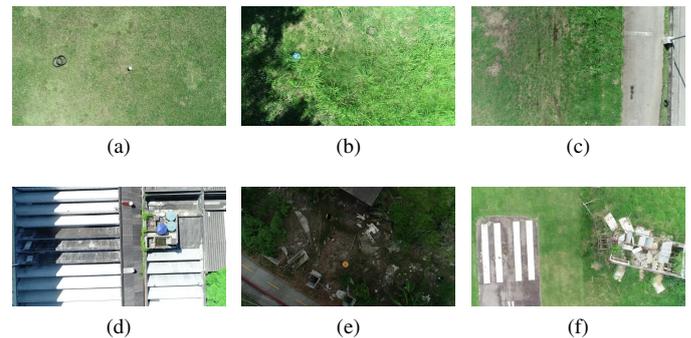


Fig. 2: Exemplos de quadros de vídeo em diferentes locais de gravação: (a) Grêmio, (b) BLI1, (c) BLI2, (d) BLI3, (e) CCMN e (f) FAU.

TABELA I: Atributos das gravações feitas em cada local.

nome do local	tipo do terreno	altura do drone	objetos anotados	duração do vídeo	área coberta
Grêmio	grama baixa	10 m	18	212 s	0.47 m ²
BLI1	grama alta	10 m	14	97 s	0.24 m ²
BLI2	rua e grama baixa	20 m	16	23 s	0.12 m ²
BLI3	rua e edifícios	40 m	0	221 s	3.99 m ²
CCMN	terreno baldio	15 m	0	75 s	0.40 m ²
FAU	terreno baldio	20 m	0	187 s	1.34 m ²
Total	-	-	48	14 min	6.56 m²

IV. AUMENTO ARTIFICIAL DE DADOS

Sistemas de classificação de imagem e detecção de objetos geralmente exigem grandes conjuntos de treinamento. Apesar de todos os esforços para construir uma base de dados extensa e diversificada, o processo de captura de dados e anotação

manual é exaustivo. Portanto, para aumentar o desempenho de tais sistemas, o aumento artificial de dados é amplamente utilizado.

Recentemente, muitos trabalhos estão adotando redes neurais para realizar o aumento de dados. Trabalhos como [8] sintetizam uma nova imagem usando GANs (do inglês *generative adversarial networks*) [9] para transferir o estilo entre duas imagens de entrada. Outros trabalhos [10], [11] estudam GLO (do inglês *generative latent optimization*) [12] para criar objetos sintéticos. Essas novas amostras podem ser usadas para aumentar o número de instâncias no conjunto de treinamento de um detector de objetos. Em [10], os autores exploram a detecção de arraias em vídeos aéreos capturados por um drone e uma rede baseada em GLO é usada para criar imagens semelhantes a arraias sob a água, realçando a detecção.

A referência [13] apresenta uma técnica de aumento de dados que corta e cola imagens de objetos em cenas de fundo utilizando transformações tradicionais, sem se preocupar com realismo. Os autores medem seu desempenho em uma base de dados para detectar objetos pequenos específicos em fotos de ambientes residenciais e de escritórios, e o uso da base de dados aumentada melhora o desempenho em 21%. Seguindo essa idéia simples, [14] adiciona informações de contexto de cena e estende a estrutura para base de dados com diferentes ambientes como cenas de fundo.

A. Base de dados de imagens de pneus

Nesta subseção, é apresentada uma base de dados de imagens segmentadas composta de fotos tiradas em ambiente fechado de seis pneus diferentes. Usando a câmera de um Moto G5S Plus, com resolução de 13-MP, a base de dados proposta contém 62 imagens com diferentes disposições dos pneus em relação à orientação e também a combinação de pneus, como mostrado na figura 3.

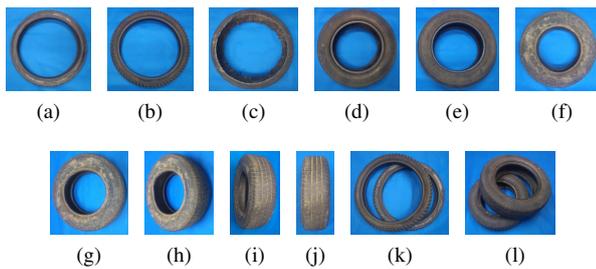


Fig. 3: As Subfiguras (a)-(f) mostram os seis diferentes pneus, (f)-(j) as orientações escolhidas para um pneu e (k)-(l) exemplos de combinação entre os pneus.

Uma lona azul foi usada como fundo para as fotos e uma simples segmentação de imagem baseada na matiz foi aplicada para remover o fundo e gerar uma máscara de segmentação. Para atenuar as bordas azuladas, uma operação morfológica de erosão [15] foi aplicada usando um elemento estrutural circular 3×3 em cinco iterações de erosão.

Quatro ações foram executadas para cada imagem segmentada: introdução de um ganho de luminância arbitrário, inversão horizontal, rotação por um ângulo arbitrário e redimensionamento. Para evitar desviar muito da realidade, a média

de luminância deslocada permaneceu dentro de $[50, 150]$ e as imagens redimensionadas ficaram restritas ao intervalo de $[0, 5; 0, 8]$. Ao girar e redimensionar a imagem, a máscara é retornada como uma imagem em tons de cinza e, portanto, um limite de 127 foi definido para manter os valores entre 0 ou 255, puramente preto e branco, respectivamente.

B. Métodos de inserção

Dada uma imagem de objeto com sua máscara e um quadro de vídeo com uma posição indicada no fundo, um método de inserção é definido neste trabalho como uma técnica para inserir a imagem no plano de fundo. Três diferentes métodos de inserção são apresentados e subjetivamente comparados neste trabalho.

1) *Colar*: O primeiro método de inserção apenas cola a imagem no plano de fundo, sem qualquer outra consideração. Em vez de aumentar o desempenho do detector, esse método introduz artefatos nas bordas que tendem a piorar os resultados de detecção.

2) *Mesclar*: O segundo método de inserção, da mesma forma que em [13], mescla a borda do objeto com o fundo ao colar a imagem. A mesclagem é feita usando a máscara do objeto depois de aplicar um filtro gaussiano para suavizar a borda do objeto.

3) *Desfocar e mesclar*: Embora o segundo método explore a mesclagem para atenuar os artefatos de borda introduzidos pelo primeiro método de inserção, em muitos casos, os objetos ainda se destacam visualmente do plano de fundo, parecendo muito artificiais. Portanto, o último método estuda o desfoque para obter uma imagem inserida com características de foco mais próximo do fundo.

Um filtro gaussiano bidimensional com a mesma variação em ambas as direções é aplicado para desfocar os objetos e um filtro laplaciano do gaussiano (LdG) é utilizado para avaliar o desfoque. O filtro LdG é definido com base na altura do drone e a medida de desfoque do plano de fundo τ_{ref}^2 é calculada e armazenada como um valor de referência. Em seguida, as iterações para desfocar o objeto são executadas para obter uma imagem com características de foco mais próximas do fundo. A variância do filtro gaussiano usado para desfocar o objeto é inicializada como $\sigma_0^2 = 0,5$ e atualizada por $\Delta\sigma = 0,1$ em cada iteração. Esse processo termina quando a medida de desfoque da imagem mesclada τ_i^2 é menor que o valor de referência τ_{ref}^2 , resultando em uma imagem com medida de desfoque equalizada.

A figura 4 mostra os resultados ao aplicar os três métodos de inserção. O método de inserção de colagem introduz artefatos de borda enquanto o método de inserção de mesclagem resolve esse problema. Além disso, o método de inserção com desfoque e mesclagem desfoca a imagem de objeto artificial para obter uma medida de desfoque que fica mais próxima do fundo, o que pode ajudar a detectar objetos reais.

C. A base de dados MBG de vídeos aumentada

Ao realizar o aumento de dados em vídeos, os quadros consecutivos devem manter consistência em relação à posição do objeto no solo. No esquema de gravação projetado, as



Fig. 4: Comparação dos métodos de inserção.

trajetórias dos drones podem ser consideradas apenas translacionais e, portanto, o uso de correlação de fase é adequado para estimar o movimento do drone entre quadros consecutivos [16]. A Figura 5 ilustra o resultado do aumento de dados de vídeo usando correlação de fase para estimar o movimento da câmera.



Fig. 5: Exemplo de inserção de objeto artificial em um determinado trecho de vídeo da base de dados MBG.

Em contraste com ampliar do número de objetos na base de dados, é importante manter parte das regiões intactas para garantir ambientes reais no conjunto de treinamento. Uma proporção aproximada de um objeto a cada 20 m² é mantida e, portanto, a base de dados aumentada tem, por exemplo, uma média aproximada de 20 objetos sintéticos por quadro quando uma altura de 20 m é utilizada. Além disso, objetos inseridos não ocluem outro objeto, especialmente um objeto real anotado. Os locais são escolhidos aleatoriamente e os três métodos de inserção usam a mesma configuração, resultando em quadros conforme o exemplo ilustrado na Figura 6.

V. DETETOR DE OBJETOS

A detecção de objetos envolve localização e classificação de objetos. A localização pode ser feita como uma máscara geral, identificando para cada pixel da imagem se pertence ou não a um objeto. No entanto, as máscaras de segmentação geralmente exigem muitos parâmetros e, se uma abordagem supervisionada é escolhida, torna a tarefa de anotação muito



Fig. 6: Exemplo de quadro do vídeo BLI2 (15 m de altura) da base de dados de vídeo MBG aumentada. Os pneus reais estão enquadrados em azul e os sintéticos em vermelho.

trabalhosa. Ao invés de usar máscaras, as caixas delimitadoras reduzem cada marcação de objeto a apenas cinco parâmetros: as duas coordenadas do ponto central, o tamanho do objeto como altura e largura e o rótulo da classe.

Os atuais detectores de objetos do estado da arte são normalmente baseados em redes neurais convolucionais profundas (CNNs) [17]. Uma CNN bem conhecida para detecção e classificação de objetos é a chamada rede neural YOLO [18]. A rede YOLO tem 19 camadas convolucionais e cinco camadas de maxpooling. Durante este trabalho, todas as redes da YOLO foram treinadas com uma taxa de aprendizado de 0,001, durante 1.000 iterações usando 64 imagens a cada iteração.

Para treinar uma rede YOLO, imagens foram extraídas da base de dados de MBG aumentada. A base de dados MBG, em sua maior resolução, contém objetos classificados como: (i) objetos pequenos: área < 32² pixels ao quadrado; (ii) objetos médios: 32² < área < 96²; (iii) objetos grandes: área > 96². Portanto, para evitar a redução de resoluções de objetos, os quadros extraídos são cortados em imagens de resolução mais baixa. Para isso, uma janela deslizante é usada com o mesmo tamanho que as entradas de rede, 832 × 832, e o maior passo possível cobrindo totalmente os quadros originais de tamanho 3.840 × 2.160.

Três vídeos da MBG foram usados para treinar a rede YOLO para detecção de pneus e outros três vídeos foram usados para testar o resultado da acurácia de detecção. Dessa forma, o conjunto de treinamento possui 54 imagens contendo caixas delimitadoras de pneus reais. O detector de pneus proposto utiliza os pesos pré-treinados da PascalVOC na arquitetura YOLOv2. A rede YOLO considera intrinsecamente algumas transformações como técnicas aumento de dados para melhorar seu desempenho. Portanto, neste trabalho, o treinamento da YOLO é executado considerando os seguintes cenários:

- Caso 1: Sem aumento de dados.
- Caso 2: Aumento de dados intrínsecos da YOLO.
- Caso 3: Aumento de dados por inserção de objetos.
- Caso 4: Aumento de dados intrínsecos da YOLO e por inserção de objetos.

VI. RESULTADOS EXPERIMENTAIS INICIAIS

A base de dados MBG aumentada possui três versões, uma para cada método de inserção. Para uma comparação justa destas três abordagens, os mesmos parâmetros de inserção de pneus são utilizados nos três métodos quando se escolhe aleatoriamente um pneu, os parâmetros de transformação e a localização do pneu. Dos vídeos do conjunto de treinamento, 330 quadros são extraídos resultando em 4.950 imagens com 18 pneus reais e 1.237 artificiais. No conjunto de teste, 450 imagens contendo 13 pneus reais foram avaliadas.

A Tabela II mostra o número total de caixas delimitadoras como falsos e verdadeiros positivos previstas pela YOLO no conjunto de testes.

TABELA II: Número de caixas demilitadoras como verdadeiros e falsos positivos (VP_b e FP_b) com probabilidade de pertencer à classe pneu maior que 0,5 em um total de 450 imagens.

	CCMN		FAU		BLI3	
	VP_b	FP_b	VP_b	FP_b	VP_b	FP_b
Caso 1	0	27	8	27	7	71
Caso 2	0	11	5	13	3	60
Caso 3 (Colar)	0	0	3	4	4	6
Caso 3 (Mesclar)	0	0	2	2	5	3
Caso 3 (Desfocar+Mesclar)	0	1	10	2	7	7
Caso 4	0	0	7	3	5	4

Os resultados experimentais mostram que, sem aumento de dados, a rede tende a detectar muitos falsos positivos. Usando o aumento de dados da YOLO, o número de falsos positivos é menor, mas também reduz o número de detecções de pneus. Em termos dos objetos em si, nenhuma das redes detectou pneus do vídeo CCMN. Um possível motivo é que esse vídeo é muito mais escuro do que os usados para treinamento, dificultando a detecção. Em contraste, todas as seis redes são capazes de detectar pelo menos dois dos três pneus do vídeo BLI3. Um estudo mais extenso sobre os parâmetros da rede pode encontrar um compromisso adequado entre verdadeiros e falsos positivos.

Comparando os métodos de inserção, o método de desfoque e mesclagem apresenta uma vantagem o número de falsos positivos é muito baixo e possui o maior número de detecções de pneus nos vídeos FAU e BLI3. Isso indica que abordagens que tentam tornar a tarefa de detecção de objetos mais difícil durante o treinamento podem ser promissoras para aumentar a acurácia geral da detecção.

VII. CONCLUSÃO

Este artigo apresenta uma abordagem para a detecção de objetos relacionados a criadouros de mosquitos. Começando por entender o problema que as organizações de saúde estão enfrentando em relação às doenças transmitidas por mosquitos, um sistema é proposto para detectar automaticamente os objetos de interesse. Uma base de dados de vídeo inicial é detalhada. Apesar de focado numa única classe de objetos (pneu), em futuras etapas o trabalho será estendido para outras classes de objetos de interesse (caixas d'água, calhas, etc.) Como os sistemas de detecção de objetos do estado da arte geralmente

requerem grandes conjuntos de treinamento, as técnicas de aumento de dados são introduzidas. Os resultados usando um detector de objetos YOLO indicam que o aumento artificial de dados reduz o sobreajuste, melhorando o desempenho geral da detecção pela rede proposta.

AGRADECIMENTO

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

REFERÊNCIAS

- [1] H. Caraballo and K. King, "Emergency department management of mosquito-borne illness: malaria, dengue, and west nile virus," *Emergency Medicine Practice*, vol. 16, no. 5, pp. 1–23, 2014.
- [2] M. T. F. Lagrotta, "Geoprocessamento de indicadores entomológicos na identificação de áreas, imóveis e recipientes (chaves) no controle do *Aedes aegypti*," Ph.D. dissertation, 2006.
- [3] W. L. Passos, T. M. Dias, H. A. Junior, B. D. Barros, G. Araujo, A. Lima, E. da Silva, and S. L. Netto, "Acerca da detecção automática de focos do mosquito *Aedes aegypti*," in *Anais do Simpósio Brasileiro de Telecomunicações e Processamento de Sinais*, 2018.
- [4] A. Agarwal, U. Chaudhuri, S. Chaudhuri, and G. Seetharaman, "Detection of potential mosquito breeding sites based on community sourced geotagged images," in *Proc. Geospatial InfoFusion and Video Analytics IV; and Motion Imagery for ISR and Situational Awareness II*, vol. 9089, 2014, p. 90890M.
- [5] M. G. Prasad, A. Chakraborty, R. Chalasani, and S. Chandran, "Quadcopter-based stagnant water identification," in *Proc. National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics*, 2015, pp. 1–4.
- [6] M. Mehra, A. Bagri, X. Jiang, and J. Ortiz, "Image analysis for identifying mosquito breeding grounds," in *Proc. IEEE International Conference on Sensing, Communication and Networking*, 2016, pp. 1–6.
- [7] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 11, pp. 1330–1334, 2000.
- [8] L. Perez and J. Wang, "The effectiveness of data augmentation in image classification using deep learning," *arXiv preprint arXiv:1712.04621*, 2017.
- [9] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Proc. Advances in Neural Information Processing Systems*, 2014, pp. 2672–2680.
- [10] Y.-M. Chou, C.-H. Chen, K.-H. Liu, and C.-S. Chen, "Changing background to foreground: an augmentation method based on conditional generative network for stingray detection," in *Proc. IEEE International Conference on Image Processing*, 2018, pp. 2740–2744.
- [11] C. Ledig, Z. Wang, W. Shi, L. Theis, F. Huszar, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani *et al.*, "Photo-realistic single image super-resolution using a generative adversarial network," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 105–114.
- [12] P. Bojanowski, A. Joulin, D. Lopez-Paz, and A. Szlam, "Optimizing the latent space of generative networks," *arXiv preprint arXiv:1707.05776*, 2017.
- [13] D. Dwibedi, I. Misra, and M. Hebert, "Cut, paste and learn: surprisingly easy synthesis for instance detection," in *Proc. IEEE International Conference on Computer Vision*, 2017, pp. 1310–1319.
- [14] N. Dvornik, J. Mairal, and C. Schmid, "On the importance of visual context for data augmentation in scene understanding," *arXiv preprint arXiv:1809.02492*, 2018.
- [15] P. Soille, *Morphological image analysis: principles and applications*. Springer Science & Business Media, 2013.
- [16] D. Pearson, "Image processing (essex series in telecommunication and information systems)," 1991.
- [17] R. C. Gonzalez, "Deep convolutional neural networks," *IEEE Signal Processing Magazine*, vol. 35, no. 6, pp. 79–87, 2018.
- [18] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 779–788.