

Assessment of public spaces through autonomous processes: digital mapping and remote sensing of tourist areas.

Carlos Eduardo Verzola Vaz¹, Pedro Oscar Pizzetti Mariano², Ana Beatriz Caetano³, Isabela Tarczewski⁴, Suelen Joaquin⁵

¹ Universidade Federal de Santa Catarina, Florianópolis, Brasil, cevv00@gmail.com

² Universidade Federal de Santa Catarina, Florianópolis, Brasil,
pedro.pm@hotmail.com

³ Universidade Federal de Santa Catarina, Florianópolis, Brasil,
anabeatriz.prado@gmail.com

⁴ Universidade Federal de Santa Catarina, Florianópolis, Brasil,
bela.constantini.t@gmail.com

⁵ Universidade Federal de Santa Catarina, Florianópolis, Brasil,
suh.zanettaj@gmail.com

Abstract. This research aims to explore the use of Unmanned Aerial Vehicles (UAV) and Computer Vision as tools to collect spatial data (qualitative and quantitative) and the dynamics of use of public spaces and generate information for public managers. During the research, drones will be used to capture videos and images of a case study selected with the help of a team of project partner managers. The collected data will be translated into information for accurate and efficient decision-making, based on the use of knowledge extraction and data mining strategies. The extracted knowledge will be presented to managers as a tool prototype that helps in the design and management of public spaces present in areas with great tourist influence. The adoption of new technologies can strengthen the means of application and revitalize the global partnership for the development of innovative built environments, thus impacting on the UN's sustainable development goals.

Keywords: Public Spaces; Leisure; Technology; Computer Vision; Sustainable Development.

1 Introdução

Krzyzanowski e Cruz (2011) afirmam que a sazonalidade é um fenômeno muito comum e frequente em milhares de destinos turísticos, no qual ocorrem dezenas de problemas e consequências socioambientais e econômicas tanto para turistas como para a comunidade receptora. As atividades ligadas ao

turismo, de Santa Catarina, correspondem a 12.5% do Produto Interno do Estado. A região litorânea concentra grande parte destas atividades, principalmente devido às características atrativas das praias. Durante o verão de 2018, em três meses Santa Catarina recebeu cerca de 2,5 milhões de turistas, sendo que grande parte destes se concentraram na região litorânea (Embratur, 2018). Este grande movimento de pessoas altera a vida das cidades litorâneas. Neste período ficam evidentes, ano a ano, problemas de mobilidade, saneamento, segurança e demanda por infraestrutura que dê suporte às atividades de lazer, não apenas de turistas, mas dos moradores locais. Contudo, para adequar os espaços públicos litorâneos de lazer é fundamental que estes sejam avaliados de modo a gerar informações adequadas para sua gestão.

Dois componentes devem ser considerados em um processo de avaliação, o primeiro deles relaciona-se às características espaciais, tais como sua configuração, equipamentos existentes, tipo de pavimentação, etc. O segundo está associado à forma com que se dá a ocupação do espaço, intensidade e frequência de uso, e os tipos de atividades realizadas pelos usuários. De acordo com Gehl e Svarre (2013), existem vários tipos de medidas utilizadas na observação de humanos no espaço público, tais como:

- Contagem: é geralmente realizada durante um longo período, bem como horários do dia ou da semana;
- Mapeamento: que serve para coletar as localizações de um sujeito em um espaço delimitado;
- Rotas: sequência de caminhada, escolha de direção, fluxo, entradas mais e menos usadas, etc.
- Rastreamento: envolve seguir pessoas e medir velocidade média de pedestres, tempo necessário para percorrer uma determinada distância e duração das atividades.

Entre as ferramentas utilizadas para a coleta de dados, Gehl e Svarre (2013) explicam que as fotografias e vídeos podem ser amplamente utilizadas no campo dos estudos da vida pública para ilustrar a interação entre a forma urbana e a dinâmica humana. Uma das possibilidades de uso de vídeo pode ser identificada na pesquisa de Whyte (1980) cujo objetivo era compreender a vida social em ambientes urbanos. Em sua pesquisa, o autor usou a observação e a captura de imagens para coletar dados sobre o fluxo e o comportamento das pessoas nos espaços públicos a partir de câmeras montadas com vista para praças em Nova York. Os estudos desenvolvidos na área urbana por Whyte (1980) e Gehl e Svarre (2013) foram baseados no pós-processamento manual dos dados coletados. No entanto, nas últimas duas décadas, a visão computacional tem se tornado um campo muito mais acessível, com bibliotecas gratuitas e ferramentas desenvolvidas para processar e analisar imagens automaticamente, permitindo a criação de aplicativos capazes de influenciar, gerenciar, administrar diretamente os cidadãos.

Este artigo apresenta uma pesquisa em andamento que visa explorar o uso de tecnologias como veículos aéreos não tripulados (VANT) e visão computacional para avaliar espaços públicos turísticos em Santa Catarina — Brasil. O objetivo principal é gerar informações precisas e relevantes para a gestão eficiente dessas áreas, além de criar um ferramental que auxilie no planejamento estratégico.

2 Aplicação da visão computacional

Moeslund e Granund (2001) afirmam que as aplicações para a captura de movimento humano pertencem ao campo do desenvolvimento de sistemas, abrangendo três áreas principais: vigilância, controle e análise. Os autores explicam que a área de vigilância engloba aplicativos nos quais um ou mais indivíduos podem ser rastreados ao longo do tempo e monitorados quanto a ações específicas. Um exemplo é a vigilância de estacionamentos, em que um sistema rastreia indivíduos para avaliar a possibilidade de comportamento suspeito.

Na área de controle, as aplicações têm o propósito de capturar movimentos e gerar funcionalidades, sendo frequentemente empregadas como interfaces em jogos, ambientes virtuais, animações ou para controle remoto de objetos. A terceira aplicação relaciona-se com a captura de dados e a análise detalhada de movimentos. Essa abordagem pode ser utilizada, conforme os autores, em estudos clínicos para diagnóstico de pacientes ortopédicos, ou como ferramenta para auxiliar atletas a compreender e aprimorar seu desempenho.

No âmbito dos estudos de desenho urbano, observa-se um aumento nas implementações que empregam visão computacional como estratégia. Zhan et al. (2008) destacam que o contínuo crescimento populacional e a rápida urbanização global tornam as situações de multidão mais frequentes nas cidades. Segundo os autores, a análise da dinâmica de multidões tem sido cada vez mais explorada, resultando em diversas aplicações com foco nos seguintes temas:

- Gerenciamento de multidão: Desenvolver estratégias para monitorar multidões, visando prevenir incidentes e garantir a segurança em eventos esportivos, por exemplo.
- Projeto de espaço público: A análise da dinâmica de uso do espaço pode fornecer dados para diretrizes de design de áreas públicas.
- Ambientes virtuais: Modelos matemáticos de multidões podem ser aplicados em ambientes virtuais para simular comportamentos humanos, enriquecendo a experiência do usuário.
- Vigilância: A análise de multidões pode ser empregada para detecção automática de anomalias e criação de alarmes. Além disso, rastrear indivíduos em grupos pode ser útil no reconhecimento de suspeitos.

- Ambientes inteligentes: Em ambientes que envolvem grandes grupos, a análise é essencial para ajudar multidões ou indivíduos em suas decisões.

3 Metodologia

A pesquisa se desenvolveu em três etapas diferentes (figura 1) que, em conjunto, resumem criar um framework de trabalho para a aplicação de ferramentas de sensoriamento na avaliação de espaços públicos. As etapas aqui apresentadas são concebidas para implementação de forma cíclica, ou seja, as informações e resultados gerados serão retroalimentados nas etapas anteriores como método de calibração das técnicas e instrumentos para um novo ciclo de aplicação. A estratégia pode corrigir inconsistências encontradas ao longo do processo.

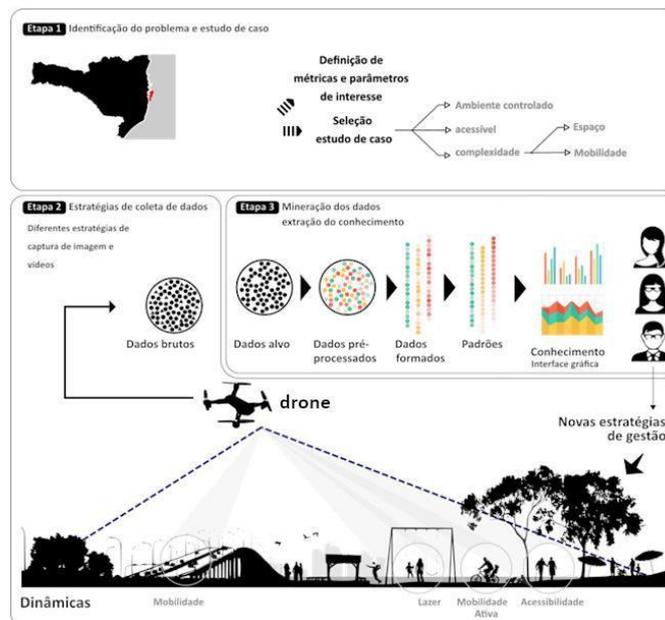


Figura 1. Diagrama geral da pesquisa. Fonte: Os autores (2023).

3.1 Identificação do problema e estudo de caso

Nesta fase, são também estabelecidos os estudos de caso, inicialmente compreendendo uma macrozona do litoral de Santa Catarina, onde as atividades e as dinâmicas espaciais variam ao longo do ano. As áreas praianas, que estão sujeitas ao turismo sazonal, possibilitam a análise da

influência das interações humanas no espaço construído e natural, gerando comparações entre momentos de presença e ausência do aumento populacional.

A coleta de dados em si foi realizada em pontos-chave da costa, onde ocorre a convergência de atividades humanas, usos do solo e modais de transporte. Os critérios utilizados para a seleção desses locais incluem:

- Ser um ambiente controlado, capaz de fornecer informações por meio de métodos alternativos para validar os instrumentos desta pesquisa, caso seja necessário;
- Ter fácil acesso para permitir que a equipe conduza suas atividades de forma contínua e sem restrições;
- Apresentar complexidade em relação à presença de diversos tipos de edificações, espaços públicos e infraestrutura de apoio a atividades de lazer em espaços abertos, além do fluxo e da intensidade dos diferentes modais de transporte motorizado e não motorizado.

3.2 Estratégia de coleta de dados

A investigação se baseia no uso da captura de imagens com o VANT, associado a ferramentas para treinamento de modelos para identificação de objetos. A partir desta estratégia, são levantadas características capazes de indicar tendências do comportamento humano e sua relação com os sistemas públicos estabelecidos e estrutura urbana, tais como:

- Mobilidade: contagem de veículos, modais de transporte, rotas (veiculares, pedonais e ciclísticas), transporte público;
- Usuários: contagem de pessoas, registro de atividades, agrupamentos (com manutenção da privacidade e anonimato de faces e outros elementos identificadores com recursos das bibliotecas utilizadas, como borrões e desfoques);

O uso de VANTs, inteligência artificial e automação de processos oferece vantagens em relação a processos tradicionais e manuais de coleta. Destacam-se como possíveis melhorias processuais com a adoção dos instrumentos desta etapa: melhor emprego do tempo atualmente despendido para treinar agentes, coletar, tabular e processar dados manualmente; a economia e redistribuição de investimentos recebidos para manter o grande corpo de agentes temporários para os processos citados anteriormente; o possível aumento do ciclo de avaliação, considerando que os altos custos e o limite do número de agentes são impeditivos ao processo contínuo; a diminuição do viés humano sobre a fonte dos dados, que pode impactar negativamente sua confiabilidade; e a capacidade humana limitada em relação ao alcance perceptivo e visual em observar padrões em grandes áreas.

3.3 Estratégia de coleta de dados

A terceira etapa corresponde à exploração de instrumentos digitais para viabilizar processos de transformação do dado bruto em informações e conhecimento acerca das dinâmicas e características do espaço avaliado, que sirvam de aporte para a criação de recomendações e ações para a gestão urbana.

A informação produzida e disponibilizada como produto das etapas anteriores pode ser consumida a depender do objetivo das consultas realizadas à plataforma. Todo o material coletado é consolidado em uma base de dados e se apresenta em diversos formatos e estruturas.

- Dado bruto: corresponde a consultas ao arquivamento de imagens e vídeos coletados in loco pelo VANT (com e sem paralelismo ao tratamento com o módulo de Visão Computacional);
- Relatórios (dinâmica humana): retornam informações de natureza mutável que representam as atividades e interações de usuários (em modalidades de transporte diferentes) com o espaço urbano. Consegue relacionar dados de uso do espaço quanto à sua quantificação, temporalidade, padronização e, potencialmente, causalidade. Estas informações são consolidadas a partir do uso de modelos algorítmicos sobre imagens e vídeos para identificação de número de indivíduos, seus fluxos, agrupamentos e a associação entre objetos, ambientes e a forma humana para inferência de atividades.

4 Descrição do processo de Pesquisa

Inicialmente, a investigação se concentrou em compreender o potencial da utilização do VANT (veículo aéreo não tripulado) como ferramenta de coleta de dados a partir de imagens. Este processo se deu de forma exploratória. Na primeira fase foram coletadas imagens segundo diferentes parâmetros de altitude e qualidade. O objetivo é compreender o nível de definição de diferentes objetos na imagem (carros, ônibus, pessoas, bicicletas, etc.), a área da cobertura (comprimento e largura) e a deformação causada nas extremidades (figura 2). Este processo se deu em diferentes espaços públicos, tais como sistemas viários, praças e áreas de orla.

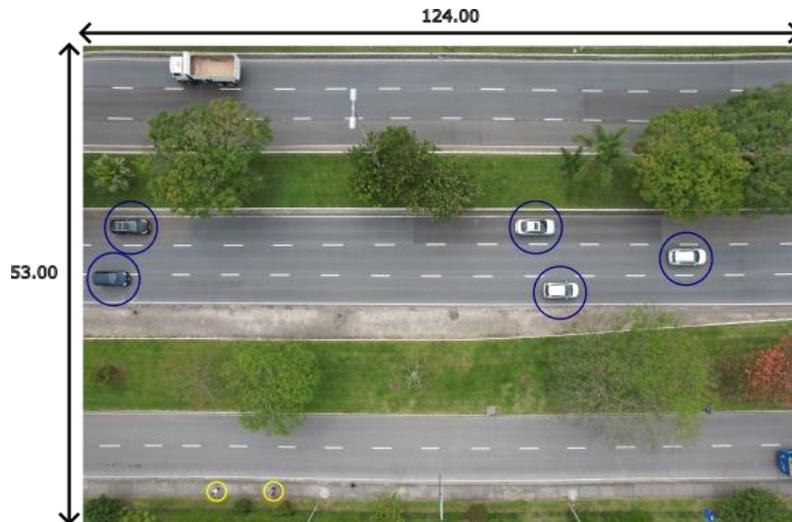


Figura 2. Amostra de imagem de drone a 50 metros de altura. Os objetos circulos em amarelo são pedestres. Fonte: Os autores (2023).

A altitude padrão para uso do VANT foi definida em 50 metros, permitindo a identificação adequada de objetos menores, como ciclistas e pedestres. Na fase subsequente, coletaram-se imagens para explorar o processamento de imagens por técnicas de visão computacional. Essa etapa envolveu a utilização de algoritmos disponíveis, exigindo apenas a formação de um banco de dados para que o modelo "aprendesse" a reconhecer objetos nas imagens. As seguintes plataformas foram empregadas para esse propósito:

- Labellmg: Ferramenta de anotação de imagens, que gera arquivos XML para processamento posterior;
- Roboflow: Plataforma online para processar e melhorar a qualidade de bancos de dados, aprimorando o treinamento de modelos por aprendizado de máquina;
- Treinamento do modelo no COLAB: Algoritmo para treinar o modelo de reconhecimento de objetos usando YOLO 5. O YOLO é uma estrutura de busca em tempo real escrita por Joseph Redmon, integrada à Darknet, uma estrutura de buscas altamente flexível escrita em linguagens de baixo nível.
- Rhinoceros: Software proprietário de modelagem tridimensional baseado em NURBS;
- Grasshopper 3D: linguagem de programação visual executada no Rhinoceros 3D.

Essas ferramentas foram empregadas para identificar e marcar objetos em imagens, enviando suas coordenadas para um ambiente de modelagem paramétrica. Dessa forma, usando um conjunto de imagens coletadas em um

mesmo percurso de mapeamento, é possível identificar a intensidade de uso de espaços públicos e fornecer informações para os gestores por meio de visualização de dados.

O processo de identificação começou com objetos de maior dimensão física, excluindo objetos estáticos como postes, mobiliário urbano e árvores, que estão presentes em mapeamentos SIG e constituem a infraestrutura urbana. Os treinamentos de modelos foram realizados para o reconhecimento de automóveis, pessoas e guarda-sóis (objetos dinâmicos).

O processo de treinamento do modelo envolveu as seguintes etapas:

- Coleta de dados: Voos de VANT a 50 metros de altitude, captando imagens de 5472 x 3648 pixels;
- Anotação das imagens: Utilização do aplicativo LabelImg;
- Processamento do banco de dados: Realizado na plataforma online Roboflow;
- Treinamento do modelo: Por meio de algoritmo para reconhecimento de objetos em YOLO 5.

5 Treinamento e identificação dos objetivos

5.1 Identificação de veículos

No primeiro caso, o levantamento das imagens se deu nas proximidades do campus da Universidade Federal de Santa Catarina Capital. Foram ao todo marcadas 136 imagens e, por meio de técnicas de aumento de dados artificialmente, o conjunto de arquivos final era formado por 240 imagens para treinamento, 24 para validação e 12 para testes. O Resultado obtido para este primeiro modelo permitiu testar as diferentes ferramentas e compreender diferentes parâmetros e estratégias de aumento do banco de dados. A partir do modelo é possível identificar carros com nível de confiança superior a 0,9 (em um domínio numérico de 0 a 1). As figuras (3 e 4) abaixo apresentam parcialmente os resultados obtidos.

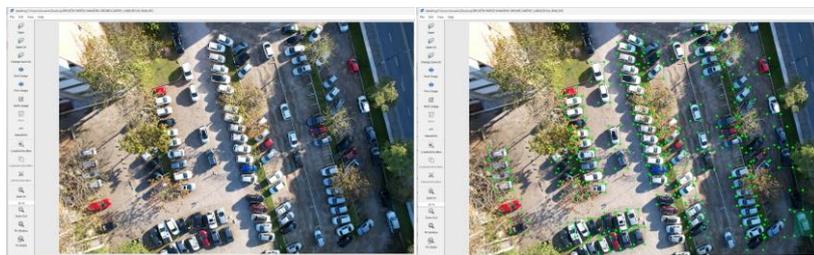


Figura 3. Imagem à esquerda sem anotações e a direita com anotações de carros.
Fonte: Os autores (2023).

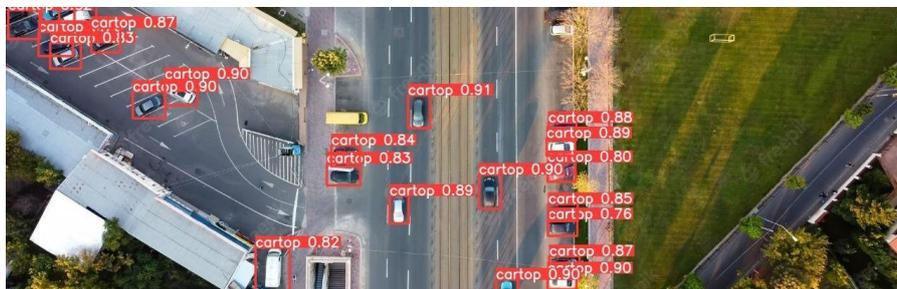


Figura 4. Imagem genérica para teste do modelo para identificar o nível de acerto do modelo treinado para veículos.Fonte: Os autores (2023).

5.2 Identificação de pessoas

Para identificar pessoas, coletamos imagens em diferentes espaços públicos com o máximo de pedestres possível, em variados cenários. Isso resultou em um amplo banco de dados para reforçar a confiabilidade do modelo. Fotografias foram capturadas no Campus da Universidade Federal de Santa Catarina — Capital, imediações, Avenida Beira Mar e praias ao norte de Florianópolis. A coleta ocorreu em etapas, totalizando 195 imagens marcadas (figura 5 e 6). O número de indivíduos identificados por imagem foi significativamente maior que no caso anterior (veículos).

Ampliamos o banco de dados para incluir 3700 imagens de treinamento, 848 de validação e 160 de teste. Esta expansão foi feita durante o pré-processamento para treinar o modelo em objetos de tamanho reduzido na imagem. O modelo treinado apresenta um nível de confiança inferior ao de veículos, contudo em etapas futuras será discutida a real necessidade de se contabilizar exatamente o número de usuários. Esta questão está relacionada com o tipo de informação que o gestor deve ter acesso e a precisão necessária.

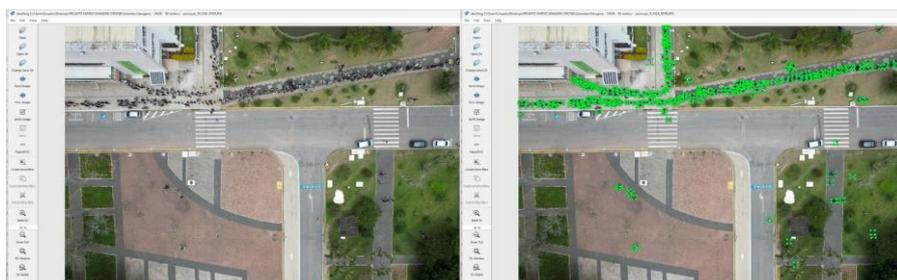


Figura 5. Imagem à esquerda sem anotações e a direita com anotações de pessoas (aplicativo Labellmg).Fonte: Os autores (2023).



Figura 6. Imagens obtidas na Praia dos Ingleses, norte da ilha de Florianópolis. Em ambos os casos o modelo foi testado para identificar o nível de acerto do modelo treinado em relação à identificação de pessoas. Fonte: Os autores (2023).

5.3 Identificação de guarda-sóis

O último cenário explorado relaciona-se diretamente ao objeto de estudo da pesquisa. Para identificar a ocupação das praias, optou-se por focar em um elemento comum nessas áreas: o guarda-sol. Dada a similaridade nos padrões de forma e cores desse objeto, foi viável construir um banco de dados sem a necessidade de uma grande quantidade de imagens. As imagens foram coletadas em duas praias do norte da ilha de Florianópolis (Praia dos Ingleses e Praia da Lagoinha), totalizando 51 imagens marcadas. O conjunto de dados final para treinar o modelo consistiu em 144 imagens de treinamento, 40 de validação e 20 para testes (figura 7 e 8).

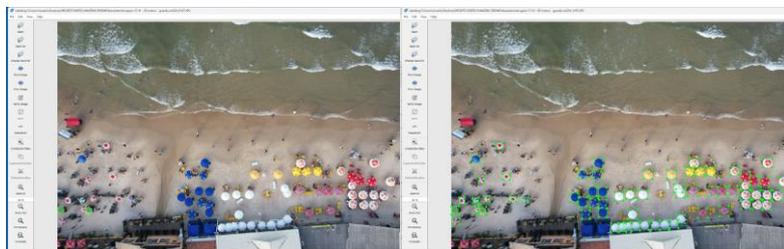


Figura 7. Exemplo de anotação de imagens no aplicativo LabelImg. A esquerda imagem original, a mesma figura anotada. Fonte: Os autores (2023).



Figura 8. Imagem obtida na Praia dos Ingleses, litoral norte da ilha de Florianópolis. Fonte: Os autores (2023).

5.4 Inserção dos resultados em ambiente de modelagem

A partir dos resultados obtidos por meio do treinamento dos modelos de reconhecimento de objetos foi elaborada uma definição no editor de algoritmos visual Grasshopper. Este visa importar os dados de posição dos objetos reconhecidos e sua *bounding box*. A figura abaixo mostra um exemplo de localização de carros identificados do lado esquerdo e, do lado direito, a definição responsável por importar os dados de um arquivo csv (figura 9).



Figura 9: Exemplo de dados importados a partir de dados obtidos por meio de modelo treinado para identificação de carros. Estes foram importados a partir de um arquivo csv por meio do ambiente de modelagem paramétrica Grasshopper (plugin da plataforma de modelagem Rhinoceros 7).: Os autores (2023).

6 Treinamento e identificação dos objetivos

O procedimento executado validou princípios do projeto, mostrando fases de coleta de dados, treinamento do modelo, extração de informações de imagens e pré-visualização de resultados em modelagem. Cada etapa traz desafios a superar em novas fases da pesquisa: Quanto à coleta de dados, é vital selecionar um estudo de caso para obtenção sistemática de imagens, considerando proximidade para visitas frequentes; No treinamento do modelo, é necessário explorar variáveis adicionais para adquirir imagens com objetos em escala reduzida, melhorando a robustez do modelo; Na visualização das informações, é crucial experimentar diferentes abordagens, correlacionando-as com informações georreferenciadas externas.

Em etapas futuras, ocorrerão oficinas com gestores da prefeitura de Florianópolis. Nestas oficinas, serão realizadas inferências sobre os dados obtidos e como as informações podem contribuir para o planejamento de áreas litorâneas com influência do turismo. Durante o processo, serão avaliados parâmetros de densidade, zoneamento, acesso e conexões, diversidade de atividades e mobilidade.

Agradecimentos. Os autores agradecem ao CNPQ pelo apoio à pesquisa a partir do financiamento mediante projeto universal e a FAPESC a partir do

financiamento por meio do Programa de Ciência, Tecnologia e Inovação aos Grupos de Pesquisa da Universidade Federal de Santa Catarina.

Referencias

Embratur, Assessoria de Comunicação da (org.) (2018). Pesquisa da FGV revela o impacto Econômico do Turismo em Santa Catarina. Pires: Inteligência em Turismo.

Gehl, J., e Birgitte S. (2013). How To Study Public Life. Island Press/Center for Resource Economics.

Krzyzanowski, A.; Cruz, G. (2011). Estudo Comparativo da Sazonalidade Turística. In: CRUZ, G. (Org.) Turismo: Desafios e especificações para turismo sustentável. Ilhéus: Editus.

Moeslund, T. B., et al. (2006). "A Survey of Advances in Vision-Based Human Motion Capture and Analysis". *Computer Vision and Image Understanding*, vol. 104, p. 90–126.

Whyte, W. H. (1980). *The social life of small urban spaces*. Washington DC: The Conservation Foundation.

Zhan, B et al.(2008). Crowd analysis: a survey. *Machine Vision And Applications*. Springer Science and Business Media LLC. v. 19, p. 345-357.