

**UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE GEOCIÊNCIAS**

SEGMENT ANYTHING MODEL INTEGRADO À ANÁLISE DE ÁREAS DE RISCOS NO
CÓRREGO DO CABORÉ, ZONA LESTE DO MUNICÍPIO DE SÃO PAULO

VINÍCIUS BRASIL OLIVEIRA
ORIENTADORA: CAMILA DUELIS VIANA

TRABALHO DE FORMATURA
TF-24/53

São Paulo

2024

RESUMO

Este trabalho tem como objetivo a aplicação de técnicas avançadas de inteligência artificial, especificamente o modelo *Segment Anything Model (SAM)*, para a segmentação de imagens aéreas de alta resolução na região do Jardim Caboré, Zona Leste de São Paulo. Diante do aumento alarmante de eventos climáticos extremos e ocorrências naturais na área, a pesquisa busca identificar e analisar os condicionantes de risco geológico e hidrológico, contribuindo para a segurança e resiliência da comunidade local. Inicialmente foi realizado um levantamento da literatura e de documentos sobre o histórico de riscos na região, para compreender os processos e fatores envolvidos. Em seguida, foram identificados os elementos passíveis de segmentação nas imagens de alta resolução, seguidos pela execução da segmentação e avaliação dos resultados utilizando abordagens *zero-shot* e *one-shot*. Além disso, foi executada uma análise integrada dos resultados, das características da área e das obras em andamento, considerando os processos geológicos e hidrológicos.

Os resultados obtidos mostraram que, apesar do foco na segmentação de construções e vegetação como elementos críticos para a classificação de risco, a segmentação de veículos apresentou os melhores resultados. O modelo SAM enfrentou limitações na identificação de construções e vegetação, com dificuldades em diferenciar objetos e inconsistências na segmentação. Embora o *Mask R-CNN* tenha melhorado a detecção de veículos, desafios persistiram na segmentação de vegetação e construções. O *MapFlow* destacou-se como uma alternativa promissora para complementar os modelos anteriores, demonstrando maior eficácia na detecção de construções e vegetação.

Além disso, os resultados ressaltaram o potencial das técnicas empregadas para embasar intervenções urbanas e políticas públicas voltadas à mitigação de riscos em áreas urbanas vulneráveis. Apesar das limitações observadas, os métodos aplicados demonstraram viabilidade para a análise de ocupação urbana e áreas de risco. Este trabalho tem como finalidade contribuir para o desenvolvimento de medidas preventivas e estratégias de adaptação que mitiguem os impactos de eventos adversos sobre a população e a infraestrutura urbana. Através da análise de dados geoespaciais e da aplicação de inteligência artificial, o trabalho buscou identificar áreas de risco e padrões que possam prever futuras ocorrências naturais.

Palavras-chave: Segmentação, Geoprocessamento, Área de Risco, Inteligência Artificial.

ABSTRACT

This project aims to apply advanced artificial intelligence techniques, specifically the Segment Anything Model (SAM), for the segmentation of high-resolution aerial images in the Jardim Caboré region, located in the East Zone of São Paulo. In light of the alarming increase in extreme weather events and natural occurrences in the area, the research seeks to identify and analyze the geological and hydrological risk factors, contributing to the safety and resilience of the local community. Initially, a literature review and analysis of documents on the region's risk history were conducted to understand the processes and factors involved. Subsequently, the elements suitable for segmentation in high-resolution images were identified, followed by the execution of segmentation and evaluation of the results using zero-shot and one-shot approaches. Additionally, an integrated analysis of the results, area characteristics, and ongoing works was performed, considering geological and hydrological processes.

The results showed that, despite focusing on the segmentation of constructions and vegetation as critical elements for risk classification, vehicle segmentation yielded the best results. The SAM model faced limitations in identifying constructions and vegetation, struggling with object differentiation and segmentation inconsistencies. Although the Mask R-CNN improved vehicle detection, challenges remained in segmenting vegetation and constructions. The MapFlow emerged as a promising alternative to complement the previous models, demonstrating greater effectiveness in detecting constructions and vegetation.

Moreover, the results highlighted the potential of the employed techniques to support urban interventions and public policies aimed at mitigating risks in vulnerable urban areas. Despite the observed limitations, the applied methods proved feasible for analyzing urban occupation and risk areas. This work aims to contribute to the development of preventive measures and adaptation strategies that mitigate the impacts of adverse events on the population and urban infrastructure. Through geospatial data analysis and the application of artificial intelligence, the project sought to identify risk areas and patterns that may predict future natural occurrences.

Keywords: Segmentation, Geoprocessing, Risk Area, Artificial Intelligence.

AGRADECIMENTOS

Desejo expressar meus mais sinceros agradecimentos a várias pessoas que contribuíram direta ou indiretamente para a realização deste trabalho.

Primeiramente à queridíssima professora Camila, pela orientação, ajuda, apoio, sugestões e discussões nas etapas deste trabalho e por tantos outros motivos que a tornam uma orientadora exemplar. Inclusive, por ter me acolhido como orientando quando as ideias na cabeça estavam nebulosas e ela conseguiu clarear tudo.

Agradeço demais à professora Maria Isabel e ao professor Márcio do MAE pelos conhecimentos e experiências compartilhados durante minha iniciação no mundo do geoprocessamento.

Digo o mesmo ao professor Cláudio Mora, que, em meio a momentos difíceis, me trouxe uma oportunidade de conhecer e me interessar por aplicações de Machine Learning.

Ao professor Boggi, por quem tenho imenso carinho. Grato pelo entendimento e apoio nos momentos difíceis dos últimos meses.

A todo o pessoal do setor Ambiental da SIURB (estágio), que teve a maior compreensão quando surgiram diversos empecilhos durante os processos do trabalho. Em especial ao Douglas e ao Tazio, que tanto me ensinaram e demonstraram paciência para esclarecimentos.

Agradeço demais a todos amigos que me acompanharam durante toda a graduação e que fizeram parte, de alguma forma, para que eu chegasse a este trabalho – das boas companhias da turma do LiGEA de 2017 até as turmas da Geologia – 57 à 64 (e contando...); ao pessoal que está na pós e trabalhando mundo afora. Em especial, a quem tenho enorme gratidão:

Giorgio, meu irmão sedimentar, que me deu forças nos tempos sombrios da pandemia.

Dodô, irmãozão que fiz em 2017, quando entramos juntos no IGc e, desde então, compartilhamos muitas ideias e rolês.

Carol, Rhu, Morfs, a tríade que mora no meu coração, a equipe linha de frente que passou tanto perrengue comigo e dividiu tantas aventuras e risadas.

Maffê e Gi, minha dupla favorita, a clássica camisa 9 e 10, que sempre me fortaleceram nas dúvidas mais cruéis que apareceram nos últimos meses e pelas boas gargalhadas que compartilhamos.

Treps, a minha querida mamãe, conterrânea e palestrina, que me fortaleceu imensamente nos primórdios do curso.

Xups, palestrino do peito, que trabalhou e me ajudou tanto na fase do estágio na geotecnia.

Lô, pelas maluquices e vivências que compartilhamos – em destaque, explorando xistos em Andrelândia – e pelos tantos mapas que realizamos, o que me fez tomar ainda mais gosto pelo mapeamento.

Coxorriha, Jay, Tche, Bidu, Livia, Camis, Caips, Naza, Ni, Chapo, Chorão, Mingau, Nômade, Picareta, Tico, Biro, Haxi, Xistose, Hantaro, Calango, Paladino, Steve, Charada, Chicote, Minalba, Mali e tantas outrxs que acompanharam essa trajetória acadêmica.

Grato demais aos meus queridos amigos da Geo FFLCH – aos dois “meninos” que dividiram as melhores experiências da graduação e de tantos outros eventos. Meu irmão Olegário, agradeço até mesmo pelo frio que passamos em janeiro.

Aos meus queridos amigos que fiz pela vida e pelas épocas de escola. Em especial ao Glauber, irmãozão desde a 3ª série (quarto ano), que, mesmo estando distante, sempre me enviou boas energias e muito apoio. E também ao Lucas, pelo companheirismo e carinho que há desde o ensino médio.

A toda a família Pereira, que sempre me incentivou e fortaleceu de diversas formas, com todo o carinho e apoio que me ajudaram a seguir firme e forte. Obrigado, Kauê, Val, Pri, tia Leda, tia Ju, tio Tifa e todos os meus primos e tios.

E à família Brasil, sem palavras! Se hoje estou realizando um sonho, é porque tive muito apoio, respaldo, amor e amparo de vocês. A todas as mulheres em quem me espelho e que para sempre me servirão de inspiração. Obrigado por tudo, vó Lena, vô Val, tia Sandra, tia Piinha, tia Ninin, tio Du e minha prima querida, Niol.

Ao companheirismo das minhas bolinhas de pelo – Otto, Olga e Mirabel – que me acompanham nessa odisseia da graduação desde a pós-pandemia.

Aos meus coroa de sangue quente, Samara e o Vando – guerreiros, fortes, corajosos e inspiradores. São os pilares de tudo em minha vida, sempre dispostos a resolver qualquer situação e que me transmitiram os mais importantes valores, fundamentais para minha formação. E à minha irmã Vale Tartaruga, sou profundamente grato por acompanhar e participar do seu crescimento. Minha trompetista, que tanto me pergunta sobre pedras e suas funções mágicas, que aprontou tantas traquinagens e fez tanta palhaçada comigo, especialmente as caretas idênticas.

E se falo da minha pequena Tartaruga, preciso lembrar da Alice, a Borboleta Bailarina que voou em 2023, nos deixando profundamente tristes com sua rápida partida, o que me fez por um momento, perder a fé em toda a caminhada que trilhei. Mas, com muito choro e saudade, posso dizer que jamais esqueci das brincadeiras em Piracaia, de todas as pedrinhas que adorava pegar na estrada e me perguntava seus nomes. Sua ida me fez tirar forças de onde não havia para arregaçar as mangas e realizar este trabalho, sempre lembrando que:

“A única luta que se perde é a que se abandona. E nós nunca, nunca abandonamos luta”

"Não tive tempo para ter medo"
(Carlos Marighella)

SUMÁRIO

1. INTRODUÇÃO E JUSTIFICATIVA	8
2. OBJETIVOS	9
3. CARACTERIZAÇÃO DA ÁREA DE ESTUDO	10
4. FUNDAMENTAÇÃO BIBLIOGRÁFICA	13
4.1. Inteligência Artificial e seus ramos	13
4.2. <i>Segment Anything Model (SAM)</i>	16
4.2.1. Text SAM	17
4.3. Arquitetura do SAM	18
4.4. SAM 2	19
4.5. <i>Mask R-CNN</i>	20
4.6. <i>MapFlow</i>	21
5. MATERIAIS E MÉTODOS	22
5.1. Revisão da Literatura	23
5.2. Identificação de Elementos Condicionantes de Risco	23
5.3. Delimitação da Área de Estudo e Segmentação de Imagens	24
5.4. Verificação, Validação e Análise Integrada dos Resultados	26
6. RESULTADOS OBTIDOS	27
6.1. Modelo pré-treinado - SAM	27
6.2. Modelo pré-treinado – <i>Tree Segmentation</i>	31
6.3. Outros modelos (<i>Mask R-CNN</i> e <i>MapFlow</i>)	33
6.3.1. Primeiros Testes e Configurações Iniciais	33
6.3.2. Evolução dos Testes: Introdução de Novas Classes e Parâmetros	34
6.3.3. Testes Adicionais para Outras Classes e Limitações do Modelo	36
6.3.4. Teste Final com o Plugin <i>MapFlow</i> no QGIS	36
7. INTERPRETAÇÃO E DISCUSSÃO DOS RESULTADOS	38
8. CONCLUSÕES	42
9. REFERÊNCIAS BIBLIOGRÁFICAS	44

1. INTRODUÇÃO E JUSTIFICATIVA

A presença crescente de eventos climáticos extremos tem gerado um aumento significativo nos riscos geológicos e hidrológicos. Este aumento é especialmente preocupante em áreas urbanas, onde há maior exposição de pessoas e infraestrutura, impactando diretamente a segurança e qualidade de vida das populações. Segundo relatório da Organização das Nações Unidas (ONU), a perspectiva futura é de que as áreas urbanas, que representavam 56% em 2021, aumentem para 68% em 2050 (UN-Habitat, 2022). Este crescimento acelerado traz diversos desafios, especialmente aqueles que se relacionam ao Objetivo de Desenvolvimento Sustentável da ONU número 11: Tornar as cidades e comunidades mais inclusivas, seguras, resilientes e sustentáveis.

Segundo dados do Instituto Nacional de Pesquisas Espaciais (INPE), entre 2019 e 2023, a Região Metropolitana de São Paulo registrou um aumento alarmante no número de ocorrências relacionadas a ocorrências naturais, com destaque para a Zona Leste da cidade. Em particular, o Jardim Caboré, localizado nessa região, tem sido afetado por uma série de eventos adversos, incluindo inundações recorrentes e deslizamentos de encostas (Rinaldi, 2023), colocando em risco a segurança e integridade da comunidade local.

Diante desse cenário, a adoção de ferramentas que automatizam o processo de análise de imagens para fins de monitoramento da ocupação e dos condicionantes de risco é extremamente valioso. A identificação e monitoramento feitos de forma automática permitem, por exemplo, melhorar modelos de suscetibilidade aos processos superficiais; a atualização em escala temporal mais próxima daquela em que ocorre a ocupação urbana e a modificação do ambiente; análises objetivas que podem ser reproduzidas de forma homogênea em todo território; melhor suporte à tomada de decisão, como no caso de execução de intervenções estruturais ou elaboração de políticas públicas setoriais; e uma resposta rápida em cenários de desastre.

O presente trabalho propõe a aplicação de um Modelo Fundacional (*Foundation Model*) para a segmentação de imagens aéreas de alta resolução, de modo a subsidiar o mapeamento objetivo e automatizado de riscos geológicos e hidrológicos na região do Jardim Caboré. Para isso, foi escolhido o *Segment Anything Model* (SAM), um *Foundation Model* (FM), também chamado de modelo base, criado para segmentação de imagens pela *Meta AI*, um laboratório de inteligência artificial que pertence à *Meta Platforms Inc.* (Kirillov et al., 2023). Ao contrário dos modelos tradicionais, que exigem treinamento extensivo em tarefas específicas, o design de trabalho do SAM adota uma abordagem mais adaptável, que potencializa a anotação de dados e permite a transferência instantânea para uma variedade de tarefas por meio de engenharia de *prompt*. Estudos recentes demonstram o potencial deste

modelo, que usa modelos de Aprendizado Profundo (*Deep Learning*) como suporte principal, na segmentação de imagens em diversos contextos, incluindo na análise e interpretação de dados espaciais (Osco et al., 2023).

Desta forma, este trabalho visa desenvolver um método de análise que integre informações provenientes de imagens aéreas de alta resolução, dados geotécnicos/topográficos e históricos de ocorrências na área de estudo.

2. OBJETIVOS

O principal objetivo deste trabalho é empregar tecnologias avançadas de inteligência artificial na segmentação de imagens aerofotogramétricas (ortofotos) integradas ao Aprendizado de Máquina (ML) para identificar e analisar áreas de risco geológico/hidrológico adjacentes ao leito do Córrego Caboré (Rio Aricanduva), localizado no Jardim Caboré, Zona Leste de São Paulo. Este trecho está em fase de obras de canalização, conduzidas pela Secretaria Municipal de Infraestrutura Urbana e Obras (SIURB), o que justifica a necessidade de análise das imagens para compreender o impacto dessas obras no contexto dos riscos geológicos/hidrológicos da região.

Para isso, serão perseguidos os seguintes objetivos específicos:

- Realizar levantamento bibliográfico e documental sobre o histórico de risco da região, de modo a entender os processos atuantes e seus condicionantes;
- Identificar os elementos condicionantes de risco passíveis de segmentação nas imagens de alta resolução;
- Realizar a segmentação dos elementos identificados em imagens de alta resolução da área de estudo;
- Avaliar o resultado da segmentação sem exemplos (*zero-shot*) e com exemplos (*one-shot*);
- Analisar de forma conjunta os resultados da segmentação, as características da área e as obras em execução, face aos processos geológicos e hidrológicos.

Como resultado, o trabalho visa contribuir para o desenvolvimento de medidas preventivas e estratégias de adaptação que possam mitigar os impactos desses eventos sobre a população e a infraestrutura urbana. Por meio da análise detalhada de dados geoespaciais e da aplicação de inteligência artificial, busca-se não apenas identificar áreas de risco, mas também compreender os padrões e tendências que possam indicar a ocorrência de futuros ocorrências naturais.

3. CARACTERIZAÇÃO DA ÁREA DE ESTUDO

A área de estudo deste trabalho está localizada na região da Zona Leste do município de São Paulo (Figura 3.1), e compreende um trecho do Rio Aricanduva, popularmente conhecido como Córrego do Caboré, que pertence à Bacia Hidrográfica do Aricanduva e delimita as subprefeituras de Itaquera (a norte) e São Mateus (a sul). O foco do estudo está na área composta geologicamente por rochas do Complexo Embu (NP1e) vista na Figura 3.2 – especificamente na Unidade predominante de sericita xisto (NP1esx)(Fernandes, 1991), vista em afloramento do corpo em situ após escavações na Figura 3.3.

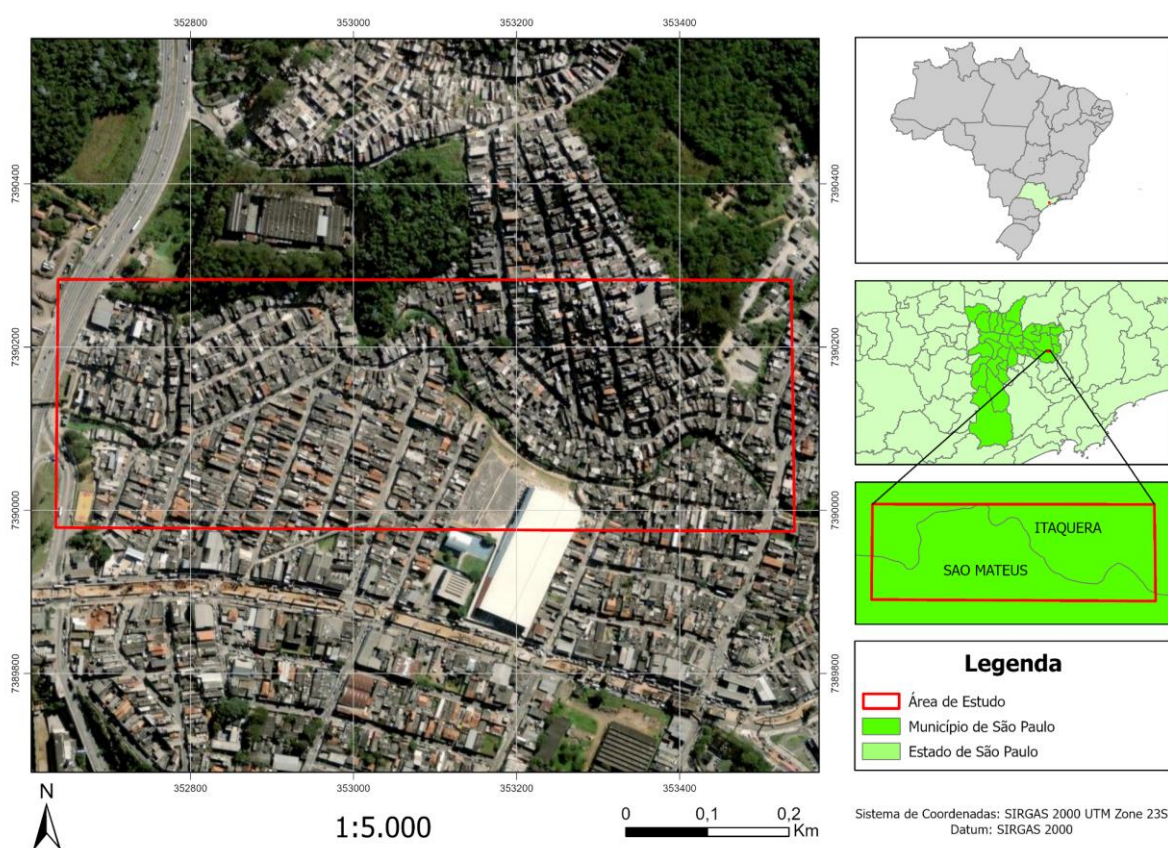


Figura 3.1 – Localização da área de estudo. Fonte: Geosampa

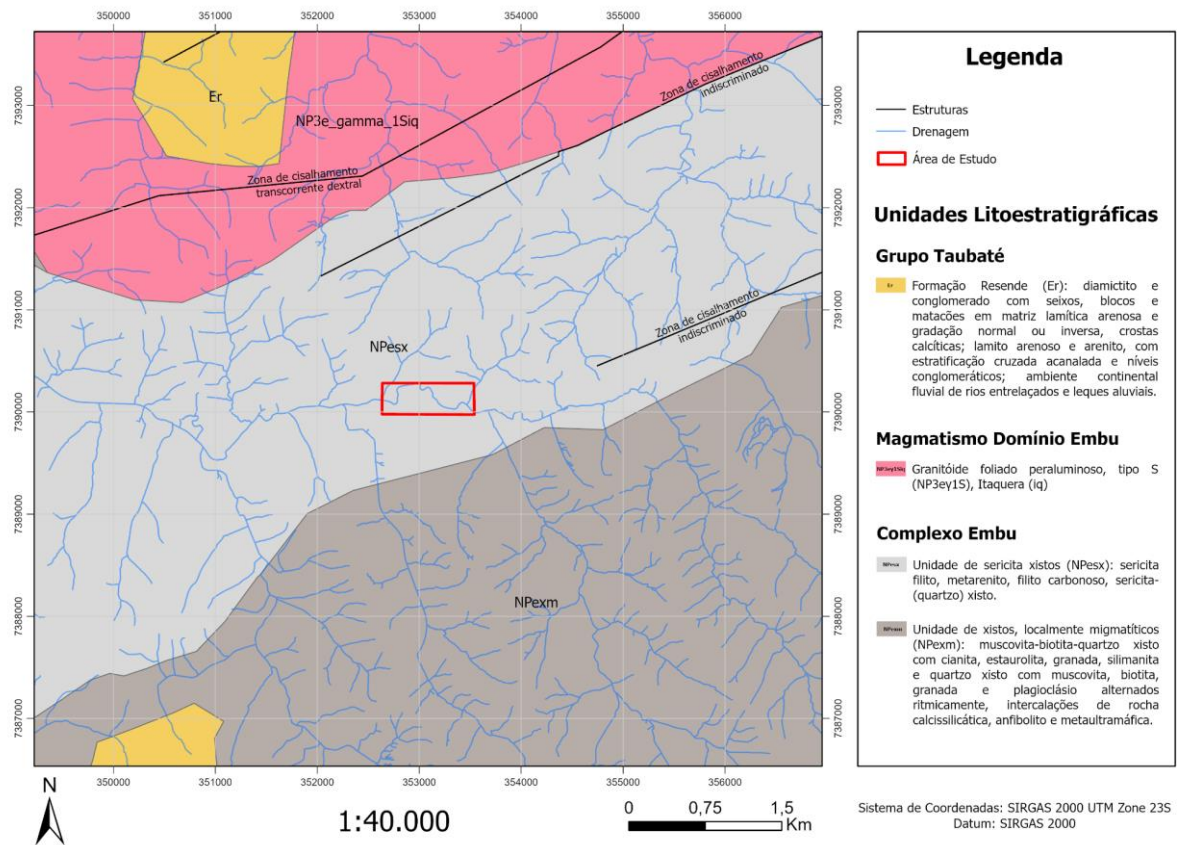


Figura 3.2 – Geologia Regional da área de estudo. Fonte: Geosampa



Figura 3.3 – Solo de alteração de rocha da Unidade - NP1esx. Foto sem escala. Fonte: Autor

Na área de estudo pode-se observar os seguintes riscos: geológicos e hidrológicos. Os hidrológicos caracterizam-se por inundações, alagamentos e enchentes que estão distribuídos ao longo do trecho e classificados por graus R2 (médio) e R3 (alto) (Figura 3.3). Segundo Relatório Técnico HSM (PMSP, 2023a), essas áreas estão setorizadas pelas siglas HSM-01 e HSM-02 para a subprefeitura de São Mateus (margem sul do trecho) e segundo Relatório Técnico HIQ (PMSP, 2023b), HIQ-06 e HIQ-07 são para as áreas localizadas na subprefeitura de Itaquera (margem norte do trecho).

Já para os riscos geológicos, os tipos são: solapamento e escorregamento, que também estão classificados como R2 (médio) e R3 (alto), respectivamente (Figura 3.4). Segundo Relatório Técnico SM (COMDEC, 2023), essas áreas estão setorizadas pela sigla SM-22 para a subprefeitura de São Mateus (margem sul do trecho) e, segundo Relatório Técnico IQ (COMDEC, 2023), pela sigla IQ-13 para as áreas localizadas na subprefeitura de Itaquera (margem norte do trecho).

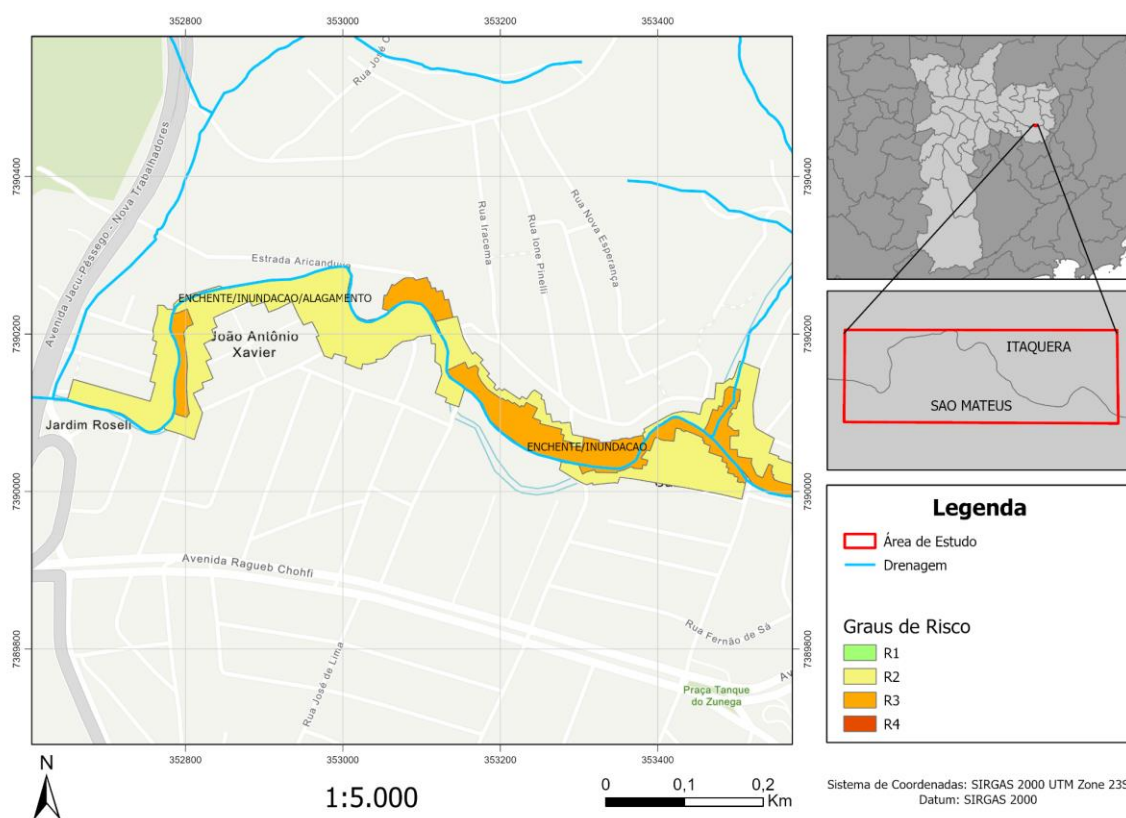


Figura 3.4 – Setores de Riscos Hidrológicos na área de estudo. Fonte: Geosampa

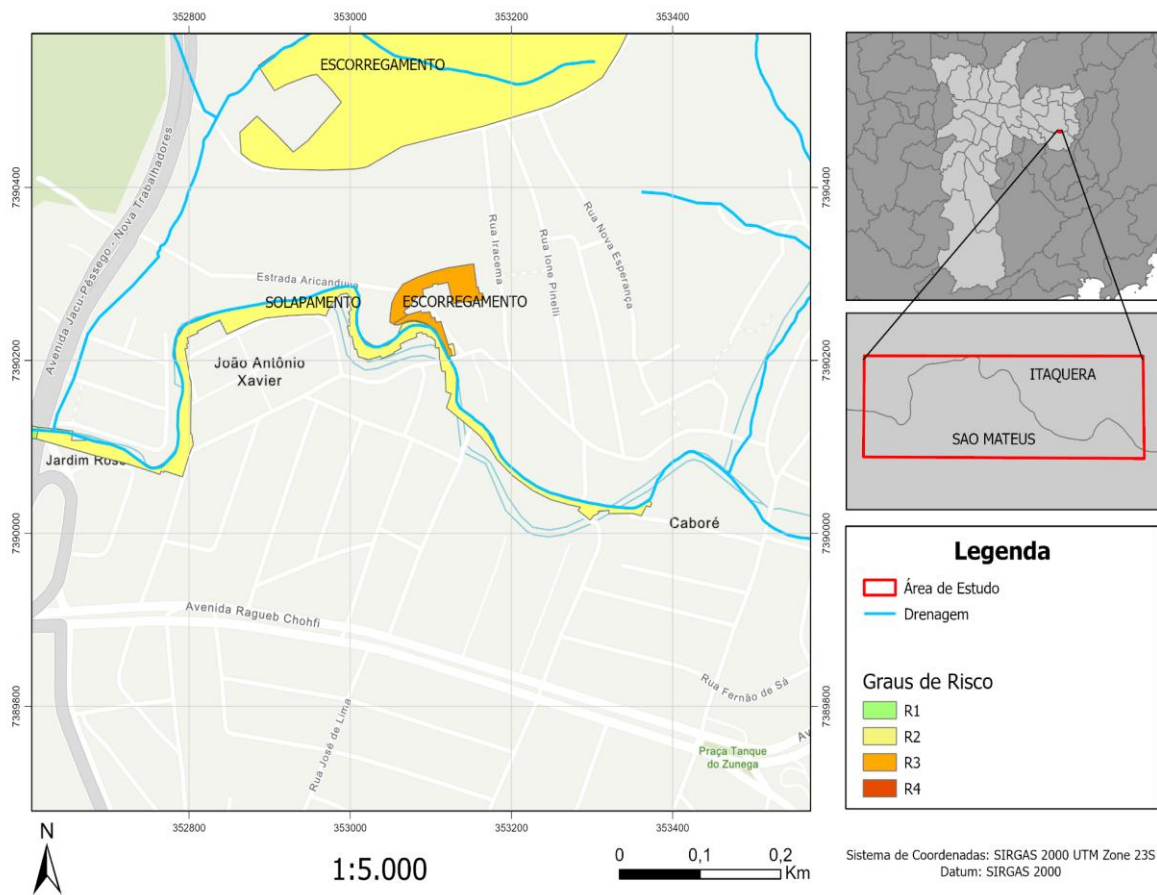


Figura 3.5 – Setores de Riscos Geológicos na área de estudo. Fonte: Geosampa

4. FUNDAMENTAÇÃO BIBLIOGRÁFICA

A inteligência artificial (IA) é um campo em constante evolução, com múltiplas ramificações que visam simular e replicar aspectos da inteligência humana em máquinas. Seus subcampos abordam diferentes desafios e aplicações, possibilitando avanços significativos em diversas áreas. A seguir, serão apresentados os conceitos fundamentais e as principais ramificações da IA, com destaque para suas características e potencialidades.

4.1. Inteligência Artificial e seus ramos

A inteligência artificial é um campo interdisciplinar da ciência da computação que se dedica ao estudo e desenvolvimento de sistemas e algoritmos capazes de simular, em parte ou na totalidade, processos cognitivos humanos. Envolve a concepção e implementação de técnicas que habilitam máquinas a realizar tarefas complexas que normalmente exigiriam inteligência humana, tais como: percepção, raciocínio, aprendizado, planejamento e tomada de decisão. Como definido por Russell e Norvig (2014), a inteligência artificial busca criar agentes

racionais que possam interagir de forma autônoma com o ambiente, adaptando-se e aprendendo com a experiência, com o intuito de alcançar objetivos específicos.

A inteligência artificial se subdivide em vários ramos e subáreas (Figura 4.1.1) e cada ramificação representa aspectos específicos do desenvolvimento e aplicação de tecnologias inteligentes (Russell & Norvig, 2014). Entre os principais ramos, destacam-se o Aprendizado de Máquina (*Machine Learning*), o Processamento de Linguagem Natural (*Natural Language Processing*), a Visão Computacional (*Computer Vision*), a Robótica Inteligente (*Intelligent Robotics*), entre outros.

O Aprendizado de Máquina (*Machine Learning*) é um dos principais ramos da inteligência artificial e tem sido amplamente utilizado na indústria. Ele envolve o desenvolvimento de algoritmos que permitem que as máquinas aprendam com dados e melhorem seu desempenho ao longo do tempo. "O aprendizado de máquina é o estudo de algoritmos que melhoram automaticamente através da experiência" (Mitchell et al., 1997). Com o aprendizado de máquina, as máquinas podem automatizar tarefas complexas, detectar padrões e tomar decisões com base nos dados disponíveis.

O Processamento de Linguagem Natural (*Natural Language Processing*) é uma área importante da inteligência artificial que se concentra em permitir que as máquinas entendam e processem a linguagem humana de forma eficaz. "PLN é um campo de inteligência artificial que dá aos computadores a capacidade de entender, interpretar e manipular a linguagem humana" (Jurafsky et al., 2009). Aplicações de PLN incluem assistentes virtuais, *chatbots* e sistemas de tradução automática.

Os Sistemas Baseados em Regras (*Rules-based Systems*) utilizam uma série de regras pré-definidas para tomar decisões ou inferir novos dados a partir de informações conhecidas (Giarratano & Riley, 2005). Estes sistemas são comuns em aplicações que requerem tomadas de decisão repetitivas e padronizadas, onde as regras podem ser claramente definidas.

O Planejamento, Agendamento e Otimização (*Planning, Scheduling, Optimization*) são áreas da IA que se dedicam à criação de estratégias para atingir objetivos específicos (Russell & Norvig, 2014). Estas técnicas são utilizadas para determinar a melhor sequência de ações ou a alocação ideal de recursos para a execução eficiente de tarefas complexas.

O Reconhecimento de Fala (*Speech*) é uma tecnologia de IA que permite a interpretação e a transcrição da fala humana (Jurafsky et al., 2009). Essa tecnologia é utilizada em assistentes virtuais, sistemas de ditado e em interfaces de usuário que dependem de comandos de voz.

A Visão Computacional (*Computer Vision*) é um campo da inteligência artificial que permite que as máquinas "vejam" e interpretem imagens ou vídeos. "Visão computacional é o campo que estuda como construir sistemas artificiais que obtêm informações úteis a partir de

imagens" (Szeliski et al., 2010). Por meio de algoritmos avançados, as máquinas podem reconhecer objetos, pessoas e até mesmo emoções em imagens.

A Robótica Inteligente (*Intelligent Robotics*) combina a inteligência artificial com a robótica para criar máquinas autônomas que podem interagir e realizar tarefas sem a necessidade de intervenção humana constante. Segundo Siciliano e Khatib (2016) - "robôs inteligentes são máquinas que podem realizar tarefas de forma autônoma ou semiautônoma, interagindo com o ambiente e executando funções complexas". Os robôs inteligentes podem ser programados para executar várias funções, desde tarefas industriais complexas até assistência em atividades diárias.

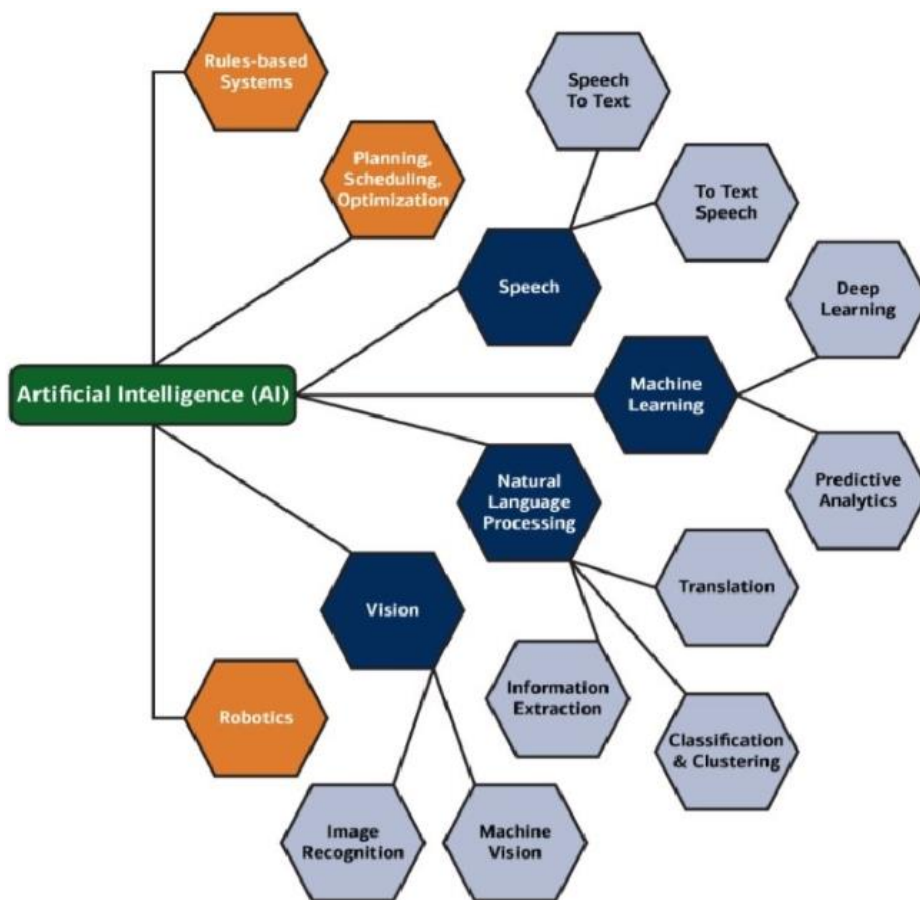


Figura 4.1.1 – Ramificações da Inteligência Artificial. Fonte: Moodle USP

Os modelos de inteligência artificial, como o *Segment Anything Model (SAM)* e o *Mask R-CNN*, têm se tornado essenciais em diversas áreas devido à sua capacidade de processar e analisar grandes volumes de dados. Esses modelos são projetados para resolver problemas complexos, permitindo que as máquinas realizem tarefas que antes exigiam inteligência humana.

O SAM, com sua natureza multimodal, permite a segmentação de imagens através de diferentes tipos de prompts, sendo particularmente útil em aplicações que requerem a extração de informações detalhadas (Kirillov et al., 2023). O *Mask R-CNN*, por sua vez, não apenas identifica objetos em uma imagem, mas também fornece máscaras de segmentação, o que é crucial em contextos que exigem precisão, como na análise de imagens médicas (He et al., 2017).

A adoção dessas tecnologias também promove a automação de processos, aumentando a eficiência operacional e permitindo que os profissionais se concentrem em atividades mais estratégicas (Russell & Norvig, 2014).

4.2. Segment Anything Model (SAM)

A inteligência artificial utilizada para o desenvolvimento deste trabalho é chamada de SAM - *Segment Anything Model*, um modelo de segmentação de imagens desenvolvido pela *Meta AI*, anteriormente conhecida como *Facebook Reality Labs*, é uma divisão da *Meta Platforms, Inc.* (anteriormente conhecida como *Facebook, Inc.*) focada em pesquisa e desenvolvimento de tecnologias avançadas de inteligência artificial, realidade virtual e aumentada. A *Meta AI* desenvolve sistemas e modelos de IA inovadores para uma variedade de aplicações, incluindo processamento de imagens, reconhecimento de padrões, processamento de linguagem natural, entre outros.

O SAM é um modelo avançado de segmentação de imagens capaz de identificar e delinear objetos em imagens com alta precisão. Ele foi treinado em um vasto conjunto de dados composto por milhões de imagens e bilhões de máscaras de segmentação, o que lhe confere uma robusta capacidade de generalização. (Kirillov et al., 2023)

A segmentação de imagens é o processo de dividir uma imagem em várias partes ou segmentos, cada um representando um objeto ou uma região específica. Essa técnica é fundamental em diversas aplicações de Visão Computacional, pois permite identificar e isolar elementos de interesse em uma imagem. A segmentação pode ser usada para posterior classificação dos objetos, facilitando tarefas como reconhecimento de padrões, detecção de anomalias e análise de cenas.

Uma característica distintiva do SAM é sua natureza multimodal, permitindo a segmentação através de diferentes tipos de *prompts*, incluindo caixas delimitadoras, pontos individuais e descrições de texto. Essa versatilidade faz com que o SAM seja capaz de processar imagens de ultra resolução, extraindo informações detalhadas que podem ser cruciais para aplicações como monitoramento ambiental e análise urbana.

O *Segment Anything Model* (SAM) é considerado um modelo *Deep Learning* que é definido como uma subárea do ramo do Aprendizado de Máquina (*Machine Learning*) que se concentra no treinamento de modelos de rede neural profunda para aprender representações complexas dos dados. Esses modelos são capazes de aprender automaticamente características ou representações dos dados em diferentes níveis de abstração, sem a necessidade de características manualmente projetadas. Isso é possível graças ao uso de algoritmos de treinamento eficazes, como *backpropagation*, que ajustam os parâmetros do modelo para minimizar a diferença entre as previsões do modelo e os rótulos reais dos dados. (Goodfellow, Bengio & Courville, 2016).

Em *Deep Learning*, os modelos de rede neural profunda têm várias camadas de processamento, permitindo que aprendam representações hierárquicas dos dados. No caso do SAM, essas camadas neurais são capazes de aprender a segmentar objetos em uma imagem de maneira progressivamente mais complexa e precisa, à medida que processam as informações de entrada.

4.2.1. Text SAM

Há também o modelo - Text SAM, que é uma abordagem inovadora que combina o *Segment Anything Model* (SAM) com o modelo de visão-linguagem *Grounding Dino*, permitindo a extração de características GIS com base em descrições textuais fornecidas pelo usuário. O Text SAM permite que o usuário forneça *prompts* textuais que descrevem o objeto de interesse, como “*tree*” ou “*building*”, permitindo que o modelo segmente e identifique com precisão esses elementos. (Tuteja, Singh & Pathak, 2024)

Essa integração entre SAM e *Grounding Dino* amplia significativamente a utilidade do modelo em aplicações GIS, possibilitando a extração de características específicas de imagens, como construções, vegetação e outros objetos bem definidos. No entanto, o Text SAM não é adequado para segmentar elementos amorfos ou com limites imprecisos, como áreas de água ou florestas. Para esses casos, modelos de classificação de pixels, como o *High Resolution Land Cover Classification*, são mais apropriados.

Além disso, a precisão do modelo pode ser aprimorada com ajustes como o tamanho da célula, que determina a quantidade de contexto que o modelo observa ao segmentar os objetos, e parâmetros como *threshold* de caixa e *threshold* de texto para refinar os resultados. O uso de ferramentas de pós-processamento, como o ArcGIS Pro, também é recomendado para filtrar e melhorar a segmentação, ajustando detecções pequenas e removendo possíveis ruídos. (Tuteja, Singh & Pathak, 2024)

4.3. Arquitetura do SAM

As capacidades de processamento do *Segment Anything Model* (SAM) (Figura 4.3.1) baseiam-se principalmente na sua arquitetura inovadora, composta por três componentes principais: o codificador de imagem, o codificador de *prompt* e o decodificador de máscara (Figura 4.3.2).

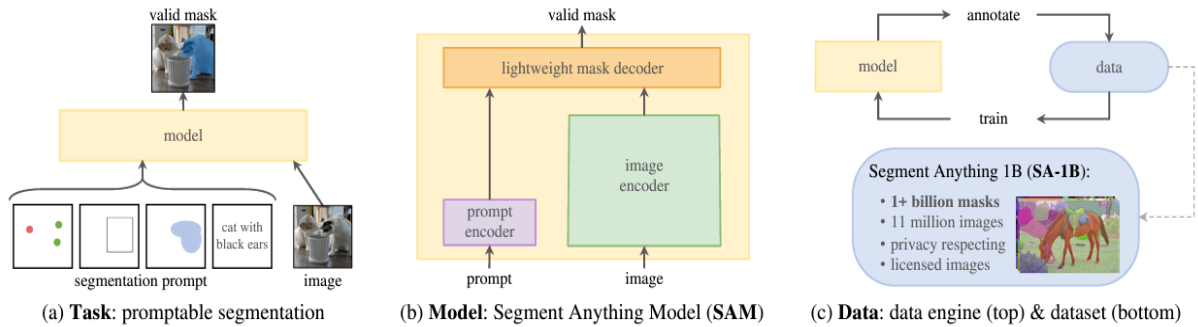


Figura 4.3.1 – O projeto *Segment Anything* (SA) apresenta uma nova tarefa, modelo e conjunto de dados para segmentação de imagens. Fonte: *Segment Anything* (Kirillov et al., 2023)

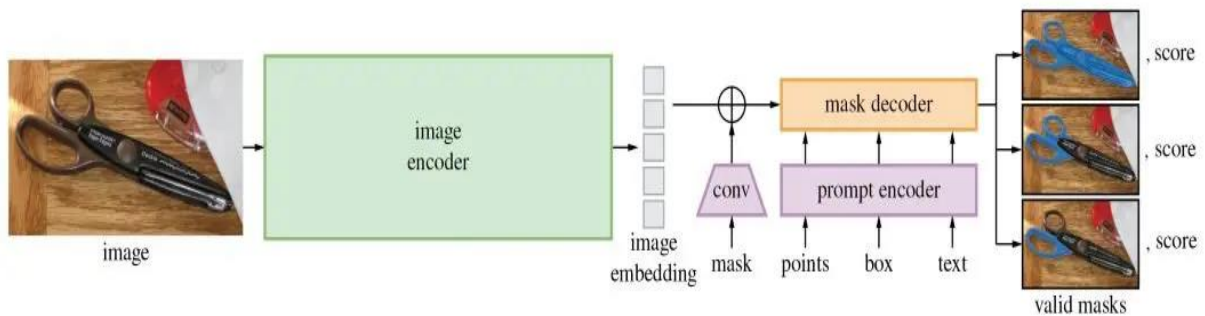


Figura 4.3.2 – A arquitetura do SAM consiste nos seguintes componentes: um codificador de imagem, um decodificador e um decodificador de máscara. – Fonte: *Segment Anything* (Kirillov et al., 2023)

O codificador de imagem é o núcleo da arquitetura do SAM, um componente sofisticado responsável por processar e transformar imagens de entrada em um conjunto abrangente de características. Utilizando uma abordagem baseada em transformadores, semelhante ao que é visto em modelos avançados de PLN, esse codificador comprime as imagens em uma matriz densa de características. Essa matriz forma a base a partir da qual o modelo identifica vários elementos da imagem (Kirillov et al., 2023).

O codificador de *prompt* é um aspecto único do SAM que o diferencia dos modelos tradicionais de segmentação de imagens. Ele interpreta várias formas de *prompts* de entrada, sejam elas baseadas em texto, pontos, máscaras aproximadas ou uma combinação desses elementos. Esse codificador traduz esses *prompts* em uma incorporação que orienta o processo de segmentação, permitindo que o modelo se concentre em áreas ou objetos específicos dentro de uma imagem conforme o indicado pelo *prompt* (Kirillov et al., 2023).

O decodificador de máscara é onde ocorre a segmentação propriamente dita. Ele sintetiza as informações dos codificadores de imagem e de *prompt* para produzir máscaras de segmentação precisas. Esse componente é responsável pelo resultado final, determinando os contornos e áreas exatas de cada segmento dentro da imagem (Kirillov et al., 2023).

A interação entre esses componentes é vital para a segmentação eficaz de imagens:

- O codificador de imagem primeiro cria uma compreensão detalhada da imagem inteira, decompondo-a em características que o motor pode analisar.
- O codificador de *prompt* então adiciona contexto, focando a atenção do modelo com base no input fornecido, seja um simples ponto ou uma descrição textual complexa.
- Finalmente, o decodificador de máscara usa essa informação combinada para segmentar a imagem com precisão, garantindo que o output esteja alinhado com a intenção do *prompt* de entrada.

4.4. SAM 2

O avanço das ferramentas de segmentação foi marcado pelo lançamento do SAM 2, anunciado em 29 de julho de 2024, como a mais recente evolução do modelo da *Meta AI*. Sucessor do *Segment Anything Model (SAM)*, o SAM 2 é um modelo unificado para segmentação de objetos em imagens e vídeos, oferecendo uma transição perfeita entre esses formatos e um desempenho consistente. Ele alcança velocidades de 44 quadros por segundo, o que o torna altamente adequado para aplicações em tempo real, como análise ambiental, mapeamento urbano e monitoramento de áreas de risco. (*Meta AI*, 2024)

Entre as principais inovações do SAM 2 está sua forte generalização *zero-shot*, que permite segmentar objetos que nunca encontrou antes, sem a necessidade de treinamento específico para cada classe. Essa capacidade é essencial em cenários dinâmicos e variados, como áreas urbanas complexas e regiões de ocupação irregular. Além disso, o modelo introduz o refinamento iterativo, possibilitando ajustes contínuos durante a segmentação para corrigir erros ou melhorar a precisão, especialmente em situações onde os objetos aparecem agrupados, sob sombra ou parcialmente ocultos.

O SAM 2 também supera muitos desafios associados à segmentação de vídeo. Com o uso de uma arquitetura baseada em memória, ele consegue rastrear objetos ao longo do tempo, mesmo diante de oclusões, reaparecimentos ou alterações de iluminação. Essa funcionalidade é particularmente útil para analisar mudanças ao longo de vídeos que retratam áreas sujeitas a ocorrências naturais ou ocupações desordenadas. Contudo, o modelo ainda apresenta limitações em vídeos muito longos ou com mudanças drásticas de perspectiva, o que exige intervenções manuais para manter a precisão. (*Meta AI*, 2024)

Outro diferencial do SAM 2 é o suporte ao conjunto de dados SA-V, que contém mais de 51 mil vídeos e 600 mil máscaras de objetos, tornando-se o maior e mais diverso *dataset* de segmentação de vídeo já criado. Esse recurso foi fundamental para melhorar a capacidade do modelo, ao fornecer uma base ampla e detalhada para seu treinamento. Em comparação com o modelo SAM anterior, o SAM 2 oferece ganhos expressivos em precisão, velocidade e redução no tempo de interação humana necessária. (*Meta AI, 2024*)

No contexto do mapeamento de risco, o SAM 2 apresenta grande potencial para identificar construções, vegetação e elementos de infraestrutura em imagens e vídeos. Ele oferece maior flexibilidade para lidar com dados visuais complexos, permitindo uma análise mais detalhada e eficiente de áreas sujeitas a ocupações irregulares ou mudanças ambientais. Apesar das limitações em cenários extremos, o SAM 2 demonstra ser uma ferramenta promissora para automatizar processos de análise visual, trazendo ganhos significativos para o planejamento urbano e a gestão de áreas de risco.

4.5. Mask R-CNN

O *Mask R-CNN* é uma extensão do modelo *Faster R-CNN*, que é uma arquitetura de detecção de objetos amplamente utilizada. O *Faster R-CNN* combina a eficiência da extração de características com a proposta de regiões para identificar e classificar objetos em imagens. Ele consiste em duas etapas principais: a primeira utiliza uma Rede de Proposta de Regiões (RPN) para gerar propostas de regiões que podem conter objetos, enquanto a segunda etapa aplica uma rede *Fast R-CNN* para classificar essas regiões e refinar as caixas delimitadoras (Ren et al., 2015). Essa abordagem permite que o modelo opere em tempo real, tornando-o adequado para aplicações em cenários dinâmicos, como vigilância e robótica.

O *Mask R-CNN* adiciona uma terceira saída ao *Faster R-CNN*, que gera máscaras de segmentação para cada instância detectada. Essa abordagem permite que o modelo não apenas identifique a presença de objetos, mas também delimite suas formas exatas em nível de pixel, resultando em uma segmentação mais precisa e detalhada (He et al., 2018). A arquitetura do *Mask R-CNN* utiliza uma técnica chamada *RoIAlign* (Figura 4.5.1), que resolve problemas de alinhamento espacial, garantindo que as saídas do modelo correspondam de forma mais precisa às entradas. Essa camada é fundamental para preservar a localização exata dos pixels, permitindo uma segmentação mais fiel (He et al., 2018).

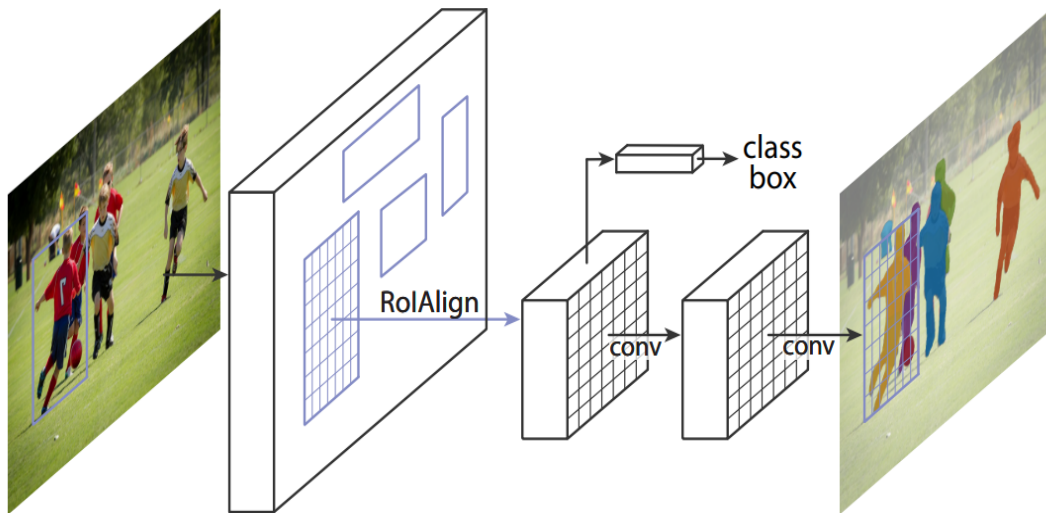


Figura 4.5.1 – A estrutura *Mask R-CNN* para segmentação de instâncias. Fonte: *Mask R-CNN* (He et al., 2018)

O *Mask R-CNN* tem uma ampla gama de aplicações, incluindo segmentação de objetos em imagens, detecção de poses humanas e análise de cenas em vídeos. Sua flexibilidade e eficiência o tornam uma ferramenta valiosa em diversas áreas, como visão computacional, robótica e análise de imagens médicas. Além disso, o *Mask R-CNN* demonstrou resultados superiores em competições de segmentação, como o COCO (*Common Objects in Context*), que é um conjunto de dados amplamente utilizado para treinamento e avaliação de modelos de detecção e segmentação de objetos, contendo imagens com anotações detalhadas de objetos em diferentes contextos e cenários (Lin et al., 2014).

4.6. *MapFlow*

O *Mapflow* é uma plataforma avançada de análise de dados geoespaciais, projetada para processar e extrair informações de imagens de satélite e aéreas utilizando ferramentas baseadas em inteligência artificial (IA). Lançada em 2021 pela *Geoalert*, a plataforma é voltada para transformar imagens em dados úteis e específicos para diversas áreas de aplicação, como mapeamento urbano, monitoramento ambiental, planejamento de infraestrutura e agricultura de precisão. Sua interface intuitiva permite que usuários sem profundo conhecimento técnico em GIS ou programação possam utilizar modelos de IA para realizar tarefas de segmentação e classificação de objetos e características geográficas.

Entre as ferramentas oferecidas, destaca-se o modelo "*Buildings*", que é treinado para identificar telhados de edifícios e fornecer opções de processamento como simplificação de contornos irregulares e integração com dados do *OpenStreetMap* para aumentar a precisão.

Outro modelo importante é o "Forest", especializado em detectar áreas de vegetação, e o "Roads", que se concentra na extração de redes viárias, especialmente em áreas rurais e suburbanas. Essas ferramentas são capazes de processar imagens de alta resolução e são projetadas para se adaptar a diferentes regiões geográficas, permitindo a escolha de modelos específicos conforme as características locais (MapFlow AI, 2023).

Além disso, o Mapflow também integra fluxos de trabalho avançados como o *Segment Anything Model* (SAM), que permite a classificação de uma vasta gama de objetos nas imagens com base em diferentes níveis de zoom. Essa funcionalidade é extremamente útil para tarefas que envolvem o uso do solo, identificação de áreas agrícolas e reconhecimento de estruturas diversas, como edifícios e rodovias (MapFlow AI, 2023).

O Mapflow oferece versões pagas e gratuitas, sendo que as opções pagas proporcionam maior capacidade de processamento e mais créditos para análise de imagens (MapFlow AI, 2023). Para utilizar a plataforma, é necessário ter imagens georreferenciadas e seguir o formato adequado, como GeoTIFF. A plataforma também se integra com softwares GIS como QGIS, permitindo que os resultados sejam exportados para análises mais detalhadas.

Em resumo, o Mapflow se destaca por sua capacidade de fornecer resultados de alta precisão e por ser uma solução acessível e poderosa para profissionais e empresas que buscam aproveitar os dados geoespaciais em diversas áreas de atuação (MapFlow AI, 2023).

5. MATERIAIS E MÉTODOS

Este capítulo apresenta os procedimentos metodológicos adotados para o desenvolvimento do trabalho, desde a revisão da literatura até a aplicação de modelos de segmentação de objetos em imagens de alta resolução. As análises focaram na identificação de elementos de risco geológico e hidrológico na região do Jardim Caboré, em São Paulo, utilizando imagens fornecidas pelo portal Geosampa.

Os principais modelos aplicados foram o SAM (*Segment Anything Model*), o *Mask R-CNN* e o complemento *MapFlow* no QGIS, além do *Tree Segmentation* no ArcGIS Pro, cada um ajustado para segmentar construções, veículos e vegetação. Esses testes foram realizados com o objetivo de avaliar a eficácia dos métodos em diferentes cenários e contribuir para o mapeamento de risco na área de estudo.

As análises foram realizadas em uma máquina com as seguintes especificações e softwares:

- Especificações:
 - **Processador:** Intel(R) Core(TM) i5-10300H CPU @ 2.50GHz
 - **RAM:** 32 GB (utilizável: 31,8 GB)

- **GPU 1:** NVIDIA GeForce GTX 1650 Ti – 4 GB
- Softwares:
 - **ArcGIS Pro (versão 3.0.2):** para execução de modelos de segmentação e análises espaciais.
 - **QGIS (versão 3.36.1):** para análises complementares e uso do complemento *MapFlow*.
 - **Adobe Fireworks CS6:** para edição e ajustes nas imagens processadas.

5.1. Revisão da Literatura

A primeira etapa do trabalho consistiu em uma revisão abrangente da literatura existente sobre riscos geológicos e hidrológicos, com foco na região do Jardim Caboré. Essa revisão incluiu a busca em bases de dados acadêmicas, como *Scopus*, *Web of Science* e *Google Scholar*, além de revistas indexadas como *Computers & Geoscience*, *Remote Sensing*, e *Artificial Intelligence in Geosciences*. Também foram consultados relatórios e documentos da prefeitura de São Paulo, que fornecem dados históricos sobre ocorrências naturais na área. O objetivo dessa revisão foi entender os processos atuantes e os condicionantes de risco que afetam a região.

5.2. Identificação de Elementos Condicionantes de Risco

A identificação dos elementos condicionantes de risco foi realizada exclusivamente a partir da análise de imagens aéreas de alta resolução. Essa abordagem permitiu observar as características físicas e ocupacionais do entorno do Córrego Caboré, correlacionando os dados com os conceitos e diretrizes descritos no Manual de Mapeamento de Riscos em Encostas e Margens de Rios (IPT, 2007). Além disso, foram consideradas classificações prévias que indicaram setores da área como pertencentes aos níveis de risco R2 (médio) e R3 (alto).

Esse processo integrou os aspectos conceituais de risco com os fatores que condicionam a ocorrência de acidentes, como enchentes e inundações em áreas urbanas.

Conforme o manual, o risco é entendido como a interação entre a suscetibilidade do terreno e a vulnerabilidade das ocupações humanas (IPT, 2007, p. 25). Já as áreas de risco – são áreas passíveis de serem afetadas por processos naturais ou induzidos, como deslizamentos e inundações, com potencial para causar danos materiais e perdas humanas (IPT, 2007, p. 26).

Condicionantes de Risco Identificados

- **Tipo de Uso do Solo:** Áreas urbanizadas apresentaram maior risco de deslizamentos devido à impermeabilização do solo e aumento da carga causada pelas construções (IPT, 2007, p. 41). Áreas naturais, com solo exposto ou cobertura vegetal insuficiente, mostraram maior suscetibilidade à erosão (IPT, 2007, p. 42).
- **Drenagem e Cursos d'Água:** Drenagem inadequada contribuiu para saturação do solo e aumento dos riscos de deslizamento e solapamento, especialmente em regiões de ocupação irregular (IPT, 2007, p. 41). Proximidade de cursos d'água apresentou risco elevado de inundações, agravado pela canalização inadequada e presença de margens instáveis (IPT, 2007, p. 47).
- **Cobertura Vegetal:** A vegetação foi identificada como estabilizadora do solo, reduzindo a erosão e a infiltração excessiva de água. Sua ausência foi associada a maior vulnerabilidade (IPT, 2007, p. 42).
- **Geometria do Terreno:** Declividades superiores a 25° foram identificadas como áreas críticas para deslizamentos, especialmente em encostas com baixa resistência do solo (IPT, 2007, p. 44).
- **Infraestrutura e Construções:** Construções em áreas vulneráveis, como encostas e margens de rios, aumentaram a carga no terreno e a probabilidade de instabilidade, especialmente em ocupações irregulares (IPT, 2007, p. 47).
- **Áreas de Contenção ou Barreiras Naturais:** Barreiras vegetativas ou construídas, como faixas de mata ciliar, foram observadas como medidas de mitigação de risco, mas estavam ausentes em muitas áreas analisadas (IPT, 2007, p. 48).

5.3. Delimitação da Área de Estudo e Segmentação de Imagens

Para a delimitação da área de estudo, foi utilizado o software QGIS, complementado pelo ArcGIS Pro. A região selecionada é o Jardim Caboré, na Zona Leste de São Paulo. O ArcGIS Pro permitiu uma análise espacial detalhada e a integração de diversas camadas de informação geográfica. Bibliotecas Python específicas foram instaladas para permitir a

execução do modelo SAM (*Segment Anything Model*) diretamente dentro do ambiente do ArcGIS Pro.

As imagens de alta resolução de 2020, disponíveis no portal Geosampa através do serviço WMS, foram importadas para ambos os softwares. Inicialmente, a abordagem *zero-shot* foi utilizada no QGIS e ArcGIS Pro, onde o modelo SAM foi empregado para segmentar as imagens diretamente, sem qualquer treinamento prévio. Essa fase teve como objetivo identificar os elementos que o modelo consegue segmentar autonomamente, tais como tipos de telhados de edifícios, cobertura vegetal, construções, cursos d'água, áreas de encosta, veículos entre outros.

Para realizar a segmentação, foram seguidos os passos abaixo:

1. **Importação das Imagens:** As imagens foram importadas do portal Geosampa para os softwares QGIS e ArcGIS Pro, utilizando o serviço WMS para garantir a resolução e atualização das imagens.
2. **Configuração do SAM no ArcGIS Pro:** Foram instaladas as bibliotecas Python necessárias para rodar o modelo SAM dentro do ArcGIS Pro. Isso incluiu a configuração do ambiente de desenvolvimento e a integração do modelo de segmentação.
3. **Segmentação Zero-Shot:** No primeiro experimento, o modelo SAM foi executado em sua forma *zero-shot*, ou seja, segmentando as imagens diretamente sem qualquer treinamento adicional. Esse processo foi conduzido tanto no QGIS quanto no ArcGIS Pro para comparar a eficácia do modelo em diferentes plataformas.
4. **Treinamento e Segmentação One-Shot:** Em seguida, foram fornecidos exemplos específicos dos objetos de interesse ao modelo SAM para realizar uma segmentação *one-shot*. Esse treinamento visou melhorar a precisão do modelo na identificação de elementos previamente definidos.

Após a análise inicial realizada com o SAM (*Segment Anything Model*), foram conduzidos novos testes com o modelo *Mask R-CNN* para segmentação de diferentes classes de interesse na área de estudo, como construções, veículos e vegetação. O *Mask R-CNN*, utilizado para comparação e ajuste da metodologia, é um modelo de segmentação de instâncias que trabalha identificando objetos e gerando máscaras pixel a pixel para cada classe detectada. Sua implementação envolveu treinamento supervisionado com rotulagem manual dos objetos, seguido por testes para avaliar a precisão da segmentação na área de estudo. Embora tenha apresentado melhorias em relação ao SAM para a detecção de veículos, ainda houve erros, como falsos positivos na identificação de telhados e objetos sobre as lajes, e dificuldades na segmentação de vegetação e construções.

Após os testes preliminares com o *Mask R-CNN*, foi utilizado o *Mapflow*, um complemento do QGIS desenvolvido para segmentação automatizada. O *Mapflow* oferece suporte à detecção de classes urbanas e naturais, incluindo construções, ruas, florestas, plantações e áreas em construção. Seu diferencial é o uso de modelos de inteligência artificial treinados para identificar padrões urbanos e rurais diretamente em imagens de satélite, facilitando análises detalhadas e automatizadas. No contexto deste trabalho, o *MapFlow* foi utilizado como ferramenta comparativa para avaliar sua precisão e eficácia em relação aos modelos previamente testados.

Além disso, foi utilizado o *Tree Segmentation*, uma ferramenta baseada em *Deep Learning* disponível no ArcGIS, especialmente projetada para a segmentação de copas de árvores em imagens de alta resolução. Conforme descrito no site oficial da Esri, o *Tree Segmentation* emprega algoritmos de aprendizado profundo para identificar e segmentar árvores em um nível individual, considerando variações de tamanho, forma e densidade. Essa ferramenta é amplamente utilizada para análises florestais, mas, neste trabalho, foi aplicada com o objetivo de mapear vegetação em áreas urbanas e avaliar sua aplicabilidade para detecção de risco. O processo envolveu o uso de imagens aéreas e configuração de parâmetros específicos para maximizar a precisão da segmentação das árvores no contexto urbano.

Essas abordagens combinadas permitiram uma análise comparativa dos métodos de segmentação, destacando as limitações e os potenciais de cada ferramenta no contexto de mapeamento de áreas de risco geológico e hidrológico.

5.4. Verificação, Validação e Análise Integrada dos Resultados

Os objetos segmentados automaticamente pelo modelo SAM foram verificados por meio de uma análise comparativa com resultados de outros trabalhos na área de segmentação de imagens aéreas. Essa comparação permitiu avaliar a precisão e eficácia do modelo na identificação de elementos de risco, destacando possíveis falhas ou inconsistências nas segmentações automáticas em relação aos resultados obtidos em estudos semelhantes.

Posteriormente, os resultados foram analisados de forma integrada, levando em consideração os riscos geológicos e hidrológicos da área. Foram avaliadas as potencialidades do modelo SAM para o monitoramento contínuo e a identificação precoce de riscos, assim como sua eficácia em diferentes cenários de treinamento.

A análise também buscou aplicar os resultados obtidos para o desenvolvimento de estratégias de mitigação de riscos, consolidando os achados do trabalho e propondo recomendações para o uso de modelos de segmentação de imagens em estudos de risco geológico e hidrológico de estratégias de mitigação de riscos. Essa etapa final visou

consolidar os produtos do trabalho e propor recomendações para a utilização de modelos de segmentação de imagens em estudos de risco geológico e hidrológico.

6. RESULTADOS OBTIDOS

Este capítulo apresenta os resultados das análises realizadas com diferentes modelos de segmentação para identificar elementos em áreas de risco. O modelo inicial SAM (*Segment Anything Model*) revelou limitações significativas, levando à aplicação de outros métodos, como o *MASK R-CNN*, o *Tree Segmentation* e o plugin *MapFlow*.

Os testes demonstraram variações nos resultados de acordo com os modelos e os ajustes aplicados, destacando avanços na detecção de objetos como veículos, mas dificuldades persistentes na segmentação de elementos críticos, como construções e vegetação. As observações obtidas reforçam a importância da escolha adequada de parâmetros e ferramentas para alcançar maior precisão na análise.

6.1. Modelo pré-treinado - SAM

Os resultados obtidos (Figura 6.1.1) a partir da segmentação das imagens aéreas revelam uma série de insights sobre a eficácia do modelo aplicado.



Figura 6.1.1 – Visão geral de todo *raster* segmentado. Fonte: ArcGIS Pro

Na figura 6.1.2, observa-se que alguns veículos foram identificados com sucesso, enquanto outros não foram reconhecidos. Essa discrepância pode ser atribuída a fatores como a presença de sombras que obscurecem a visibilidade dos veículos ou a iluminação direta do sol, que pode dificultar a segmentação. Ao centro da imagem, a sombra foi segmentada corretamente, mas, ao lado dela, uma sombra não foi identificada, evidenciando a variabilidade na capacidade do modelo de reconhecer elementos sob diferentes condições de iluminação.

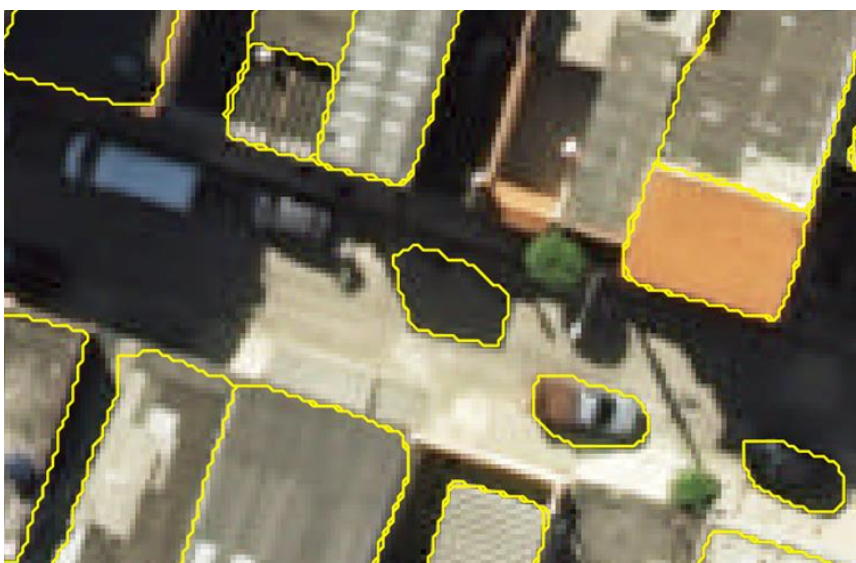


Figura 6.1.2 – Veículos segmentados e outros não. Fonte: ArcGIS Pro

Já na figura 6.1.3, a análise da mata densa mostra algumas distinções nas copas das árvores, com algumas sendo segmentadas corretamente, enquanto outras não foram identificadas. Essa limitação pode indicar a necessidade de um treinamento mais robusto para que o modelo aprenda a diferenciar melhor entre as copas das árvores e o fundo denso da vegetação.



Figura 6.1.3 – Vegetação segmentada. Fonte: ArcGIS Pro

Adicionalmente, na figura 6.1.4 é possível observar que, em um terreno baldio, alguns bancos de areia foram identificados, enquanto outros não foram reconhecidos. Essa inconsistência sugere que a segmentação pode ser influenciada pela variação na textura e na cor do solo, o que pode dificultar a identificação uniforme de todos os elementos presentes.



Figura 6.1.4 – Terreno baldio com bancos de areia segmentados. Fonte: ArcGIS Pro

Ademais, na figura 6.1.5, um erro de segmentação é evidente, onde a vegetação foi dividida erroneamente, indicando que o modelo pode ter dificuldades em lidar com áreas de vegetação complexa.



Figura 6.1.5 – Erro de segmentação em meio a vegetação. Fonte: ArcGIS Pro

Por fim, a análise da segmentação revelou que as ruas, estradas e sistemas de drenagem (figura 6.1.6) não foram identificados de forma eficaz. Essa falha é preocupante, pois esses elementos são cruciais para a gestão de águas pluviais e a mitigação de inundações. A ausência de segmentação adequada pode comprometer a avaliação dos riscos geológicos e hidrológicos.



Figura 6.1.6 – Erro de segmentação em meio a vegetação. Fonte: ArcGIS Pro

Para explorar outras abordagens de segmentação, foi testado o modelo de segmentação textual citado anteriormente - o TEXTSAM. Este método, que utiliza *prompts* de texto para guiar a segmentação também não obteve sucesso, com todos os resultados gerados sendo vazios.

6.2. Modelo pré-treinado – *Tree Segmentation*

O *Tree Segmentation* é um modelo de aprendizado profundo desenvolvido para a detecção de árvores em imagens aéreas ou capturas de drones de alta resolução. Este modelo baseia-se na arquitetura *DeepForest*, utilizando técnicas de aprendizado profundo treinadas com dados da Rede Nacional de Observatórios Ecológicos (NEON).

A detecção de árvores com o *Tree Segmentation* é particularmente útil em aplicações como manejo de vegetação, silvicultura e planejamento urbano, permitindo análises precisas em diferentes contextos espaciais e temporais. Sua capacidade de segmentar copas de árvores individuais, mesmo em ambientes com alta densidade de vegetação, torna-o ideal para estudos ambientais e urbanos. Além disso, a utilização de imagens aéreas e de drones facilita o monitoramento eficiente devido à ampla cobertura espaço-temporal oferecida por essas tecnologias (*DeepForest*, 2020).

Em um teste afim de analisar outro modelo, foi utilizado *Tree Segmentation*, aonde o processamento levou aproximadamente 1 hora e 50 minutos para segmentar 547 objetos. Entre os elementos detectados, o modelo conseguiu identificar várias árvores de diferentes cores e tamanhos, incluindo árvores isoladas e agrupadas. No entanto, também foram segmentados telhados de diversas composições e cores, além de objetos variados nas lajes das casas e alguns veículos.



Figura 6.2.1 – Primeiro teste de segmentação usando o modelo *Tree Segmentation*. Fonte: ArcGIS Pro

Observou-se que algumas árvores claramente visíveis e isoladas não foram identificadas pelo modelo, enquanto árvores em agrupamentos menos nítidos e sob sombra foram corretamente detectadas. Os parâmetros configurados para esse teste incluíram: *padding* de 100, *threshold* de 0,1, *nms overlap* de 0,1, *batch size* de 4, com *exclude pad detections* ativado, *test time augmentation* desativado e *prompt* configurado como box.



Figura 6.2.2 – Alguns objetos exemplificando o primeiro teste. Fonte: ArcGIS Pro

6.3. Outros modelos (*Mask R-CNN* e *MapFlow*)

Até o relatório de progresso, a análise de risco foi realizada com o modelo SAM (*Segment Anything Model*), que apresentou limitações significativas na segmentação. Para superar essas dificuldades, foram realizados testes com o modelo *Mask-RCNN*, que, embora tenha mostrado melhorias, ainda apresentou imprecisões. Isso levou à realização de ajustes sucessivos e à inclusão de outros modelos comparativos para avaliar suas capacidades no contexto da área de estudo.

6.3.1. Primeiros Testes e Configurações Iniciais

Inicialmente, a classificação dos objetos foi feita manualmente. Os veículos, por exemplo, foram rotulados desenhando polígonos sobre carros e caminhões presentes na imagem da área de estudo. Esses rótulos foram exportados para um primeiro teste do modelo de aprendizado profundo, treinado no ArcGIS Pro. Para esse primeiro teste, houve a necessidade de atualizar os pacotes CONDA e clonar o ambiente Py (Python) compatível com a plataforma ArcGIS, o que foi fundamental para configurar o processamento.

Esse primeiro modelo não obteve êxito. Após avaliar os parâmetros, o segundo teste incluiu ajustes no processamento, como a seleção da GPU para acelerar a execução, a redução do número de épocas de 20 para 10, a alteração do *chip size* para 512 e o tamanho do lote para 1. Com essas modificações, o tempo de processamento reduziu para aproximadamente 10 minutos, porém, o resultado ainda foi vazio.

6.3.2. Evolução dos Testes: Introdução de Novas Classes e Parâmetros

Com o objetivo de aumentar a precisão e capturar mais elementos da área de estudo, foi realizado um novo modelo com rotulagem adicional, introduzindo três classes: veículos, construções e vegetação. Esse terceiro teste utilizou os mesmos parâmetros anteriores, resultando em uma detecção parcial. O modelo conseguiu identificar aproximadamente 42 veículos, mas não foi capaz de segmentar as classes de vegetação e construções.



Figura 6.3.2.1 – Primeiro teste de segmentação de veículos usando o modelo *MASK R-CNN*. Fonte: ArcGIS Pro

Com a configuração inicial estabilizada no *MASK R-CNN*, novos ajustes foram testados para ampliar a quantidade de objetos detectados. As alterações incluíram o aumento do *padding* de 64 para 128, o *batch size* de 4 para 8, o *threshold* reduzido de 0,9 para 0,7 e o *tile size* expandido de 256 para 512. Esses ajustes resultaram em um aumento significativo nas detecções, com mais de 200 objetos identificados. Contudo, o modelo apresentou uma taxa

elevada de falsos positivos, identificando erroneamente telhados, caixas d'água e lonas plásticas nas lajes das casas como veículos e corretamente alguns veículos sob sombras.



Figura 6.3.2.2 – Segundo teste de segmentação identificando telhados e caixas d'água. Fonte: ArcGIS Pro

Para melhorar a seleção de objetos, um novo teste foi realizado, ajustando novamente o *threshold* para 0,8 e ativando a supressão máxima para evitar duplicações em objetos próximos. Esse ajuste resultou em um total de 131 veículos identificados com maior precisão, excluindo a maioria dos objetos incorretos. No entanto, alguns veículos, especialmente aqueles localizados sob sombra, deixaram de ser reconhecidos, o que indica uma sensibilidade do modelo às condições de iluminação e posicionamento dos objetos.



Figura 6.3.2.3 – Terceiro teste com resultados divergentes do anterior para os mesmos objetos. Fonte: ArcGIS Pro

6.3.3. Testes Adicionais para Outras Classes e Limitações do Modelo

Observando a variabilidade de reconhecimento entre diferentes objetos, foram realizados novos testes segmentando separadamente veículos, caixas d'água e vegetação. Os testes focados na vegetação não obtiveram êxito, reforçando os resultados anteriores em que o modelo demonstrou baixa eficácia para essa classe. Em um teste específico para construções e moradias, foram desenhados manualmente novos rótulos sobre as casas, tentando capturar a variedade de configurações arquitetônicas e características visuais. Ainda assim, o modelo, mesmo após múltiplas tentativas com variações de parâmetros iniciais e finais, gerou produtos vazios.



Figura 6.3.3.1 – Rótulos desenhados manualmente para treinar modelo para a detecção de moradias e construções no geral. Fonte: ArcGIS Pro

6.3.4. Teste Final com o Plugin *MapFlow* no QGIS

Diante das limitações observadas nas tentativas com o ArcGIS, um teste adicional foi conduzido no QGIS, utilizando o complemento - o *MapFlow*. Este plugin com uma interface mais interativa, mas com poucas possibilidades de configurações nos ajustes, também é baseado em IA para segmentação de objetos. Esta ferramenta apresentou uma capacidade

diferenciada de reconhecer construções, ruas, florestas, plantações e locais em construção (Figura 6.3.4.1).

Nos testes realizados com o *MapFlow* foi possível identificar moradias, ruas/estradas e áreas de vegetação na imagem da área de estudo sendo classificadas por cores diferentes em categorias pré-definidas do modelo – como densidade das vegetações e o tipo de construções. Nota-se também como o modelo erroneamente confundiu o córrego com uma rua (Figura 6.3.4.2).



Figura 6.3.4.1 – Teste de segmentação para vegetação, ruas e construções usando o complemento *MapFlow*.

Fonte: QGIS

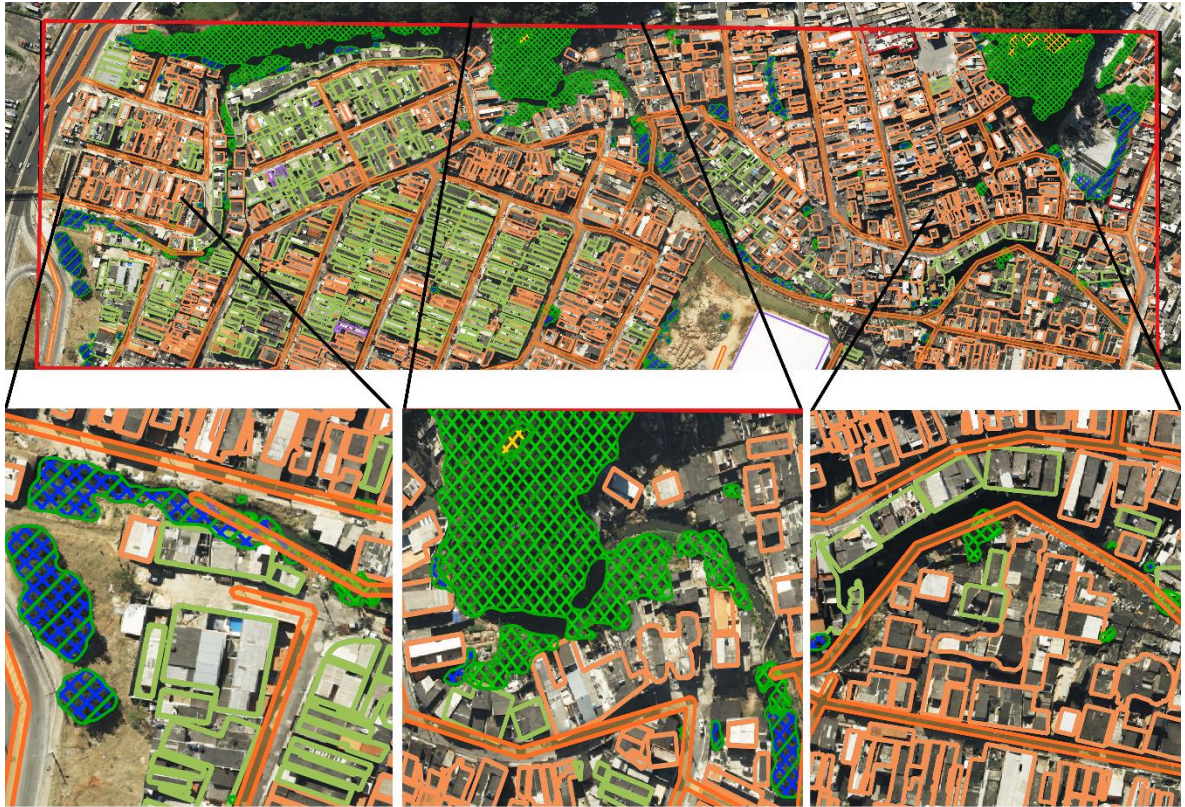


Figura 6.3.4.2 – Composição com ampliação do teste de segmentação para vegetação, ruas, construções e edifícios usando o complemento *MapFlow*. Fonte: QGIS

7. INTERPRETAÇÃO E DISCUSSÃO DOS RESULTADOS

Os resultados obtidos destacam avanços e limitações nos diferentes modelos de segmentação aplicados ao mapeamento de risco no Jardim Caboré. Inicialmente, o SAM (*Segment Anything Model*) foi utilizado para a análise de veículos, construções e vegetação. No entanto, apresentou limitações consideráveis, como dificuldade em diferenciar objetos de interesse e inconsistências na segmentação, particularmente em condições de iluminação adversas ou com objetos parcialmente visíveis.

A introdução do *Mask R-CNN* trouxe melhorias pontuais, com maior precisão na detecção de veículos e construções, mas manteve limitações ao lidar com vegetação e elementos de drenagem. Além disso, a execução prolongada, frequentemente superior a duas horas, restringiu ajustes mais aprofundados nos parâmetros. O uso do *MapFlow*, por sua vez, mostrou-se eficaz na detecção de construções e vegetação, sendo uma alternativa promissora para complementar os modelos anteriores, apesar de sua configuração limitada e dependência de um modelo pré-treinado.

As dificuldades técnicas surgidas ao longo do trabalho incluíram ajustes complexos no ambiente Python do ArcGIS, limitações no uso eficiente da GPU e altos índices de falsos

positivos em alguns testes. Esses desafios impactaram o ritmo dos experimentos, evidenciando a necessidade de ferramentas mais otimizadas e intuitivas para análises futuras. Adicionalmente, o tempo prolongado de processamento foi um fator limitante. Com tempos que ultrapassaram duas horas em alguns testes, mesmo com uma máquina de boa capacidade, o processo de ajuste contínuo dos parâmetros foi afetado, restringindo a possibilidade de explorar configurações e modelos adicionais com a profundidade desejada. Esses períodos de processamento destacam a necessidade de otimizações futuras nos modelos, para viabilizar o uso prático das ferramentas em mapeamentos de risco.

Essas limitações levaram a ajustes nos objetivos originais. A proposta inicial, que buscava a segmentação multiclasse abrangente, foi redirecionada para priorizar a detecção de elementos mais diretamente associados à ocupação urbana em áreas de risco, como construções e moradias. Esse foco mais direcionado permitiu avanços na análise, embora ainda tenha sido limitada pela necessidade de parâmetros mais ajustados e pela dificuldade de explorar todas as configurações possíveis devido ao tempo de processamento.

A segmentação desempenha um papel crucial no monitoramento de áreas dinâmicas como o Jardim Caboré, onde as mudanças frequentes no uso do solo e nas ocupações urbanas representam desafios para a gestão de riscos geológicos e hidrológicos. A identificação precisa de elementos como construções em expansão, remoção de vegetação e alterações em sistemas de drenagem pode contribuir significativamente para intervenções mais eficazes e direcionadas.

A análise dos resultados sugere que, apesar das limitações, uma abordagem combinada, utilizando o SAM para uma visão inicial, o MASK R-CNN para detecções mais refinadas e o MapFlow para validações finais, pode oferecer uma base sólida para estudos futuros. Essa combinação de ferramentas não apenas amplia a precisão das análises, mas também otimiza os recursos disponíveis, permitindo um mapeamento mais robusto e adaptado às particularidades de áreas urbanas de risco.

Além disso, a inclusão de fotos de campo como as figuras abaixo (7.1, 7.2, 7.3 e 7.4), ou até mesmo imagens aéreas, ambas de datas distintas, poderia enriquecer o monitoramento dessas áreas dinâmicas, evidenciando mudanças significativas, como obras e ocupações recentes. Esse tipo de análise comparativa pode complementar as segmentações automáticas, validando os resultados e direcionando futuras intervenções.

Este processo incluiria fotos de campo e imagens aéreas de diferentes datas para enriquecer o monitoramento de áreas dinâmicas, complementando as segmentações automáticas realizadas pelos modelos de inteligência artificial.

As fotos de campo desempenham um papel importante ao fornecer uma visão local detalhada, capturando elementos específicos que podem não ser evidentes em imagens

aéreas. Exemplos incluem sinais de instabilidade, como trincas no solo ou desníveis recentes, que são informações essenciais para validar ou corrigir as segmentações automáticas. Já as imagens aéreas, especialmente aquelas capturadas em períodos diferentes (multitemporais), permitem analisar a evolução dos riscos e alterações na paisagem ao longo do tempo. Comparar imagens de antes e depois de uma obra de canalização, por exemplo, pode revelar impactos na estabilidade do terreno ou no comportamento das margens do córrego.

A análise comparativa entre essas fontes complementa as segmentações automáticas, pois fornece um contexto mais rico para interpretar os resultados. Por exemplo, uma área identificada como vegetação pelo modelo pode ser corrigida ao se observar, em uma imagem recente ou foto de campo, que a vegetação foi removida. Essa integração de informações não apenas valida os modelos, mas também ajuda a identificar tendências, como o aumento de ocupações em áreas vulneráveis ou a degradação de barreiras naturais, como taludes e faixas de vegetação.



Figura 7.1 – Rua Ângelo Sampaio, altura do número 30 em fevereiro de 2010. Fonte: Google Street View



Figura 7.2 – Rua Ângelo Sampaio, altura do número 30 em março de 2023. Fonte: Google Street View



Figura 7.3 – Rua Ângelo Sampaio, altura do número 30 em abril de 2024. Obras de canalização já iniciadas. Fonte: Google Street View



Figura 7.4 – Rua Ângelo Sampaio, altura do número 30 ao final de outubro de 2024. Obras de canalização já avançadas. Fonte: Autor

8. CONCLUSÕES

O presente trabalho teve como objetivo empregar técnicas de inteligência artificial, especificamente o modelo *Segment Anything Model (SAM)*, para segmentar imagens aéreas de alta resolução e identificar elementos de risco geológico e hidrológico na região do Jardim Caboré, Zona Leste de São Paulo. Utilizando uma abordagem que combinou revisão bibliográfica, aplicação de modelos pré-treinados e validação integrada, buscou-se explorar o potencial desses métodos para o mapeamento e gestão de áreas de risco. As etapas envolveram análises *zero-shot* e *one-shot*, bem como testes complementares com os modelos *Mask R-CNN* e *MapFlow*, visando refinar a identificação de construções, vegetação e outros elementos críticos para a classificação de riscos.

Em termos de uso prático, os resultados obtidos têm grande potencial para embasar intervenções urbanas e políticas públicas voltadas à prevenção de enchentes e deslizamentos. A detecção de construções, majoritariamente irregulares, auxilia no planejamento urbano, especialmente em locais onde a vegetação tem sido substituída por essas edificações, aumentando a vulnerabilidade da área a eventos hidrológicos e geológicos

– escorregamento, solapamento e erosão. Apesar das limitações observadas, como segmentação imprecisa em condições adversas e identificação parcial de elementos como vegetação, os métodos empregados demonstraram viabilidade para análise de ocupação urbana em áreas vulneráveis. A integração das técnicas empregadas mostrou-se promissora para o mapeamento de áreas de risco, consolidando-se como uma abordagem que pode ser ampliada e refinada em aplicações futuras.

Para o aprimoramento das metodologias de análise e mapeamento de riscos, recomenda-se explorar o uso de fotos oblíquas em estudos futuros. Esse tipo de imagem, que capta os elementos construtivos de diferentes ângulos, apresenta um potencial significativo para áreas com alta declividade, onde as perspectivas aéreas convencionais podem não capturar detalhes essenciais. Além disso, estudos adicionais poderiam incluir modelos pré-treinados mais robustos, análise de imagens multitemporais para compreender a evolução dos riscos e a implementação de infraestrutura computacional otimizada para processamento em tempo reduzido. Esses avanços podem proporcionar maior precisão e agilidade na identificação de áreas críticas, subsidiando ações de prevenção e mitigação mais eficazes.

9. REFERÊNCIAS BIBLIOGRÁFICAS

- Coordenação Municipal de Defesa Civil. Relatório HIQ-06 São João - Caboré - 02/05/2022. São Paulo: COMDEC, 2022.
- Coordenação Municipal de Defesa Civil. Relatório HIQ-07 Angelo Sampaio - 02/05/2022. São Paulo: COMDEC, 2022.
- Coordenação Municipal de Defesa Civil. Relatório HSM-01 Caboré – 19/04/2022. São Paulo: COMDEC, 2022.
- Coordenação Municipal de Defesa Civil. Relatório HSM-02 São João – 19/04/2022. São Paulo: COMDEC, 2022.
- Coordenação Municipal de Defesa Civil. Relatório SM-22 – 19/07/2023. São Paulo: COMDEC, 2023.
- Coordenação Municipal de Defesa Civil. Relatório IQ-13 – 21/06/2023. São Paulo: COMDEC, 2023.
- DeepForest. Deep Learning for Individual Tree Detection. National Ecological Observatory Network (NEON). 2020. <https://deepforest.readthedocs.io> (Acessado em Novembro 2024).
- Fernandes, A. J. 1991. O Complexo Embu no leste do Estado de São Paulo: contribuição ao conhecimento da litoestratigrafia e da evolução estrutural e Metamórfica. Instituto Geológico de São Paulo: <https://www.infraestruturameioambiente.sp.gov.br/institutogeologico/1991/12/o-complexo-embu-no-leste-do-estado-de-sao-paulo-contribuicao-ao-conhecimento-da-litoestratigrafia-e-da-evolucao-estrutural-e-Metamorfica/> (Acessado em Março 2024).
- Freitas, A. L. S.; Corsi, A. C.; Sandre, L. H. Análise temporal da evolução da ocupação em duas áreas suscetíveis a movimentos gravitacionais de massa e inundação na cidade de São Paulo. 17º Congresso Brasileiro de Geologia de Engenharia e Ambiental: https://schenautomacao.com.br/cbge2022/envio/files/trabalho1_114.pdf (Acessado em Março 2024).
- Giarratano, J. C., & Riley, G. Expert Systems: Principles and Programming. Cengage Learning, 2005.
- Goodfellow, I., Bengio, Y., & Courville, A. 2016. Deep Learning. Cambridge, MA: The MIT Press.
- Instituto Nacional de Pesquisas Espaciais (INPE). 2010. Estudo aponta que enchentes e deslizamentos serão mais frequentes na capital paulista. INPE - Notícias, 15 de junho de 2010: [http://www.inpe.br/noticias/noticia.php?Cod_Noticia=2215#:~:text=Estudos%20preliminares%20do%20Instituto%20Nacional,10%20mil%C3%ADmetros\)%20na%20capital%20paulista.](http://www.inpe.br/noticias/noticia.php?Cod_Noticia=2215#:~:text=Estudos%20preliminares%20do%20Instituto%20Nacional,10%20mil%C3%ADmetros)%20na%20capital%20paulista.) (Acessado em Abril 2024).
- He, K., Gkioxari, G., Dollar, P., Girshick, R. *Mask R-CNN*. 2018. DOI: <https://doi.org/10.48550/arXiv.1703.06870>
- Jurafsky, D.; Martin, J. H. *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*. 2nd ed. Upper Saddle River: Prentice Hall, 2009.
- Kirillov, A., Berg, A., Rolland, C., Mintun, E., Mao, H., Gustafson, L., Ravi, N., Dollar, P., Girshick, R., Whitehead, S., Lo, W., Xiao, T. *Segment Anything* (2023). https://scontent.fcgh68-1.fna.fbcdn.net/v/t39.2365-6/10000000_900554171201033_1602411987825904100_n.pdf?_nc_cat=100&ccb=1-7&_nc_sid=3c67a6&_nc_ohc=e7bYPRaiQscQ7kNvgFrR57U&_nc_zt=14&_nc_ht=sconte

- nt.fcgh68-1.fna&_nc_gid=A6mspf-BHAszrnYsQSm2z8-
&oh=00_AYAZICVu6Ty1jceo2MP56ixmNn4LF8Y2ah2vuJ4I9d--bw&oe=673E9DA7.
(Acessado em Abril 2024).
- Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., ... & Girshick, R. (2023). Segment anything. In Proceedings of the IEEE/CVF International Conference on Computer Vision (pp. 4015-4026). DOI: <https://doi.org/10.48550/arXiv.2304.02643>
- Lin, Y.; Mair, M., Gupta, S., Belongie, S.; Hariharan. Microsoft COCO: Common Objects in Context. In: European Conference on Computer Vision (ECCV). 2014. p. 740-755.
- MapFlow. Plataforma de extração de dados geoespaciais. <https://docs.mapflow.ai>. (Acessado em Outubro 2024).
- Meta AI. Introducing SAM 2: The next generation of Meta Segment Anything Model for videos and images (2024). <https://ai.Meta.com/blog/segment-anything-2/>. (Acessado em Julho 2024).
- Ministério das Cidades / Instituto de Pesquisas Tecnológicas – IPT. Mapeamento de Riscos em Encostas e Margem de Rios. Celso Santos Carvalho, Eduardo Soares de Macedo e Agostinho Tadashi Ogura, organizadores – Brasília: Ministerio das Cidades; Instituto de Pesquisas Tecnológicas – IPT, 2007.
- Mitchell, T. M. Machine Learning. New York: McGraw-Hill, 1997.
- Oscó, L. P., Wu, Q., Lemos, E. L. de, Gonçalves, W. N., Ramos, A. P. M., Li, J., & Marcato Junior, J. 2023. The Segment Anything Model (SAM) for Remote Sensing Applications: From Zero to One Shot.
- Prefeitura do Município de São Paulo. Secretaria Municipal de Infraestrutura Urbana e Obras. (2022). Caderno de Bacia Hidrográfica: Bacia do Rio Aricanduva (2a ed.). São Paulo: https://www.prefeitura.sp.gov.br/cidade/secretarias/upload/obras/cadernos_de_drenagem/CBH_2022_Aricanduva_ed2.pdf (Acessado em Março 2024).
- Prefeitura do Município de São Paulo (PMSP). (2023). Mapeamento de risco de enchentes e inundações em assentamentos precários do Município de São Paulo. Relatório técnico. São Mateus, São Paulo: Divisão de Prevenção (DPREV), Coordenação Municipal de Defesa Civil (COMDEC), Secretaria Municipal de Segurança Urbana (SMSU).
- Prefeitura do Município de São Paulo (PMSP). (2023). Mapeamento de risco de enchentes e inundações em assentamentos precários do Município de São Paulo. Relatório técnico. Subprefeitura: Itaquera. UNIDADE RESPONSÁVEL: Divisão de Prevenção (DPREV), Coordenação Municipal de Defesa Civil (COMDEC), Secretaria Municipal de Segurança Urbana (SMSU).
- Prefeitura Municipal de São Paulo. 2023. INFORMES urbanos: Nº 56 - Janeiro 2023: Áreas de risco geológico e o Plano Diretor Estratégico do Município de São Paulo: https://gestaourbana.prefeitura.sp.gov.br/wp-content/uploads/2023/02/56_IU_AREAS-DE-RISCO.pdf (Acessado em Março 2024).
- Prefeitura Municipal de São Paulo. 2023. Com investimento de R\$ 60,7 milhões, obras contra enchentes beneficiam mais de mil moradores da Comunidade Caboré. Prefeitura Municipal de São Paulo - Secretaria de Obras, 04 de agosto de 2023. <https://www.prefeitura.sp.gov.br/cidade/secretarias/obras/noticias/?p=352107> (Acessado em Março 2024).
- Ren, S., He, K.; Girschick, R., Sun, J. Faster R-CNN: Towards real-time object detection with region proposal networks. In: Advances in Neural Information Processing Systems. 2015. p. 91-99.

- Rinaldi, L. (03/02/2023). Vídeos: Temporal provoca alagamentos, cobre ambulância, desaba casa e deixa moradores com água na cintura em SP. TV Globo e G1-SP.
- Russel, S., & Norvig, P. 2014. Inteligência Artificial: Uma Abordagem de Aprendizado de Máquina. Rio de Janeiro: Elsevier.
- Siciliano, B.; Khatib, O. (Eds.). Springer Handbook of Robotics. 2nd ed. Berlin: Springer, 2016.
- Szeliski, R.; Shum, H. Y.; Szummer, M.; Freeman, W. T.; Sawhney, H. S.; Leonardis, A. Vision Algorithms: Theory And Practice. Berlin: Springer, 2010.
- Tuteja, P., Singh, R., Pathak, S. Text SAM: Extracting GIS Features Using Text Prompts. Meta's Segment Anything Model and Grounding Dino. <https://link.com>. (Acessado em Setembro 2024).
- UN-Habitat. 2022. World Cities Report 2022: https://unhabitat.org/sites/default/files/2022/06/wcr_2022.pdf (Acessado em Março 2024).
- Viso AI. (s.d.). Segment Anything Model (SAM) Explained: <https://viso.ai/deep-learning/segment-anything-model-sam-explained/> (Acessado em Março 2024).