

Visualização Panorâmica Georreferenciada de Potenciais Focos de Mosquito *Aedes aegypti*

Thadeu L. B. Dias¹, Lucas S. Meyer², Wesley L. Passos³, Eduardo A. B. da Silva^{1,2} e Sergio L. Netto^{1,2}.

Resumo—A detecção e localização de possíveis focos de proliferação de mosquitos podem direcionar ações de combate às doenças relacionadas, aumentando a eficiência no controle e diminuindo a carga sobre as equipes de campo. Neste trabalho é apresentado um sistema de mapeamento georreferenciado de possíveis focos do mosquito *Aedes aegypti*, a partir de reconhecimento aéreo obtido por veículos não tripulados. Utilizando técnicas de combinação de quadros, uma vista panorâmica de alta resolução da região monitorada é sintetizada, sinalizando e georreferenciando pontos de interesse, além de apresentar estatísticas como densidade de focos por região, permitindo uma interpretação intuitiva para direcionamento de equipes.

Palavras-Chave—Costura de imagens, Georreferenciamento.

Abstract—The detection and localization of potential mosquito breeding sites can guide actions to combat related diseases, increasing control efficiency and reducing the burden on field teams. This work presents a georeferenced mapping system for potential breeding sites of the *Aedes aegypti* mosquito, based on aerial recognition obtained through unmanned vehicles. Using frame stitching techniques, a high-resolution panoramic view of the monitored area is synthesized, highlighting and georeferencing points of interest, as well as providing statistics such as breeding site density by region, enabling an intuitive interpretation to guide field teams.

Keywords—Image stitching, georeferencing.

I. INTRODUÇÃO

Dengue, Zika, Chikungunya e febre amarela urbana, arboviroses transmitidas pelo mosquito *Aedes aegypti*, causam impactos significativos na saúde e na economia brasileira. Somente em 2023, 1,6 milhão de casos prováveis de dengue foram registrados no Brasil [1], um aumento de 15,8% em relação ao ano anterior. Estimativas dos custos relacionados à epidemia de dengue são superiores a 1,4 bilhão de dólares anuais [2].

Entre as estratégias mais eficazes para mitigar as infecções transmitidas pelo *Aedes aegypti*, destaca-se a eliminação de focos de proliferação destes mosquitos. No Brasil, essa abordagem é tradicionalmente implementada por meio de visitas domiciliares realizadas por agentes de saúde e de controle de endemias, que buscam potenciais criadouros. Esses locais são caracterizados pela presença de água limpa, parada e sombreada, especialmente em ambientes urbanos, onde objetos comuns como vasos de plantas, entulhos, pneus descartados e caixas d'água abertas frequentemente tornam-se habitats ideais para reprodução. Considerando a abundância e dispersão desses possíveis focos, a vigilância precisa ser abrangente, o que implica em maiores custos operacionais, recursos humanos

significativos e elevada carga de trabalho para os agentes de saúde responsáveis pelo monitoramento.

Para ampliar e tornar mais eficiente o monitoramento de possíveis focos do *Aedes aegypti*, o estudo descrito em [3] introduz um conjunto de dados constituído por imagens aéreas obtidas por Veículos Aéreos Não Tripulados (VANTs), popularmente conhecidos como *drones*. Essas imagens são capturadas sobre regiões reconhecidamente afetadas pela presença do mosquito, permitindo que possíveis focos sejam anotados e, posteriormente, identificados automaticamente utilizando métodos de aprendizado de máquina e visão computacional. Os VANTs possuem a capacidade de realizar varreduras automáticas em áreas específicas, integrando o georreferenciamento via satélite às suas operações. Associando essa funcionalidade à detecção automatizada, torna-se possível realizar avaliações rápidas e precisas sobre quais locais devem ser priorizados para visita, facilitando a identificação de regiões críticas e, conseqüentemente, reduzindo a carga operacional das equipes de campo envolvidas nas ações de controle vetorial.

Embora a detecção automática de potenciais focos do *Aedes aegypti* utilizando VANTs seja uma tarefa já amplamente consolidada na literatura [3]–[9], um sistema de monitoramento prático deve também apresentar claramente ao operador as informações obtidas durante uma varredura em campo. Neste contexto, o presente trabalho propõe um método para visualização panorâmica e georreferenciada dos possíveis focos do mosquito *Aedes aegypti* utilizando detecções obtidas por meio da análise de imagens aéreas capturadas por VANTs.

A partir de um vídeo capturado por uma varredura aérea, o método proposto visa fornecer aos operadores um mapa panorâmico e de alta resolução da região monitorada, integrando informações georreferenciadas sobre cada detecção a serviços de mapas como o Google Maps. Além disso, o sistema oferece estatísticas complementares, tais como mapas de calor, permitindo a pronta identificação e visualização das áreas críticas com maior incidência de potenciais focos do mosquito *Aedes aegypti*.

Para melhor apresentar as contribuições, o restante do presente trabalho está organizado da seguinte maneira: a Seção II apresenta trabalhos relacionados à localização de focos de mosquitos por varreduras de VANTs, introduzindo o conjunto de dados que é o objeto deste trabalho; já a Seção III apresenta o método proposto para a combinação de quadros e síntese de vista panorâmica; finalmente, a Seção IV sumariza as contribuições deste trabalho e denota possíveis direções para pesquisas futuras.

II. TRABALHOS RELACIONADOS

A utilização de imagens aéreas obtidas por VANTs e técnicas de visão computacional vêm ganhando espaço na

¹PEE/COPPE, UFRJ, Rio de Janeiro-RJ. ²DEL/Poli, UFRJ, Rio de Janeiro-RJ. ³CEFET-RJ, Nova Iguaçu-RJ. emails: {thadeu.dias, lucas.meyer, eduardo, sergioln}@smt.ufrj.br, wesley.passos@cefet-rj.br. O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

literatura científica como soluções eficazes para a identificação e o monitoramento automatizado de focos do mosquito *Aedes aegypti* [3]–[9].

Diversos estudos têm investigado o uso de VANTs para monitoramento ambiental e controle de vetores. Por exemplo, o estudo em [9] utiliza imagens aéreas capturadas por VANTs e aprendizado profundo para detecção de caixas d’água e piscinas, objetos conhecidos como potenciais criadouros em áreas urbanas. Da mesma forma, o trabalho em [8] aborda a detecção de possíveis criadouros de mosquitos usando um VANT, visando caixas d’água e áreas com resíduos inorgânicos. Os autores de [7] exploram o uso de imagens de VANTs e aprendizado profundo para identificar possíveis locais de reprodução do *Aedes albopictus*, detectando nove classes de objetos associados à reprodução desses mosquitos, incluindo baldes, tigelas, lixeiras, plantas aquáticas, potes, tampas, vasos, caixas e pias.

Naturalmente, para que a detecção automática de potenciais focos de proliferação do mosquito seja eficaz, é necessário contar com detectores precisos para os objetos específicos de interesse. Em geral, esses métodos requerem grandes volumes de dados anotados para treinamento, a fim de garantir um bom desempenho. Para atender a essa demanda, o trabalho em [3] propõe um conjunto de dados anotado, denominado *Mosquito Breeding Grounds* (MBG), que inclui vídeos capturados por VANTs que retratam possíveis criadouros de mosquitos. Além disso, ele fornece resultados para a detecção de pneus e caixas d’água usando aprendizado profundo. O estudo apresentado em [4] amplia o trabalho em [3] aplicando técnicas de aumento de dados ao conjunto de dados MBG para melhorar o volume de dados disponíveis e o desempenho da detecção. Vale mencionar que o conjunto de dados MBG constitui o fundamento para o desenvolvimento e validação do trabalho proposto neste artigo.

Apesar dos avanços nos estudos mencionados, ainda há uma carência de soluções integradas que combinem robustez na detecção automática, visualização panorâmica intuitiva e georreferenciamento detalhado especificamente voltados para o controle vetorial do *Aedes aegypti*. O presente trabalho busca preencher essa lacuna, apresentando uma abordagem integrada que facilita a identificação rápida de áreas críticas, a fim de reduzir a carga operacional sobre as equipes responsáveis pelo controle vetorial.

III. MÉTODO PROPOSTO

A. Métodos “transitivos”

Dado que os vídeos utilizados neste trabalho são obtidos por VANTs durante varreduras aéreas com rotas pré-programadas, em altitudes fixas e com o plano da câmera predominantemente paralelo ao solo [3], é razoável assumir que, excetuando-se pequenos efeitos de paralaxe, transformações simples baseadas no modelo de movimento euclidiano são suficientes para realizar a composição das imagens.

Considerando que o movimento entre pares consecutivos de quadros em um vídeo é relativamente pequeno – resultando em alta sobreposição entre os quadros – a correlação de fase se apresenta como um método eficiente para estimar o deslocamento em pixels entre as imagens. A aplicação da correlação de fase em cada par de quadros consecutivos, (I_i, I_{i+1}) fornece um vetor de translação $\Delta \mathbf{x}_{i,i+1}$, que representa o movimento relativo entre as imagens. Assumindo



Fig. 1: Costura de quadros via correlação de fase. Apesar da aparente continuidade entre imagens, o modelo de translação não acomoda rotações, criando falsas disparidades. Eventualmente, a acumulação de erros leva a falhas visíveis entre passadas. A faixa preta na região superior da imagem revela um desvio nas estimativas de posição.

a transitividade desse deslocamento ao longo da sequência, é possível computar o vetor de translação entre quaisquer pares de quadros i e $i + N$ na sequência, como

$$\Delta \mathbf{x}_{i,i+N} = \sum_{j=i}^{N-1} \Delta \mathbf{x}_{j,j+1}, \quad (1)$$

tornando possível a combinação incremental de múltiplas imagens em um mosaico panorâmico.

A aplicação do método da correlação de fase em uma das cenas do conjunto de dados está ilustrada na Fig. 1. Apesar da correlação de fase ser capaz de estimar movimentos pequenos entre pares de quadros, a costura insatisfatória via transitividade revela uma deficiência dessa abordagem: o movimento do VANT não é perfeitamente paralelo ao solo. Durante o percurso, pequenas variações na altitude e na orientação da câmera causam movimentos aparentes entre quadros que não são bem modelados apenas por translações, comprometendo a precisão do mosaico gerado.

Considerando as insuficiências observadas no uso da correlação de fase, uma alternativa mais robusta seria permitir transformações projetivas mais gerais entre as imagens. Uma estratégia comumente aplicada consiste em encontrar uma homografia entre duas imagens sobrepostas I_i e I_{i+1} . Usando detectores/descriptores de pontos salientes, como o SIFT, é possível identificar pontos do espaço visíveis simultaneamente nas duas imagens e, partir de suas correspondências, estimar a matriz de homografia $H_{i,i+1}$, que mapeia as coordenadas dos pontos identificados da imagem I_{i+1} para I_i , conforme ilustrado na Fig. 2. Considerando também a transitividade dessas transformações, é possível compor a homografia entre imagens não adjacentes I_i e I_{i+N} , a partir do produto das homografias individuais

$$H_{i,i+N} = \prod_{j=i}^{N-1} H_{j,j+1}, \quad (2)$$

o que permite posicionar múltiplas imagens no mosaico.

A Fig. 3 ilustra o método baseado na composição de homografias ao longo de uma sequência de vistas. Embora a transição entre as imagens seja suave – sem cortes ou descontinuidades perceptíveis – observa-se que, assim como no método da correlação de fase, distorções geométricas tendem a se acumular à medida que o mosaico se estende. Essas distorções comprometem a planicidade do panorama, dificultando a composição de mosaicos com trechos mais longos. Isso ocorre porque métodos transitivos se baseiam em

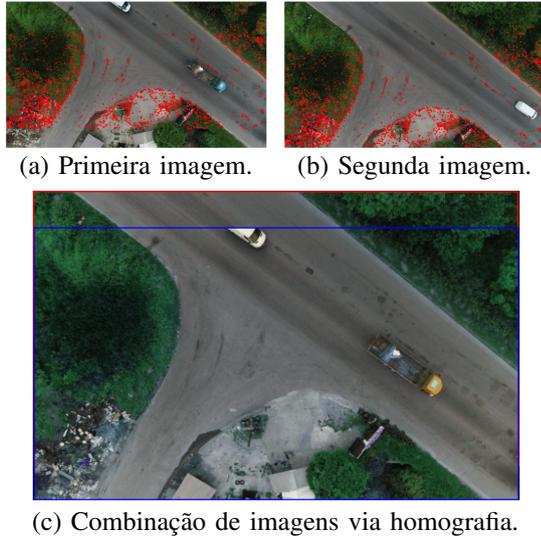


Fig. 2: Exemplo de combinação de imagens via homografia. Os pontos vermelhos denotados nas figuras originais foram detectados e descritos usando SIFT. Uma homografia é computada para o alinhamento, movendo a primeira imagem (borda azul) sobre a segunda (borda vermelha).

estimativas indiretas para posicionar cada quadro em relação a um quadro de referência inicial, mesmo que haja sobreposição entre eles. Essa abordagem leva a uma dependência estrutural em torno de um quadro preferencial e, inevitavelmente, à propagação de erros acumulativos ao longo da costura.



Fig. 3: Trecho do vídeo combinado através da composição de homografias. Apesar da costura suave, distorções tendem a se acumular, impedindo a combinação de intervalos mais longos.

B. Método da vista sintética

Enquanto os métodos baseados em transitividade falham em fundir uma quantidade significativa de quadros, ferramentas de Structure-from-Motion (SfM) [10] podem ser úteis nessa tarefa. SfM é uma técnica capaz de extrair informação 3D de uma cena (*structure*) a partir de uma série de imagens obtidas de posições variadas (*motion*). Embora o objetivo primário do SfM é a extração de informação da cena (modelo 3D de objetos, por exemplo), a estimação da posição de pontos 3D em cenas requer, simultaneamente, a estimação da posição e orientação das vistas utilizadas na reconstrução. Uma das ferramentas de SfM mais populares é a ferramenta de código aberta COLMAP [11]–[13], capaz de realizar as tarefas relacionadas à estimação de múltiplas vistas, como calibração, triangulação e *bundle adjustment*. A Fig. 4 demonstra o resultado de uma localização obtida com a ferramenta COLMAP, denotando posições estimadas das câmeras durante a varredura



Fig. 4: Vista lateral de uma localização obtida com a ferramenta COLMAP. A estrutura da cena pode ser visualizada com a nuvem de pontos, na parte inferior da imagem, enquanto as pirâmides vermelhas no topo da imagem denotam a posição e orientação das vistas.

aérea, além de uma nuvem de pontos 3D representando a estrutura extraída da cena.

O método proposto para a síntese da vista panorâmica pode ser descrito da seguinte forma: Dada uma triangulação da área observada, através do SfM, contendo estimativas K dos parâmetros intrínsecos da câmera utilizada, as localização e orientação C_i das vistas, e a posição $\mathcal{X} = \{\mathbf{x}_i\}$ de um conjunto esparsos de pontos na cena, os seguintes passos são realizados:

- 1) Obtém-se uma estimativa da posição central da cena, e se posiciona uma nova câmera, C_P , de forma que o campo de visão de C_P englobe a área observada completamente, eventualmente sendo associada a imagem panorâmica I_P . Uma forma simples de obter uma vista que englobe toda a região é posicionar a câmera C_P em uma altitude mais elevada relativa às imagens originais.
- 2) Para cada vista original I_i , os pontos triangulados e visíveis em I_i são projetados via C_P , resultando em estimativas para as coordenadas dos pontos triangulados em I_M .
- 3) Como a associação entre os pontos já triangulados em I_i e suas projeções em I_P é automática, isso é o suficiente para se obter uma estimativa de homografia H_{iP} , posicionando uma vista I_i na imagem panorâmica I_P , de forma similar ao método transitivo anterior.

Exemplos de posicionamento de C_P , assim como a localização das vistas em I_P podem ser vistos nas Figs. 5 e 6. Como as transformações aplicadas nas vistas originais são homografias gerais, o método é capaz de lidar com pequenas variações de altura e orientação na plataforma sem que os efeitos de distorção sejam significativos. Além disso, com as vistas mutuamente localizadas, num mesmo quadro de referência, apenas uma transformação é aplicada por imagem, evitando a acumulação de erros observada nos métodos anteriores.

Além de fornecer as informações suficientes para o posicionamento das vistas na imagem panorâmica, uma estimativa da resolução da imagem final também pode ser extraída. No método proposto, utilizando os pontos triangulados na vista posicionada mais próxima ao centro da região observada, é possível obter uma estimativa da razão entre a área da vista e a resolução original. Extrapolando a razão para a área completa da gravação, a resolução adequada para a vista panorâmica pode ser calculada, e os quadros posicionados, formando o mosaico panorâmico como na Fig. 7.

C. Georreferenciamento

Durante a gravação das imagens, a posição do VANT ao longo de sua trajetória é registrada, via sistema de navegação por satélite (GNSS). Apesar da localização não ser atualizada



Fig. 5: Vista lateral da cena, com o posicionamento da câmera panorâmica. As pirâmides menores denotam as posições das vistas originais, assim como na Fig. 4, e seus campos de visão. A maior pirâmide denota a localização de C_P , com seu campo de visão englobando toda a área observada.



Fig. 6: Posicionamento de uma seção de vistas sobre I_M , com as bordas de cada imagem delimitadas. É possível notar diferentes acomodações para cada vista original, preservando a continuidade da imagem global apesar dos movimentos da plataforma.

na mesma taxa dos quadros filmados, como na Fig. 8, é possível estimar a localização das vistas intermediárias através de interpolações. Como as características da câmera, assim como a altitude do VANT são conhecidas, podemos estimar a razão métrica entre distância aparente na imagem, isso é, em pixels, e a distância física de pontos em uma imagem através de

$$e = \frac{2H \operatorname{tg}\left(\frac{\operatorname{FoV}}{2}\right)}{\sqrt{h_i^2 + w_i^2}}, \quad (3)$$

onde H é a altitude do VANT em metros, h_i e w_i são as dimensões da imagem, em pixels, e FoV é o campo de visão (diagonal) da câmera.

Além disso, relacionando a estimativa de posição das câmeras (orientadas com as imagens) obtidas pelo SfM com o posicionamento reportado pelo sistema de navegação do VANT, os versores de orientação podem ser calculados através de um simples sistema linear,

$$\begin{bmatrix} \Delta x & \Delta y \\ -\Delta x & \Delta y \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \Delta\phi \\ \Delta\lambda \end{bmatrix}, \quad (4)$$

onde Δx e Δy são deslocamentos em pixels entre os centros das imagens, e $\Delta\phi$ e $\Delta\lambda$ são os deslocamentos medidos no sistema de coordenadas geográficas entre vistas (latitude e longitude, respectivamente).

Finalmente, os vetores de orientação de uma imagem podem

ser obtidos, sendo

$$\partial_\phi = \frac{1}{\sqrt{a^2 + b^2}} \begin{bmatrix} a \\ b \end{bmatrix} \quad \text{e} \quad \partial_\lambda = \frac{1}{\sqrt{a^2 + b^2}} \begin{bmatrix} b \\ -a \end{bmatrix} \quad (5)$$

os versores de latitude e longitude, respectivamente. A Fig. 9 demonstra a sobreposição de um quadro localizado e orientado sobre uma captura de satélite.

Finalmente, aplicando a estimação de deslocamento, Eq. (3), sobre as detecções de possíveis focos em cada vista, a localização das detecções pode ser também posicionada em um mapa, como na Fig. 10, e o mapa de calor das regiões mais críticas pode ser visualizado, como na Fig. 11.

IV. CONCLUSÕES

Neste artigo, apresentamos métodos para a visualização panorâmica e georreferenciada de detecções de possíveis focos de proliferação do mosquito *Aedes aegypti*. Utilizando técnicas de visão computacional, as imagens obtidas por VANTs durante inspeções aéreas são combinadas em mosaicos panorâmicos de alta resolução, permitindo uma representação visual detalhada das regiões monitoradas. A abordagem proposta integra ferramentas de SfM para reconstrução da trajetória e orientação das vistas.

Adicionalmente, incorporamos mecanismos de georreferenciamento, possibilitando o mapeamento dos potenciais focos em plataformas como Google Maps, além da geração de mapas de calor que destacam áreas críticas. A integração entre visualização, detecção automática e geolocalização fornece subsídios para o planejamento e execução de ações de combate ao mosquito em campo.

Como trabalho futuro, propomos a expansão do sistema para suportar análises temporais com diferentes campanhas de voo, a incorporação de algoritmos de detecção mais robustos treinados com dados multirresolução, e a implementação de uma interface interativa que permita a navegação e anotação colaborativa das regiões identificadas como prioritárias.

REFERÊNCIAS

- [1] Ministério da Saúde, “Com previsão de aumento de casos, ministério da saúde coordena ações de enfrentamento das arboviroses,” <https://www.gov.br/saude/pt-br/assuntos/noticias/2023/dezembro/com-previsao-de-aumento-de-casos-ministerio-coordena-acoes-de-enfrentamento-das-arboviroses>, [Acessado em 30/04/2025].
- [2] A. Laserna, J. Barahona-Correa, L. Baquero, C. Castañeda-Cardona, and D. Rosselli, “Economic impact of dengue fever in latin america and the caribbean: a systematic review,” *Revista Panamericana de Salud Pública*, vol. 42, p. e111, 2018.
- [3] W. L. Passos, G. M. Araujo, A. A. de Lima, S. L. Netto, and E. A. da Silva, “Automatic detection of aedes aegypti breeding grounds based on deep networks with spatio-temporal consistency,” *Computers, Environment and Urban Systems*, vol. 93, p. 101754, 2022.
- [4] W. L. Passos, C. S. Barreto, G. M. Araujo, U. Haque, S. L. Netto, and E. A. B. da Silva, “Toward improved surveillance of *Aedes aegypti* breeding grounds through artificially augmented data,” *Engineering Applications of Artificial Intelligence*, vol. 123, pp. 1–12, 2023.
- [5] W. L. Passos, G. M. Araujo, U. Haque, F. Cruz-Roldan, and S. L. Netto, “IEEE ICIP 2023 challenge on the automatic detection of mosquito breeding grounds,” in *IEEE International Conference on Image Processing Challenges and Workshops*, Kuala Lumpur, Malaysia, 2023, pp. 1–5.
- [6] C. Laranjeira, D. Andrade, and J. A. dos Santos, “Yolov7 for mosquito breeding grounds detection and tracking,” in *IEEE International Conference on Image Processing Challenges and Workshops*, 2023, pp. 1–5.
- [7] K. Yu, J. Wu, M. Wang, Y. Cai, M. Zhu, S. Yao, and Y. Zhou, “Using UAV images and deep learning in investigating potential breeding sites of *aedes albopictus*,” *Acta Tropica*, vol. 255, pp. 1–12, 2024.



Fig. 7: Mosaico panorâmico de uma das regiões do dataset, obtido pelo método da vista sintética. Formado pela combinação de 244 quadros, e com resolução nativa de 40912×9978 px., essa vista permite uma observação ampla e detalhada a região sobrevoada.



Fig. 8: Exemplo de trajetória do VANT ao longo de uma observação. Os pontos assinalados no mapa correspondem às posições reportadas pelo VANT.



Fig. 10: Localização de possíveis focos detectados. As marcações no mapa podem ser usadas para identificar e traçar rotas para os endereços que devem ser vistoriados pelos agentes.



Fig. 9: Quadro do conjunto de dados (borda vermelha), sobreposto a imagem de satélite da mesma região. A conexão uniforme entre as áreas de asfalto do VANT e a imagem de satélite evidenciam o bom alinhamento e orientação do quadro sobreposto, apesar do ângulo não paralelo ao chão na imagem de satélite. A diferença na resolução e riqueza de detalhes da filmagem com o VANT também pode ser observada.



Fig. 11: Mapas de calor apontando a densidade de pontos de interesse localizados na região. Esse tipo de visualização permite uma análise intuitiva de pontos críticos, e quais áreas requerem maior atenção.

[8] D. T. Bravo, G. A. Lima, W. A. L. Alves, V. P. Colombo, L. Djogbénou, S. V. D. Pamboukian, C. C. Quaresma, and S. A. de Araujo, "Automatic detection of potential mosquito breeding sites from aerial

images acquired by unmanned aerial vehicles," *Computers, Environment and Urban Systems*, vol. 90, no. 9, pp. 1–13, 2021.

- [9] H. S. Cunha, B. S. Schlauser, P. F. Wildemberg, E. A. M. Fernandes, J. A. dos Santos, M. de Oliveira Lage, C. Lorenz, G. L. Barbosa, J. A. Quintanilha, and F. Chiaravalloti-Neto, "Water tank and swimming pool detection based on remote sensing and deep learning: Relationship with socioeconomic level and applications in dengue control," *PLOS One*, vol. 16, no. 12, pp. 1–24, 2021.
- [10] R. Szeliski, *Structure from motion and SLAM*. Cham: Springer International Publishing, 2022, pp. 543–594.
- [11] J. L. Schönberger and J.-M. Frahm, "Structure-from-motion revisited," in *Conference on Computer Vision and Pattern Recognition*, 2016.
- [12] J. L. Schönberger, E. Zheng, M. Pollefeys, and J.-M. Frahm, "Pixel-wise view selection for unstructured multi-view stereo," in *European Conference on Computer Vision*, 2016.
- [13] "COLMAP - Structure-from-Motion and Multi-View Stereo — github.com," <https://github.com/colmap/colmap>, [Acessado em 30/04/2025].