



RIO

PUC

Tese de Doutorado

Análise de imagens satelitais usando deep learning para planejamento de operações militares

Raphael de Souza e Almeida

Pontifícia Universidade Católica do Rio de Janeiro
Centro Técnico Científico
Departamento de Informática

Rio de Janeiro, outubro de 2025



Pontifícia
Universidade
Católica do
Rio de Janeiro

Tese de Doutorado

Análise de imagens satelitais usando deep learning para planejamento de operações militares

Raphael de Souza e Almeida

Orientação: Professor Alberto Barbosa Raposo

Tese apresentada como requisito parcial para a obtenção do grau de
Doutor em Ciências - Informática pelo programa de Pós-Graduação em
Informática, no Departamento de Informática da PUC-Rio

Rio de Janeiro, 07 de outubro de 2025



Pontifícia
Universidade
Católica do
Rio de Janeiro

Análise de imagens satelitais usando deep learning para planejamento de operações militares

Raphael de Souza e Almeida

Tese apresentada como requisito parcial para a obtenção do grau de Doutor em Ciências – Informática pelo Programa de Pós-Graduação em Informática. Aprovada pela Comissão examinadora abaixo:

Professor Alberto Barbosa Raposo

Orientador

Departamento de Informática – PUC-Rio

Professor Manuel Eduardo Loaiza Fernandez

Universidad Católica San Pablo - UCSP

Professor Leonardo Cardia da Cruz

Área de Computação e Big Data – SENAI-CIMATEC

Professor Paulo Ivson Netto Santos

Departamento de Informática – PUC-Rio

Cesar Augusto Sierra Franco

Instituto TecGraf – PUC-Rio

Professor Thiago Malheiros Porcino

Laboratório Nacional de Computação Científica – LNCC

Rio de Janeiro, 07 de outubro de 2025



Pontifícia
Universidade
Católica do
Rio de Janeiro

Todos os direitos reservados. A reprodução, total ou parcial, do trabalho é proibida sem autorização da universidade, da autora e do orientador.

Raphael de Souza e Almeida

Graduou-se em Ciências Navais pela Escola Naval (2006) e mestrado em Ciência da Computação pela Pontifícia Universidade Católica do Rio de Janeiro (2020).

Ficha Catalográfica

Almeida, Raphael de Souza e

Análise de imagens satelitais usando deep learning para planejamento de operações militares / Raphael de Souza e Almeida; orientador: Alberto Barbosa Raposo. – Rio de Janeiro: PUC-Rio, Departamento de Informática, 2025.

v., 147 f: il. color. ; 30 cm

Tese (doutorado)- Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Informática.

Inclui bibliografia

1. Informática – Teses. 2. Sensoriamento Remoto;. 3. Ambiente de aprendizagem;. 4. Planejamento Militar;. 5. Segmentação Semântica.. I. Raposo, Alberto Barbosa . II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Informática. III. Título.

CDD: 004

À minha mãe, pelo seu exemplo em vida de bondade e dedicação.

Agradecimentos

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

Resumo

Almeida, Raphael de Souza e ; Raposo, Alberto Barbosa . **Análise de imagens satelitais usando deep learning para planejamento de operações militares**. Rio de Janeiro, 2025. 147p. Tese de Doutorado – Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

O planejamento de uma operação militar é difícil e exige uma série de procedimentos e coordenações para se analisar e interpretar todos os dados disponíveis. A este processo chamamos de Processo de Planejamento Militar. Nele, obter informações detalhadas da área de operações é uma atividade complexa, demorada e essencial para viabilizar sua condução. Além disso, entender os aspectos táticos do terreno e transmitir a análise realizada encontram como barreira a abstração necessária para tal. Outra dificuldade está em se testar o planejamento confeccionado na mesma região analisada. Sendo assim, este trabalho apresenta um pipeline baseado em técnicas modernas de deep learning para segmentação semântica de imagens de sensoriamento remoto, com o propósito é automatizar e aprimorar o processo de geração de cartas de trafegabilidade, terrenos virtuais e ambientes simulados para apoio ao planejamento de operações militares. Nossa abordagem consiste no treinamento de uma arquitetura baseada em redes neurais UNet, utilizando segmentação semântica binária com um método comitê, a fim de detectar classes relacionadas à trafegabilidade. Para analisar o terreno virtual, propomos o Simulador Virtual para Estudo Topotático do Terreno (SVETT), que utiliza recursos encontrados em sala de aula para facilitar o estudo do terreno. Como resultados, os dados obtidos das predições realizadas pela arquitetura nos mostraram resultados satisfatórios se comparados com classificadores multiclasse, apresentando um IoU de 0.7259 na predição de algumas das classes. Referente ao SVETT, nós o avaliamos com 54 alunos da Escola Naval. Os resultados sugerem que os alunos se sentiram mais seguros participando do Exercício no Terreno tendo recebido a instrução antes no simulador, com média de 4,69 e desvio padrão de 0.53, em uma escala Likert de 5 pontos.

Palavras-chave

Sensoriamento Remoto; Ambiente de aprendizagem; Planejamento Militar; Segmentação Semântica.

Abstract

Almeida, Raphael de Souza e ; Raposo, Alberto Barbosa (Advisor). **Analysing satellite images using deep learning for planning military operations**. Rio de Janeiro, 2025. 147p. Tese de Doutorado – Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

Planning military operations is difficult and requires a series of procedures and coordinations to analyze and interpret all available data. We call this process the Military Planning Process. During this process, obtaining detailed information from the area of operations is a complex, time-consuming and essential activity to enable its conduct. In addition, understanding the tactical aspects of the terrain and transmitting the analysis carried out find as a barrier the abstraction necessary for this. Another difficulty is in testing the planning made in the same region analyzed. Thus, this work presents a pipeline based on modern deep learning techniques for semantic segmentation of remote sensing images, with the purpose to automate and improve the process of generating trafficability letters, virtual terrains and simulated environments to support the planning of military operations. Our approach consists of training an architecture based on UNet neural networks, using binary semantic segmentation with an ensemble method, in order to detect classes related to trafficability. To analyze the virtual terrain, we propose the Virtual Simulator for Topotactic Land Study (SVETT), which uses resources found in the classroom to facilitate the study of the terrain. As a result, the data obtained from the predictions made by the architecture showed us satisfactory results compared to multiclass classifiers, showing an IoU of 0.7259 in the prediction of some of the classes. Regarding SVETT, we evaluated it with 54 students from the Naval School. The results suggest that the students felt safer participating in the Field Exercise having received the instruction before in the simulator, with an average of 4.69 and standard deviation of 0.53, on a 5 degree Likert scale.

Keywords

Remote Sensing; Learning Environment; Military Planning; Semantic Segmentation.

Sumário

1	Introdução	18
1.1	O problema de pesquisa	23
1.2	Justificativa para o problema de pesquisa	24
1.3	As Hipóteses de pesquisa	25
1.4	O objetivo deste trabalho	26
1.5	Contribuições	27
2	Referencial teórico	29
2.1	Sobre o Corpo de Fuzileiros Navais	29
2.2	Do Processo de Planejamento Militar	32
2.3	Questionário aplicado aos instrutores do C-ApA-CFN	39
2.4	O uso da simulação nos treinamentos e planejamentos militares	43
2.5	Uso de sensoriamento remoto no estudo do terreno	47
2.6	Inteligência Artificial	48
2.7	Deep Learning no sensoriamento remoto	56
3	Trabalhos Relacionados	61
4	Segmentação Semântica em imagens de Sensoriamento Remoto	78
4.1	Sobre a busca pelo dataset	78
4.2	Experimento 01 - Usando Feature Pyramid Network em imagens satelitais Sentinel 2	81
4.3	Experimento 02 - Usando UNet em 3 classes do banco de dados Deepglobe	84
4.4	Experimento 03 - Usando UNet em 6 classes do banco de dados Deepglobe	86
4.5	Experimento 04: o experimento proposto por este trabalho	87
4.6	Metodologia para o treinamento do experimento 4	89
4.7	Resultados obtidos com o treinamento do modelo de IA	90
5	Simulador Virtual para Estudo Topotático do terreno - SVETT	96
5.1	Interfaces do SVETT	98
5.2	Metodologia de experimento com o SVETT	101
5.3	Resultados obtidos com o SVETT	103
6	Criação da Carta de Trafegabilidade e terreno para o SJD	112
6.1	Processo manual de criação do terreno para o SJD e a utilização da solução proposta	112
6.2	Uso das imagens concatenadas como Carta de Trafegabilidade	114
7	Conclusões	120
8	Referências bibliográficas	124

A	Respostas do questionário qualitativo e quantitativo aplicado aos instrutores antigos e atuais do C-ApA-CFN	134
B	Respostas do questionário qualitativo e quantitativo aplicado a Oficiais	139

Lista de figuras

Figura 1.1	Exemplo de CT.	20
Figura 1.2	Pipeline proposto como objetivo deste trabalho.	26
Figura 2.1	Emprego do Grupamento Operativo de Fuzileiros Navais em Apoio à Defesa Civil. Fonte: Portal BIDS, disponível em https://portalbids.com.br/2024/02/26/marinha-mobiliza-grupamento-de-fuzileiros-navais-para-apoiar-vitimas-de-temporal-no-rj/	31
Figura 2.2	O Desembarque anfíbio tem por finalidade projetar poder, a partir de uma Força Naval, para um território hostil ou potencialmente hostil. Fonte: Livro Branco de Defesa Nacional, ano 2012 [8]	32
Figura 2.3	Sequência de ações a serem realizadas no planejamento de uma operação militar, destacando a etapa onde se faz necessário obter informações sobre o terreno. Não é o foco deste trabalho esgotar o assunto, sendo assim, estas etapas não serão explicadas aqui.	34
Figura 2.4	Comparativo entre imagens. A esquerda observamos uma imagem satélite obtida por meio de fontes abertas. A direita, uma carta de trafegabilidade mostrando as restrições de deslocamento da imagem a esquerda. Cabe destacar que este tipo de carta de trafegabilidade é utilizado no SJD.	35
Figura 2.5	Sistema de Jogos Didáticos exibindo simbologias das forças amigas, em azul, e inimigas, em vermelho.	45
Figura 2.6	À esquerda, um exemplo de CT usada no SJD. À direita, em cima, legendas das ações executadas pelas peças de manobra. À direita, em baixo, as trafegabilidades existentes na região da carta.	47
Figura 2.7	O espectro eletromagnético. Fonte: http://imagers.gsfc.nasa.gov/ems/waves3.html .	48
Figura 2.8	Neurônio encontrado no cérebro humano. Fonte: https://escolakids.uol.com.br/ciencias/neuronios.htm .	50
Figura 2.9	Neurônio artificial. Fonte: https://medium.com/@avinicius.adorno/redes-neurais-artificiais-418a34ea1a39 .	51
Figura 2.10	Estruturas de redes neurais artificiais. Em a) uma rede de 1 camada, em b) uma rede multi-camada e em c) uma rede recorrente [51].	52
Figura 2.11	Deep Learning e Machine Learning são ramos específicos dentro da Inteligência Artificial [54].	54
Figura 2.12	Arquitetura U-net (exemplo para 32x32 pixels na resolução mais baixa). Cada caixa azul corresponde a um mapa de recursos multicanal. O número de canais é indicado na parte superior da caixa. O tamanho x-y é fornecido na borda inferior esquerda da caixa. As caixas brancas representam mapas de recursos copiados. As setas indicam as diferentes operações [60].	54
Figura 2.13	Comparação entre parte de uma carta de trafegabilidade e uma imagem satélite. As áreas urbanas, estradas e rios correspondem à trafegabilidades específicas segundo as cores estabelecidas.	57
Figura 2.14	Processo de convolução, com filtro de 3x3, aplicado em uma matriz. Fonte: https://www.kaggle.com/code/akshitsharma1/unet-architecture-explained-in-one-shot-tutorial .	59

Figura 2.15	Processo de convolução em 3 camadas consecutivas, com filtro horizontal de 3x3, aplicado à imagem de um gato. Fonte: imagem gerada por Inteligência Artificial.	59
Figura 2.16	Em a), a imagem representa o processo de pooling em 3 camadas consecutivas aplicado à imagem de um gato. Nele, a resolução é reduzida (pooling) preservando os elementos importantes. Já em b), a imagem representa o processo de pooling usando uma região de 2x2 pixels. Fonte: imagens do gato geradas por Inteligência Artificial.	60
Figura 3.1	Framework referente ao método de Comitê de modelos segundo Kumar et al. [73].	66
Figura 3.2	Gráfico relativo à teoria do fluxo.	71
Figura 4.1	Faixas do espectro eletromagnético exibindo os respectivos comprimentos de ondas. A região identificada como “visible light” identifica a faixa contendo as cores visíveis. Já a região identificada como “NIR region”, descreve a região contendo a faixa do NIR. Fonte: https://bbpsales.com/near-infrared-measurements-how-do-they-work/ .	82
Figura 4.2	À esquerda, imagem concatenada para ser usada como dado de entrada no modelo de IA. À direita, a anotação referente à esta imagem.	83
Figura 4.3	Em a), observamos uma imagem e sua respectiva máscara do banco de dados Land Cover Classification Challenge. Em b) observamos a imagem e a máscara do banco Road Extraction Challenge.	87
Figura 4.4	Fluxograma para treinamento da segmentação.	90
Figura 4.5	Predição realizada com a imagem do Experimento 3 contendo 06 classes.	93
Figura 4.6	Comparação de uma mesma imagem prevista pelos 3 tipos diferentes de concatenação existentes neste trabalho.	93
Figura 5.1	Interfaces do SVETT. Nela, os 3 tipos de equipamentos interagem dentro do terreno virtualmente criado	98
Figura 5.2	Interface de projeção mostrando a mesma representação esboçada na Interface Construtiva na Figura 5.3.	99
Figura 5.3	Interface construtiva mostrando uma representação da interface do projetor em amarelo no centro da imagem e alguns esboços no terreno.	100
Figura 5.4	Interface de realidade virtual apontando para uma posição inimiga representada no terreno.	101
Figura 5.5	Cenário “sala de aula”.	102
Figura 5.6	Questionário 2, utilizado para avaliar o uso da ferramenta nas instruções.	106
Figura 5.7	Média e desvio padrão das respostas do subitem “Satisfação com a aprendizagem atual” do questionário “Student Satisfaction and Self-Confidence in Learning” com escala Likert de 5 pontos.	107
Figura 5.8	Média e desvio padrão das respostas do subitem “autoconfiança na aprendizagem” do questionário “Student Satisfaction and Self-Confidence in Learning” com escala Likert de 5 pontos.	110
Figura 5.9	Questionário 3, que avalia o processo de aprendizagem iniciando em sala com a teoria, efetuando o uso do SVETT e participando do ET.	111

Figura 6.1	A direita observamos a imagem satélite original e, a direita, a imagem criada para uso como trafegabilidade no SJD.	114
Figura 6.2	Respostas sobre o posto dos planejadores, onde obtemos o tempo de experiência dentro das Forças Armadas.	115
Figura 6.3	Respostas sobre a quantidades de dias de manobra e exercício, que denota experiência em treinamentos e operações militares.	116
Figura 6.4	Respostas sobre a quantidades planejamentos, que significa que os militares possuem experiencia em realizar o PPM.	116

Lista de tabelas

Tabela 2.1	Gastos necessários para conduzir um grupo contendo 30 OfA e 22 militares da equipe de instrução.	39
Tabela 3.1	Resultados dos modelos considerando Accuracy, Precision, Recall, F-score, e mIoU.	62
Tabela 3.2	Desempenho da classificação com relação à quantidade de dados em pequena escala.	65
Tabela 3.3	Os desempenhos em dados de pequena escala	66
Tabela 3.4	Resultados de mIoU para o conjunto de dados de classificação LCLU com o banco de dados DeepGlobe.	69
Tabela 3.5	Resumo dos trabalhos relacionados neste capítulo	77
Tabela 4.1	Resultados das predições realizadas com o modelo do experimento 3 contendo 6 classes.	92
Tabela 4.2	Resultados obtidos nas diversas classes do dataset DeepGlobe através da abordagem da metodologia de Comitê de modelos, utilizando segmentações binárias.	92
Tabela 4.3	Resultados de mIoU alcançados por Desai e Ghose, em 2022, com treinamentos realizados com o Dataset DeepGlobe Land Cover.	92
Tabela 5.1	Perguntas usadas no questionário 2.	105
Tabela 5.2	Questões do subitem “Satisfação com a aprendizagem atual” do questionário “Student Satisfaction and Self-Confidence in Learning”.	105
Tabela 5.3	Questões do subitem “autoconfiança na aprendizagem” do questionário “Student Satisfaction and Self-Confidence in Learning”.	108
Tabela 5.4	Perguntas do questionário 3.	109
Tabela 6.1	Trafegabilidades existentes no Sistema de Jogos Didáticos (SJD).	113

Lista de siglas e abreviaturas

AOp – Área de Operações

ASA – Ambiente Sintético de Aprendizagem

C-ApA-CFN – Cursos de Aperfeiçoamento Avançado de Oficiais do Corpo de Fuzileiros Navais

C-EMOI – Curso de Estado-Maior para Oficiais Intermediários

CGAN – Conditional Generative Adversarial Networks

CSimCFN – Centro de Simulação do Corpo de Fuzileiros Navais

CM – Carta Militar

CT – Carta de Trafegabilidade

DL – Deep Learning

EATT – Estudo dos Aspectos Topotáticos do Terreno

EM – Estado-Maior

EPI – Estimativa Preliminar de Inteligência

ET – Exercício no Terreno

GH – Giro do Horizonte

GLO – Garantia da Lei e da Ordem

FN – False Negative

FP – False Positive

FFE – Força de Fuzileiros da Esquadra

FPN – Feature Pyramid Network

HP – Hipóteses de Pesquisa

IA – Inteligência Artificial

IoT – Internet of Things

IoU – Intersection over Union

IDMC – Centro de Monitoramento de Deslocamentos Internos

LA – Linhas de Ação

LCLU – Land Cover & Land Use

MDE – Modelo Digital de Elevação

MINUSTAH – Missão de Paz das Nações Unidas para a estabilização do Haiti

ML – Machine Learning

NIR – Infravermelho Próximo

OfA – Oficiais-Alunos

OA – Overall Accuracy PE – Ponto Estação

PO – Pontos de Observação

PPM – Processo de Planejamento Militar

PUC-Rio – Pontifícia Universidade Católica do Rio de Janeiro

RAS – Radar de Abertura Sintética

RNA – Rede Neural Artificial

RNC – Redes Neurais Convolucionais

RUNet – Robust UNet

RV – Realidade Virtual

SACEM – Sequência de Ações de Comando e Estado-Maior

SAE – Squeeze-and-Excitation

SANT – Sistemas Aéreos não Tripulados

SE – Sinais de Entrada

SGD – Stochastic Gradient Descent

SJD – Sistema de Jogos Didáticos

SR – Sensoriamento Remoto

SVETT – Simulador Virtual para Estudo Topotático do Terreno

TSR – Técnicas de Sensoriamento Remoto

TN – True Negative

TP – True Positive

UNIFIL – Força Interina das Nações Unidas no Líbano

“A persistência é o caminho do êxito.”

Charles Chaplin, *Vida e Pensamentos*.

1

Introdução

Vivemos em um mundo onde as organizações precisam manter sua produtividade e vantagem competitiva. A agilidade e a adaptabilidade tornaram-se habilidades cada vez mais necessárias para que os líderes sejam bem-sucedidos. Essas mudanças bruscas forçam as empresas a frequentemente se reinventarem, mudando o comportamento e as necessidades por capacitações e novas tecnologias [1].

Neste contexto, o acrônimo em inglês VUCA, definido pela Volatilidade, Incerteza, Complexidade e Ambiguidade, surge em meio à globalização e desenvolvimentos no âmbito da tecnologia da informação [2]. Este tipo de ambiente incentiva pessoas a aprenderem a lidar com vários cenários socioeconômicos [3]. Porém, não somente no mundo empresarial, mas também no âmbito das Forças Armadas, observamos constantes corridas tecnológicas.

No mundo militar, observamos uma eterna busca por novas tecnologias voltadas à Defesa Nacional. Esta procura visa fazer frente ao constante surgimento de novas ameaças, sejam elas na doutrina ou em equipamentos. O Escritório de Pesquisas Navais (do inglês Office of Naval Research, sigla ONR), vem se destacando por coordenar pesquisas junto à laboratórios pelo mundo todo, com foco em projetar os Fuzileiros Navais Americanos como uma das forças mais sofisticadas tecnologicamente [4]. Estas buscas baseiam-se em uma diretiva, publicada em 2023, chamada “Force Design 2030”, com metas estabelecidas para o respectivo ano [5].

Fruto destas pesquisas, podemos citar como exemplo o conceito de “Arms Room”. Nele, um militar, antes especializado em apenas um meio ou equipamento, com foco em um tipo específico de tarefa, agora precisa se capacitar em diversos outros sistemas. A intenção é permitir maior flexibilidade para o cumprimento de diferentes missões por uma única Unidade, bastando apenas “escolher o equipamento necessário para a missão” (tradução nossa) [6].

Em nosso país, a Marinha do Brasil estabelece em sua Doutrina Militar Naval a seguinte missão:

“Preparar e empregar o Poder Naval, a fim de contribuir para a defesa da Pátria; estar pronta para atuar na garantia dos poderes constitucionais e, por iniciativa de qualquer destes, da lei e da ordem; atuar em ações sob a égide de organismos internacionais e em apoio à política externa do País; e cumprir as atribuições

subsidiárias previstas em Lei, com ênfase naquelas relacionadas à Autoridade Marítima, a fim de contribuir para a salvaguarda dos interesses nacionais” [7].

O cumprimento desta grande quantidade de tarefas é possível graças à amplitude de ambientes operacionais ¹ em que esta Força Singular é capaz de operar. De forma a atender às diferentes tarefas citadas, um Comando subordinado específico, com meios e pessoal preparados para tal, recebe a responsabilidade de planejar e conduzir as ações.

As operações terrestres cabem à Força de Fuzileiros da Esquadra (FFE) [8], que é composta por militares do Corpo de Fuzileiros Navais (CFN), tropa permanentemente em condição de pronto emprego para assegurar sua capacidade de projeção de poder do mar para a terra [9].

Sendo assim, em que pese a grande quantidade de ambientes operacionais citados e, considerando a necessidade dos militares estarem em constante prontidão para uma rápida resposta às diferentes ameaças, um planejamento rápido e detalhado se faz necessário. Para isso, seu preparo é feito por pessoal altamente treinado e capaz de interpretar todos os dados disponíveis de forma rápida e precisa.

Neste sentido, a maneira padronizada para se analisar e escriturar todos estes detalhamentos é através do Processo de Planejamento Militar (PPM) [10]. Para tal, existem várias etapas e fases a serem cumpridas, que serão detalhadas no Capítulo 2. Neste contexto, destacam-se em importância a etapa da “Orientação inicial do Estado-Maior” e a do “exame de situação do Comandante”.

Na “Orientação inicial do Estado-Maior”, os planejadores receberão informações iniciais importantes sobre os aspectos atinentes à missão a ser cumprida. Este momento cresce em importância à medida que são recebidas as primeiras informações para iniciar os trabalhos, e também pelo fato de que todo o planejamento será realizado com base nas informações recebidas neste momento.

Entre os dados e análises recebidos, destacamos a Estimativa Preliminar de Inteligência (EPI), que aborda assuntos como forças adversas, condições climáticas e diversos outros. Uma das principais análises recebidas nesta Estimativa são as características da área de operações, que abordam o Estudo dos Aspectos Topotáticos do Terreno (EATT) [11].

Entendemos como aspectos topotáticos a união entre as características topológicas, que consistem no estudo das conformações dos relevos de uma

¹São ambientes operacionais: espaço aéreo, superfície das águas, abaixo da superfície das águas e terrestre [8].

determinada região, incluindo sua cobertura (vegetal, rios, construções), e os aspectos táticos, que são estudos voltados à utilização desse relevo. Este estudo visa aproveitar as características da área de operações, para se obter sucesso no cumprimento de uma missão com o mínimo de danos colaterais possível [12].

Como parte do resultado do EATT, um documento conhecido como carta de trafegabilidade (CT) é gerado. Ele consiste na tradução do conjunto de informações dispostas, tais como cartas militares, imagens de satélites ou fotografias aéreas de uma determinada área, no tipo de mobilidade possível em se deslocar por esta área em questão [14].

Sua importância está no fato dela orientar quais tipos de meios, mobilidade e, em consequência disso, qual a velocidade e condições que a tropa terá para avançar em direção aos seus objetivos. Além disso, permite levantar regiões importantes que podem canalizar, dissociar ou impedir o movimento da tropa durante a operação, dificultando, assim, o cumprimento da missão.

Observamos uma CT feita de forma manual na Figura 1.1. Este documento é bastante demorado para se confeccionar à medida que deve contemplar toda uma área de operações, que pode conter desde 60 km², a depender do tipo de operação.

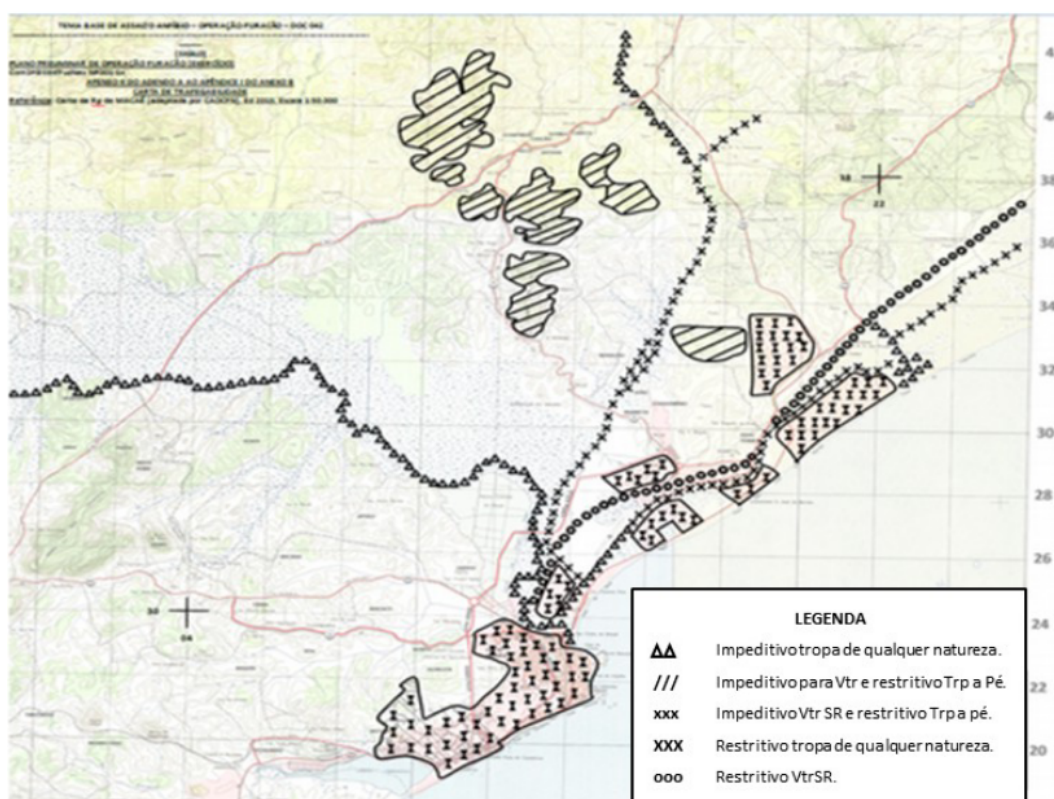


Figura 1.1: Exemplo de CT.

Tanto a EPI como a CT são muito importantes à medida que analisam

as características da região. Por exemplo, elas permitem mostrar se o uso de viaturas é recomendado em uma determinada ação que se pretende realizar, dada a instabilidade na área[15]. É de posse destas informações que serão definidos limites de atuação entre as peças de manobra (Pelotões e Companhias), os locais onde serão montadas as instalações logísticas, entre outros.

Porém, realizar este tipo de análise é muito difícil quando estamos em um ambiente fechado, valendo-se apenas de cartas militares e imagens aéreas. Para isso, quando possível, os planejadores deslocam-se para as regiões onde se pretende realizar as ações, ocupando os pontos mais altos dessas áreas, de forma a se obter melhor observação da região.

Posicionado na elevação aonde se pretende fazer o estudo, é realizado inicialmente um reconhecimento conhecido como Giro do Horizonte (GH), que consiste em identificar pontos ou regiões importantes como referência a serem utilizados. Após esta ambientação inicial, a análise continua com o estudo dos aspectos topológicos e táticos.

Para ser possível a realização deste estudo, existem uma série de dificuldades. Uma delas está no alto custo referente a deslocamento, hospedagem e alimentação.

Outro aspecto a se considerar em uma região real está na grande abstração existente quando se realiza a análise do terreno onde se pretende operar. Em questionário aplicado aos instrutores do Curso de Aperfeiçoamento Avançado do Corpo de Fuzileiros Navais, descrito na Seção 2.2.1, uma das dificuldades relatadas, de maneira geral, foi a de conseguir mostrar como o planejamento na carta se traduz no terreno. Este problema ocorre quando se busca transmitir aos militares a ideia de manobra, que pode incorrer em erros de interpretação.

Já no “exame de situação do Comandante”, são criadas 3 opções de ações a serem executadas pelas tropas, com o foco em resolver o problema militar. Estas opções chamamos de Linhas de Ação (LA), e são testadas através de confrontos, que consistem na realização de iterações, considerando as possibilidades que o inimigo possui na área de operações [14].

Sendo assim, cada LA é confrontada individualmente, através de Jogos de Guerra (JG), fornecendo conclusões de melhor qualidade para auxiliar na decisão do Comandante. Segundo Peter Perla, JG é “um modelo ou simulação de guerra, cujo funcionamento não envolve as atividades de forças militares reais, e cuja sequência de eventos afeta e, por sua vez, é afetada pelas decisões tomadas pelos jogadores representando os lados adversários” [16].

Cabe considerar, ainda no âmbito dos aspectos táticos, os aspectos relacionados aos domínios híbridos existentes, onde podemos citar a guerra urbana,

ameaças cibernéticas, tecnologias baseadas no espaço e táticas assimétricas. Sendo assim, o sucesso da missão não depende apenas de mão de obra ou poder de fogo, mas da velocidade, da qualidade dos planos confeccionados e da consciência situacional em tempo real.

Uma grande dificuldade para esta etapa está na falta de regiões mapeadas recentemente, dificultando os testes das sugestões das LA em nosso nível de planejamento. Desta forma, o Comandante decidirá sobre a ação a ser tomada e expedirá suas ordens a todos os elementos subordinados com base em informações sobre a região, normalmente desatualizadas.

Outra dificuldade está na necessidade de se controlar diversos fatores, tais como tempo, condições de trafegabilidade, clima etc., além dos testes do plano, o que pode ocasionar perda de dinamismo e foco nos aspectos negativos do planejamento.

Cabe destacar aqui que o Corpo de Fuzileiros Navais possui um simulador chamado Sistema de Jogos Didáticos (SJD), que possibilita a realização de JG assistidos por computador. Neste simulador, tanto o ambiente como as tropas são simuladas, cabendo aos planejadores se dedicarem apenas ao teste do planejamento. A utilização do SJD traria um dinamismo maior e melhores condições para a realização dos confrontos através dos JG. Mais detalhes sobre o SJD na seção 2.4.1.

Porém, uma dificuldade está na criação da área virtual para o SJD. Dados empíricos obtidos no Centro de Simulação do Corpo de Fuzileiros Navais (CSimCFN) mostram que o tempo médio gasto para a criação de um terreno com as trafegabilidades corretas, para uso em JG, com a área de $60km^2$, é de cerca de 2 a 3 meses. O resultado obtido normalmente é uma imagem em “.bmp” onde cada pixel equivale a uma área de $10m^2$.

Para se obter informações sobre uma determinada região rapidamente, planejadores normalmente buscam cartas militares, fotografias, imagens de satélites, entre outros subterfúgios. Dentre as informações com maior probabilidade de estarem atualizadas, destacam-se as imagens de satélites.

Os dados de imagens satélites, obtidos por meio de sensoriamento remoto, abrangem várias fontes, incluindo as imagens ópticas, imagens multiespectrais, imagens hiperespectrais e imagens de Radar de Abertura Sintética (do inglês, Synthetic Aperture Radar - SAR). As imagens SAR utilizam bandas do espectro eletromagnético, tecnologias de sensoriamento remoto ativo, para capturar imagens de superfície de alta resolução, que não são afetadas pelas condições atmosféricas, nuvens ou noite. Sendo assim, elas oferecem vantagens não encontradas em imagens de luz visível [17]. Uma possibilidade está na utilização das faixas do espectro características das cores vermelho, verde e

azul para a criação de imagens, não sendo estas imagens ópticas.

Um aspecto importante está no tempo em que um satélite pode levar para revisitar um mesmo local, que pode ser diário, como é o caso das constelações WorldView e Pléiades, a alguns dias, neste caso QuickBird-2 e GeoEye-1 [18]. Esta característica é importante pois permite monitorar e obter imagens em tempo hábil para planejar e conduzir ações em pouco tempo, tendo em vista que uma constelação de satélites pode retornar imagens diárias ou com frequências até maiores.

Nos últimos anos, a tecnologia de reconhecimento de sensoriamento remoto encontrou aplicações em diversos campos, como agricultura moderna, gestão florestal, planejamento urbano, gestão de recursos naturais e monitoramento de desastres [18].

Uma destas aplicações está em iniciativas como a do one world terrain [19], que utilizam imagens satelitais e outras fontes para a criação de terrenos para simulação, treinamento e experimentação. A simulação possui o potencial de permitir um treinamento com maior qualidade e o consequente preparo da Tropa [20].

Desta forma, um estudo do terreno poderia ocorrer primeiramente em Realidade Virtual (RV), observando e interagindo no terreno virtualmente criado baseado nesta imagem satélite antes de chegar ao local de reconhecimento. Uma segunda aplicação está no uso deste terreno em JG por meio do SJD, por ocasião do confronto das LA durante o planejamento, antes de deslocar para a área de operações.

De forma a obtermos imagens de qualidade e atualizadas, uma possibilidade para a confecção deste tipo de material estaria na utilização de Inteligência Artificial. Neste sentido, o uso de técnicas de segmentação semântica em imagens de sensoriamento remoto poderia gerar as camadas necessárias para a criação da CT, do terreno virtual e do terreno para ser utilizado nos JG, com maior rapidez.

Este resultado permitiria chegar ao local de operações já possuindo um primeiro contato com a região através da sua análise em um simulador, bem como realizados os testes ao planejamento confeccionado para o cumprimento da tarefa proposta por meio dos JG.

1.1

O problema de pesquisa

Considerando os aspectos citados anteriormente, podemos relacionar uma série de dificuldades para se realizar o planejamento de uma operação militar. Entre elas, destacamos:

- A criação da CT é um processo demorado e trabalhoso, porém muito importante para o planejamento de uma operação militar;
- Existe uma dificuldade em se realizar a análise da região onde se pretende operar por demandar conhecimento detalhado sobre o assunto;
- Existem dificuldades para se ir ao terreno realizar o EATT;
- O EATT tem que ser rápido, preciso e com poucos erros;
- Existe muita abstração envolvida no entendimento e na transmissão do EATT para os demais planejadores;
- O confronto do planejamento normalmente se baseia em informações pouco detalhadas sobre a região; e
- No confronto devem ser considerados diversos outros aspectos além do teste do planejamento em si, dificultando o foco nos seus pontos negativos.

Estas dificuldades aumentam principalmente quando a missão envolve grandes alterações no relevo da região, como em tragédias envolvendo desastres naturais. Podemos citar como exemplo o deslizamento das barragens da região de Brumadinho [21]. Nestas operações, o tempo de resposta deve ser o mais rápido possível e com o máximo de conhecimento atualizado sobre a região, sob pena de aumentar a quantidade de mortos fruto da demora em se chegar às regiões mais afastadas.

Baseado nas informações acima, a questão em que este trabalho está concentrado em responder está relacionado à dificuldade em se realizar três das tarefas do PPM, com o foco na realização de operações militares.

Neste sentido, elaboramos a seguinte pergunta: “Como desenvolver uma abordagem que forneça melhores condições de planejamento, com recursos atualizados e de maneira rápida, capaz de auxiliar no entendimento do estudo dos aspectos topológicos e táticos do terreno, bem como auxiliar na realização dos confrontos deste plano?”

1.2

Justificativa para o problema de pesquisa

Este problema pode ser observado até hoje, onde Moreira (2024) destaca o trabalho do Centro de Monitoramento de Deslocamentos Internos (IDMC):

“o trabalho de coleta e processamento de dados IDMC, que é alimentado por fontes nacionais da Defesa Civil e do Ministério da Integração e do Desenvolvimento Regional, poderiam ser explorados com maior intensidade na ciência nacional” [22].

Ele cita ainda que existe uma escassez de estudos sobre a elaboração de políticas públicas ante desastres naturais no país, o que denota que este problema ainda está em tela e não foi tratado em âmbito nacional.

Além disso, quando falamos de operações militares em conflitos armados, o assunto ainda encontra-se relacionado às Forças Armadas e, com isso, à Defesa Nacional. Desta forma, existe uma grande dificuldade em se obter fontes de pesquisa similares que abordem soluções para o PPM. O que se observa, no entanto, são empresas desenvolvendo tecnologias e vendendo-as em plataformas onde os conteúdos dos códigos não são revelados.

Estes softwares, conhecidos como “caixas pretas”, podem conter códigos capazes de copiar a metodologia de uso ou diversos outros tipos de informações sigilosas, armazenando-as de forma a serem compartilhadas indevidamente [23]. Este tipo de atividade é perigosa, pois permite que uma empresa privada conheça o *modus operandi* de nossas Forças, possibilitando a venda desse conhecimento para terceiros.

A importância de tratarmos as dificuldades na realização do PPM está na necessidade por informações o mais rápido possível, de forma a iniciar o planejamento da operação como um todo e de maneira correta. Outro aspecto é a busca pela transmissão das informações táticas e análise do terreno de maneira padronizada e que todos entendam, sem margens para erros de entendimento devido a abstrações. Desta forma, as missões serão cumpridas explorando ao máximo as características da área de operações e evitando ao máximo os danos colaterais causados por erros de planejamento.

1.3

As Hipóteses de pesquisa

De forma a responder à questão proposta, apresentamos nossas Hipóteses de Pesquisa (HP) com suas justificativas. Estas HP foram elaboradas em consonância aos diferentes problemas enunciados na seção 1.1.

O método de Comitê de modelos (do inglês, “Ensemble methodology”), para segmentação semântica binária, apresenta desempenho superior à abordagem multiclases tradicional, permitindo identificar com maior precisão as classes relacionadas à trafegabilidade do terreno em imagens de sensoriamento remoto. Desta forma, o método proposto contribuirá com a utilização de segmentação semântica na obtenção das diferentes classes existentes em imagens de Sensoriamento Remoto.

Os produtos gerados pela segmentação semântica – particularmente as classes identificadas – podem ser utilizados para criar cartas de trafegabilidade que atendem aos requisitos operacionais das Ope-

rações militares, acelerando o processo de análise e planejamento do terreno. Com isso, as classes serão segmentadas, gerando uma CT que auxiliará os analistas no seu planejamento. Assim, teremos maior velocidade em uma das etapas do planejamento, a saber, a EPI.

A integração de ambientes virtuais para estudo topotático do terreno, combinando recursos didáticos tradicionais com simulação virtual, potencializa o aprendizado e a confiança dos usuários (alunos e planejadores), sendo eficiente no treinamento antes do deslocamento ao campo real para análise de missão. Considerando nesta HP que o desenvolvimento de diferentes categorias de simuladores fornece um treinamento com maior qualidade e o consequente preparo da Tropa em melhores condições [20], um planejamento valendo-se de recursos instrucionais pode auxiliar no entendimento da análise do terreno para o planejamento de operações militares.

O pipeline desenvolvido, que abrange desde a segmentação semântica até a geração de terrenos virtuais para simuladores e jogos de guerra, é capaz de fornecer dados suficientes e condizentes com situações reais, possibilitando a realização de testes efetivos de planejamento militar em ambientes simulados. Com esta capacidade, os planejadores poderiam focar nos aspectos mais importantes do confronto, resultando em ajustes ao planejamento melhores e com maior qualidade.

1.4

O objetivo deste trabalho

O presente trabalho tem como objetivo melhorar a análise detalhada do terreno, a realização de treinamentos e testes de planos estratégicos em contextos reais e simulados. Para isso, desenvolveremos e validaremos um pipeline baseado em técnicas modernas de deep learning para segmentação semântica de imagens de sensoriamento remoto (Figura 1.2). O propósito é automatizar e aprimorar o processo de geração de cartas de trafegabilidade, terrenos virtuais e ambientes simulados para apoio ao planejamento de operações militares. Para atingir esse objetivo, abordamos o problema proposto em 6 fases.

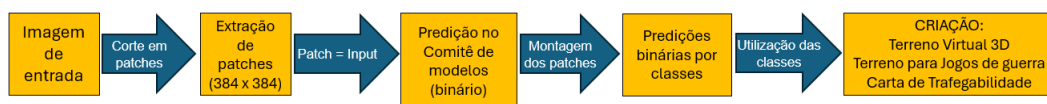


Figura 1.2: Pipeline proposto como objetivo deste trabalho.

Em uma primeira fase, foram realizados treinamentos com arquiteturas de redes neurais e diversos bancos de dados obtidos, de forma a se obter

a melhor combinação de parâmetros. Estes treinamentos buscam, primeiro, levantar as classes necessárias e, depois, obter as trafegabilidades necessárias. As imagens SAR procuradas são oriundas de diversas fontes, sendo algumas em alta resolução. A maior parte dos bancos de dados encontrados são utilizados em competições no site Kaggle [24].

Já na segunda fase, foi realizada a predição com imagens de teste e obtidas as classes necessárias para a carta de trafegabilidade. Posteriormente, essas classes foram organizadas juntas de forma a criar a carta de trafegabilidade.

Na terceira fase, foi criado virtualmente um terreno e um visualizador para observá-lo. Este simulador foi chamado de Simulador Virtual para Estudo Topotático do Terreno (SVETT). Neste simulador, foram inseridos recursos de forma a facilitar o estudo da região criada virtualmente em 3 interfaces distintas. Mais detalhes sobre esta solução proposta podem ser encontrados no Capítulo 5.

O terreno foi desenvolvido com base nas classes obtidas por ocasião das predições realizadas na quarta fase, utilizando o modelo treinado na terceira fase. Usando softwares como Blender, ArcGIS e Unity, essas classes foram utilizadas na criação do terreno virtual e do SJD.

Finalizando, na quarta fase, foi conduzida uma pesquisa sobre o uso do SVETT como ferramenta de análise do terreno. Este estudo foi realizado em um contexto onde seu emprego não é reservado, permitindo a obtenção de subsídios para responder às hipóteses de pesquisa. Sendo assim, buscamos os cursos de formação, onde a instrução sobre o estudo dos aspectos topológicos e táticos do terreno é ministrada sem conteúdo sigiloso. Isso permitiu a realização de pesquisas para obter respostas que validassem nosso trabalho do ponto de vista do uso da simulação no processo de análise do terreno.

1.5

Contribuições

A contribuição científica deste trabalho consiste no emprego do método de comitê de modelos com segmentação semântica binária em imagens satélite para a classificação dos diferentes tipos de áreas existentes.

Este método está inserido dentro de um pipeline, capaz de receber como dados de entrada imagens satélites de regiões do país e gerar cartas de trafegabilidade para emprego em operações militares de forma rápida, precisa e padronizada, substituindo etapas manuais lentas e suscetíveis a erros humanos. Esta automação promove avanços na análise de terreno e no planejamento de operações militares.

No campo da simulação militar, a integração do pipeline com ambientes

virtuais e simuladores possibilita a elaboração automática de terrenos virtuais, facilitando a análise didática (aplicações educacionais) e o teste de planos de operação em simuladores de guerra. Este tipo de treinamento majora sensivelmente a capacitação dos planejadores militares.

Além disso, embora o foco principal seja militar, a metodologia pode ser aplicável em outras áreas que demandem análise detalhada do terreno. Este emprego promove inovação em áreas como engenharia civil, defesa civil, gestão ambiental, agronegócio, entre outros.

O trabalho estabelece também um framework metodológico e um conjunto de soluções técnicas que podem ser expandidas por novos estudos em segmentação semântica, simulações computacionais e aplicações militares, tornando-se uma referência no campo.

Uma última contribuição consiste no simulador para estudo do terreno, com a possibilidade de interação em 3 diferentes interfaces. Este software possibilita o uso de simulação para a análise inicial do terreno, visando o planejamento com foco em operações reais.

Este trabalho foi organizado em 7 Capítulos. O segundo capítulo fornecerá uma fundamentação teórica sobre as diferentes atividades envolvidas. O terceiro capítulo elucidará trabalhos desenvolvidos previamente que basearam e estão relacionados à nossa solução proposta. O quarto capítulo versará sobre o modelo escolhido e treinado em segmentação semântica com as imagens de sensoriamento remoto, abordando os treinamentos realizados e os resultados obtidos. O quinto capítulo abordará o desenvolvimento da ferramenta de observação do terreno virtual, os testes e experimentos realizados, exibindo e analisando os resultados. No sexto capítulo, escreveremos sobre a criação da CT e criação do terreno para o SJD. Finalizando, o sétimo capítulo concatenará a sequência do pipeline, abordando seus principais tópicos e complementando o texto com trabalhos futuros.

2

Referencial teórico

Este capítulo visa definir conceitos e procedimentos, sejam eles do ponto de vista militar ou acadêmico. Para isto, vamos iniciar escrevendo sobre o Corpo de Fuzileiros Navais, tropa envolvida no escopo deste trabalho, e suas características.

Posteriormente abordaremos alguns aspectos do Processo de Planejamento Militar, ferramenta importante para a preparação para operações militares. Ainda neste contexto, elucidaremos aspectos sobre a análise do terreno, com o foco na condução de operações militares.

De forma a orientar melhor o nosso problema proposto, aplicamos um questionário a instrutores antigos e atuais do C-ApA-CFN. Este questionário tem como foco as dificuldades encontradas tanto para ir aos Exercícios no Terreno (ET) como para conduzi-los no alto das elevações.

Posteriormente, versamos sobre o uso da simulação nos treinamentos e planejamentos militares, abordando o processo manual de criação da carta de trafegabilidade e do terreno para o SJD.

Ainda neste contexto, falaremos sobre a Inteligência Artificial, passando por Deep Learning e técnicas de segmentação semântica em sensoriamento remoto. Finalizando, abordaremos o uso de sensoriamento remoto no estudo do terreno.

Cabe destacar que não será possível esgotar o assunto nem fornecer informações sobre tópicos de cunho sigiloso sobre qualquer um dos capítulos escritos, dado o caráter estratégico de alguns dos conteúdos.

2.1

Sobre o Corpo de Fuzileiros Navais

O CFN é uma tropa profissional e de caráter voluntário, onde os militares são admitidos por meio de concurso público. Sua origem data de 07 de março de 1808, data em que a Brigada Real da Marinha portuguesa chegou ao Rio de Janeiro, acompanhando a Família Real portuguesa que transmigrava para o Brasil [8].

Ao longo dos anos, o CFN recebeu diversas denominações: Batalhão de Artilharia da Marinha do Rio de Janeiro, Corpo de Artilharia da Marinha, Batalhão Naval, Corpo de Infantaria de Marinha, Regimento Naval e, finalmente, desde 1932, Corpo de Fuzileiros Navais. Na década de 1950, o CFN estruturou-se para emprego operativo como Força de Desembarque, passando a constituir

parcela da Marinha destinada às ações e operações terrestres necessárias a uma campanha naval.

Seus militares estão em permanente condição de pronto emprego. Seu caráter anfíbio e expedicionário por excelência é parcela do Conjugado Anfíbio da Marinha do Brasil¹ [7]. Além das Operações Anfíbias², seu emprego envolve a defesa das instalações navais, das instalações portuárias, dos arquipélagos e ilhas oceânicas nas águas jurisdicionais brasileiras. São capazes também de atuar com diferentes entidades em operações subsidiárias e, no contexto internacional, podem atuar em operações de paz e em ações humanitárias [8].

O próprio CFN é responsável pela formação, preparo e emprego de seus militares. Sendo assim, suas atividades envolvem tanto a constante busca pela capacitação como a participação em operações reais. Desta forma, cabe, respectivamente, aos Centros de Instrução e à Força de Fuzileiros da Esquadra (FFE) a sua realização.

A FFE é uma Força organizada, treinada e equipada para executar Operações Anfíbias. Ela é caracterizada pela integração de forças com funções de combate diferentes. Tais operações, apesar de geralmente serem finalizadas em terra, visam à obtenção dos objetivos de uma campanha naval e terrestre, de amplitude limitada. Porém, devido às ações serem concentradas em terra, é possível seu emprego em uma gama de atividades neste ambiente.

As tropas vêm sendo empregadas ultimamente em diversos tipos de operações. Primeiramente, podemos destacar a Operação São Francisco, onde os militares foram empregados no complexo da maré, realizando operações de Garantia da Lei e da Ordem (GLO) [26]. Em outras operações, tais como Jornada Mundial da Juventude e Copa das Confederações [27], os militares foram posicionados com a finalidade de executar a segurança em regiões tanto marítimas como terrestres. Em um terceiro tipo, focando a atuação sob a égide de Organizações Internacionais, como a Missão de Paz das Nações Unidas para a estabilização do Haiti (MINUSTAH) [28] e a Força Interina das Nações Unidas no Líbano (UNIFIL) [29], a tropa foi empregada sob a égide da ONU em Operações de Paz tanto em terra quanto no mar. Ao longo dos últimos anos, fruto das demandas com foco em ações subsidiárias, no período de dezembro a março, são ativados Grupamentos Operativos de Fuzileiros Navais em Apoio à Defesa Civil (fig. 2.1), com o foco no eventual emprego em situações de desastres naturais [30].

¹“Força Naval com um Grupamento Operativo de Fuzileiros Navais (GptOpFuzNav) embarcado, juntamente com meios aeronavais adjudicados, em condições de cumprir missões relacionadas às tarefas básicas do Poder Naval”

²Operações Anfíbias consistem na projeção de poder, a partir de uma força naval localizada no mar, para uma região litorânea hostil, contendo locais de importância operativa, normalmente portos ou aeroportos [25]



Figura 2.1: Emprego do Grupamento Operativo de Fuzileiros Navais em Apoio à Defesa Civil. Fonte: Portal BIDS, disponível em <https://portalbids.com.br/2024/02/26/marinha-mobiliza-grupamento-de-fuzileiros-navais-para-apoiar-vitimas-de-temporal-no-rj/>

2.1.1

Características

O CFN (fig. 2.2) possui como principais características a flexibilidade, versatilidade, mobilidade e permanência. Seus soldados-marinheiros são aptos tanto para a vida de bordo dos navios de guerra como para o combate em terra, e os seus meios são apropriados para o embarque em navios e posterior desembarque em terra [31]. Esta parcela da Marinha do Brasil, por estar em condição permanente de pronto emprego e, somando-se à flexibilidade que permite operar em diversas regiões, está constantemente se preparando para a condução de operações militares, sejam elas envolvendo conflitos ou operações humanitárias.

Entendemos como flexibilidade, segundo a Estratégia Nacional de Defesa, como sendo “a capacidade de organizar grupamentos operativos de diferentes valores, em função da missão, possibilitando a gradação no seu emprego”. Já a versatilidade “permite alterar a postura militar, mantendo a aptidão para executar uma ampla gama de tarefas”. Sobre a mobilidade, esta refere-se à infraestrutura logística de transporte do País que permita às Forças Armadas “deslocar-se, rapidamente, para a área de emprego, no território nacional ou no exterior, quando assim impuser a defesa dos interesses nacionais”. Finalizando, a permanência indica a “capacidade de operar, continuamente,



Figura 2.2: O Desembarque anfíbio tem por finalidade projetar poder, a partir de uma Força Naval, para um território hostil ou potencialmente hostil. Fonte: Livro Branco de Defesa Nacional, ano 2012 [8]

com independência e por longos períodos, em áreas distantes e de grandes dimensões” [9].

Para a manutenção do comando e do controle por ocasião do cumprimento de suas missões, em que pese a necessidade de se ter flexibilidade, é importante que todos entendam suas tarefas. Para isso, torna-se imprescindível a criação de Documentos Operativos [14].

Estes documentos são capazes de fornecer um detalhamento das ações a serem executadas, bem como fornecer uma análise completa sobre diversos aspectos, onde podemos citar os aspectos topotáticos do terreno [11]. Para que todos os requisitos e informações necessárias sejam fornecidos de forma que todos entendam, é executado o Processo de Planejamento Militar (PPM), visto com mais detalhes a seguir.

2.2

Do Processo de Planejamento Militar

Para a execução de uma operação militar, existe a necessidade de se pensar em todos os detalhes e possibilidades. Esta preocupação visa estar pronto, tanto em termos de treinamento como de equipamentos, para o máximo de intercorrências possíveis.

Sendo assim, peculiaridades sobre a complexidade da operação a ser executada, sobre a especificidade da logística para o cumprimento das tarefas, sobre as naturezas distintas das forças envolvidas (aeronaves, meios sobre rodas e meios sobre lagarta) entre outros, precisam ser discutidas e coordenadas de

forma a se obter o sucesso no cumprimento da missão atribuída. A maneira padronizada para se analisar e escriturar todos estes detalhes é através do PPM [10].

O PPM teve sua origem na Prússia, na metade do século XIX, sendo adotado e aperfeiçoado pelo Naval War College. Foi utilizado na primeira e segunda guerras mundiais, recebendo novos conceitos e processos desde então. Na Marinha do Brasil, foi adotado e utilizado nos cursos da Escola de Guerra Naval e utilizado doutrinariamente pelos Comandantes de Forças Navais. Em 2001, este processo foi documentado e aprovado pelo Ministério da Defesa de forma a ser utilizado por todas as Forças Singulares (Marinha, Exército e Aeronáutica).

Apesar de seguir as diretrizes básicas do PPM, o planejamento de Operações Anfíbias, característico do CFN, se distingue por meio da simultaneidade, paralelismo e detalhamento. Ele é simultâneo pois é realizado concomitantemente por dois ou mais níveis de um mesmo comando. Por exemplo, ele ocorre quando uma Brigada precisa realizar um planejamento juntamente com o escalão superior (uma Divisão) e com seus subordinados (os Batalhões) para uma operação de GLO.

Além disso, ele é paralelo, pois é realizado concomitantemente por escalões correspondentes de forças diferentes, onde a Divisão de Fuzileiros Navais, responsável pelas operações em terra, planeja juntamente com a Divisão da Esquadra Brasileira, responsável pelas operações no mar. E é detalhado, pois é desenvolvido minuciosamente por todos os níveis de comando participantes, sem prejudicar a flexibilidade das Unidades subordinadas [14].

Para que estas características sejam empregadas sem gerar incongruências durante o processo, existe uma ordem padronizada a orientar os trabalhos, conhecida como Sequência de Ações de Comando e Estado-Maior (SACEM).

A SACEM funciona como um guia para o desenvolvimento de ações por um comandante e os planejadores de seu Estado-Maior (EM), compreendendo cerca de 12 etapas, mostradas na fig. 2.3. Estas são divididas entre reuniões, disseminação de informações, processos decisórios, elaboração de documentos, entre outros, todas com o foco em resolver um problema militar. Todo este processo é sistemático, seguindo uma sequência correta, e tem como foco a orientação das atividades dos planejadores militares dos diferentes setores e níveis de decisão [14].

Tão logo a missão seja recebida, a "Orientação Inicial do EM" deve ocorrer o mais breve possível, de forma a permitir o início dos trabalhos. Sendo assim, o comandante reúne seu EM para divulgar o seu entendimento sobre a missão e expressar suas recomendações e intenções para o prosseguimento do

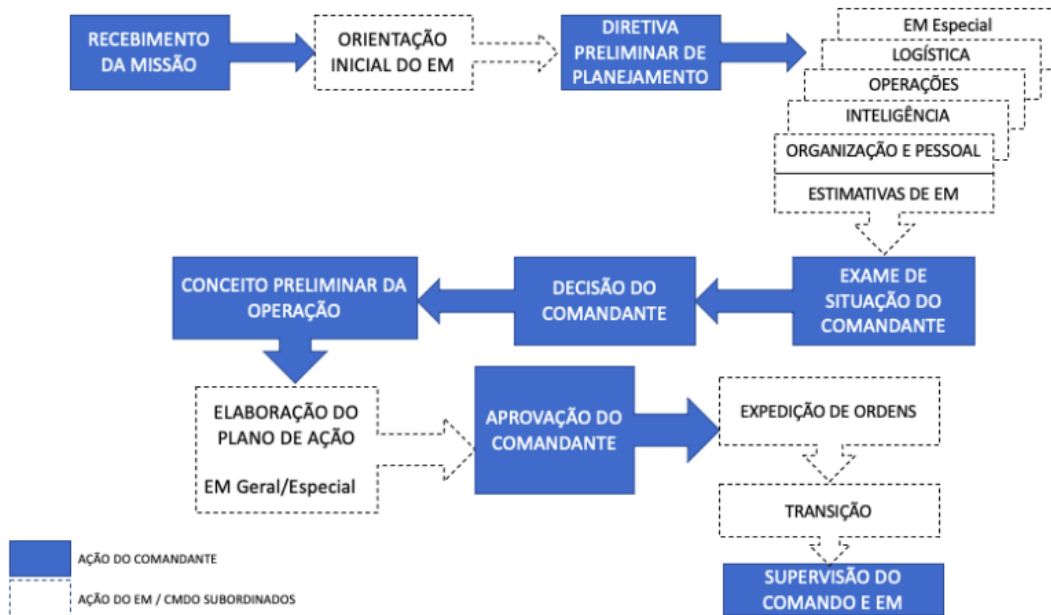


Figura 2.3: Sequência de ações a serem realizadas no planejamento de uma operação militar, destacando a etapa onde se faz necessário obter informações sobre o terreno. Não é o foco deste trabalho esgotar o assunto, sendo assim, estas etapas não serão explicadas aqui.

planejamento.

Nesta etapa, o EM apresenta todos os dados disponíveis até então, que envolvem os obtidos por ele e por outros escalões (superiores, adjacentes ou subordinados). Entre esses dados, avulta de importância a Estimativa Preliminar de Inteligência (EPI), por ser básica para o prosseguimento dos trabalhos. Esta EPI não constitui propriamente um documento, pois é apresentada verbalmente. Nela são abordados diversos assuntos, onde destacamos as principais características da área de operações (AOp).

Dentro das características da AOp, são analisadas as peculiaridades da região que influenciarão direta ou indiretamente na condução das operações militares. Entre elas, podemos destacar a vegetação, relevo, rios e lagos, vias de transporte terrestres, entre outros. Como produto final, obtemos, entre uma série de informações, um documento chamado carta de trafegabilidade (CT). Este documento é a representação gráfica produzida fruto da análise dos fatores referentes aos aspectos topográficos, sendo ferramenta essencial para a análise tática da AOp e cumprimento das outras etapas do PPM [32]. Podemos observar um exemplo de carta de trafegabilidade na fig. 1.1.

Como exemplo da fig. 2.4, citamos uma região onde, por meio de fotografias aéreas e imagens de satélites, constatamos ser de uma região com gramínea. Esta região é classificada como possível área de passagem de viaturas sobre rodas, pessoal a pé, viaturas que utilizam esteiras, entre outras, devido

à sua facilidade de deslocamento.

Este documento, que consiste basicamente em um mapa de uma determinada região contendo as manchas com as mobilidades possíveis em cada área, deve ser confeccionado o mais rápido possível, de forma a possibilitar um planejamento adequado, exequível e aceitável. Uma opção para a confecção deste tipo de produto está no uso de imagens de satélites.

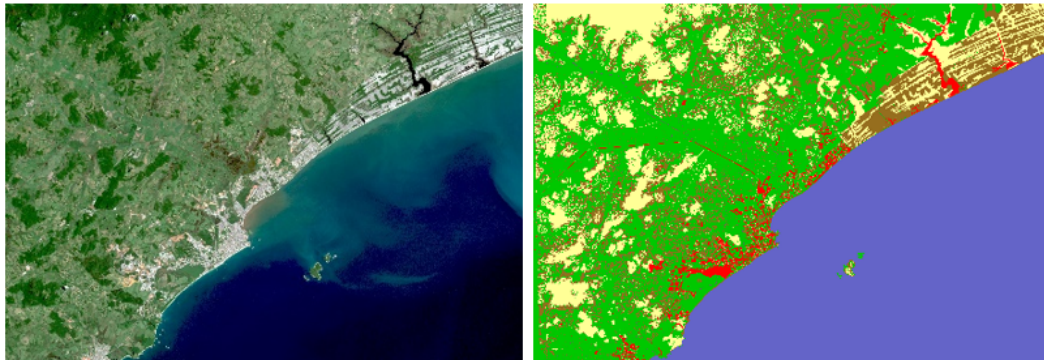


Figura 2.4: Comparativo entre imagens. A esquerda observamos uma imagem satélite obtida por meio de fontes abertas. A direita, uma carta de trafegabilidade mostrando as restrições de deslocamento da imagem a esquerda. Cabe destacar que este tipo de carta de trafegabilidade é utilizado no SJD.

Tanto o exemplo da fig. 2.4 como o da fig. 1.1 são maneiras de se confeccionar a carta de trafegabilidade. Este material, dada a necessidade de estar pronto bem no início dos trabalhos dos planejadores do EM, demanda um grande esforço por parte do responsável por este tipo de planejamento, o Oficial de Inteligência, sendo necessária sua constante atualização.

Após efetuar a orientação inicial do EM, o Comandante, assessorado por seu EM, prossegue pelas diferentes fases do planejamento, até a etapa chamada Exame da Situação do Comandante, onde são analisadas Linhas de Ação (LA) e testadas individualmente através do Confronto. A etapa da LA consiste na criação de 2 a 3 esboços de soluções para o problema militar, que são analisados pelo Comandante. Após sua análise, as LA são confrontadas uma a uma com possíveis planejamentos de forças adversas, onde podem ser empregados militares para fazer esse papel. Após selecionada a LA, o Comandante formula a sua decisão.

Após a formulação da decisão do Comandante, são elaborados os diversos planos e ordens para sua aprovação ou assinatura, e posterior expedição. Em seguida, são realizadas diversas atividades, tais como a MAPEX, atividade onde o Estado-Maior vale-se de cartas para apoio visual à dinâmica do briefing para seus comandados [14]. Esta atividade visa permitir ao comandante apresentar pessoalmente, discutir e ensaiar o planejamento final com o EM e os comandantes subordinados antes da execução.

Durante todo o processo, não somente o Oficial de Inteligência, mas também os demais Oficiais planejadores do EM, devem saber como interpretar as informações atinentes ao terreno, sendo capazes de realizar uma análise correta da região e identificar seus aspectos topográficos e táticos. A este estudo damos o nome de Estudo dos Aspectos Topotáticos do Terreno (EATT).

2.2.1

Sobre a Análise do Terreno

Como mencionado na seção anterior, o EATT é uma das etapas iniciais da SACEM e uma das mais importantes. A partir da análise do terreno, será possível explorar as vantagens existentes na região, obtendo o máximo de aproveitamento de suas características e evitando possíveis danos colaterais.

Como estamos falando sobre operações militares em diferentes regiões, podemos destacar aqui os diferentes Biomas do Brasil, que por si só são bastante complexos. Além disso, temos a Amazônia azul, que é o conjunto formado pelo Mar territorial, Zona Econômica Exclusiva e Extensão da plataforma continental, totalizando 8,5 milhões de Km² [33]. Além disso, devemos considerar também nosso espaço aéreo.

Por isso, a Análise de uma região para o planejamento e a condução de operações militares é muito difícil de ser realizada dentro de uma sala fechada, principalmente pelo fato de que deve-se levar em consideração as características atuais da região.

Uma maneira possível para isto seria observar uma imagem aérea e compará-la com uma carta topográfica, o que, na maioria das ocasiões, não transmite uma noção espacial adequada. A melhor opção, em que pesem as condições climáticas e táticas, seria ocupar posições em locais que permitam a melhor observação do terreno aonde se pretende operar. Normalmente, os melhores locais para observação são o topo das elevações mais altas da área de operações.

Para iniciar o estudo, precisamos entender o processo em que se busca observar as características do terreno que podem influenciar decisivamente as ações em um determinado local, conhecido como Giro do Horizonte (GH). O GH é a identificação em uma Carta Militar (CM) dos diversos pontos de referência importantes para a operação no terreno.

Este estudo está limitado a quatro quilômetros, a partir da posição de observação, devido à dificuldade de observação, além dessa distância, sem o equipamento adequado [34]. Os critérios ou precauções que devem ser levados em consideração são os seguintes:

1. Ocupar uma posição que possua a melhor observação possível da região a ser estudada. Esta posição chamaremos de Ponto Estação (PE);
2. Determinar o PE (localização atual) e orientar a carta militar (CM), que consiste em posicioná-la em conformidade com o terreno, alinhando seu Norte ao Norte na região a ser observada;
3. Realizar uma inspeção sumária dos acidentes próximos mais notáveis, identificando-os com a carta militar, para garantir que a orientação esteja correta;
4. Ficar ao lado daqueles que assistirão a apresentação do “Giro do Horizonte”;
5. Apresentar o PE;
6. Apresentar a direção geral do Norte (ou Sul);
7. Dividir o terreno a ser identificado em linhas (faixas ou setores);
8. Para cada linha, iniciar a identificação da esquerda para a direita e do mais próximo para mais afastado; e
9. Apresentar os detalhes altimétricos (elevações) e depois os detalhes planimétricos (estradas, rios, lagos etc.).

Este estudo facilita o entendimento à medida que observamos as principais características da região e sua conformação, padronizando pontos de referência e locais importantes a serem considerados. Após esta fase, iniciamos o estudo tático do terreno, onde efetuamos uma série de análises e comparações entre faixas existentes em uma região ³ com a finalidade de melhor explorá-lo.

Importante ressaltar que existe a necessidade de se “imaginar” estes aspectos mais táticos. Um dos motivos é a impossibilidade de se posicionar os meios no terreno antes do início de uma operação militar. Outra dificuldade está em marcar nitidamente as medidas de coordenação, tais como limites de atuação entre as Unidades militares. Esta abstração, apesar de ser amenizada pelos pontos de referência estabelecidos durante o GH, é dificultada à medida que são realizadas análises detalhadas tanto em regiões próximas como em mais afastadas do PE.

Outro agravante está nos possíveis erros de indicações, a partir do PE, dos locais a serem estudados. Ao se conduzir a direção de observação dos demais

³Faixas do terreno são setores de observação que podem variar em distância e direção utilizados para analisar os detalhes de toda a região [31]

para o local aonde se pretende estudar, existe uma grande possibilidade de os observadores estarem focando em pontos próximos, porém diferentes do esperado pelo analista. Desta forma, o conteúdo a ser entendido pode conter erros que incorrerão em aprendizagem equivocada.

Outros aspectos relevantes que consomem tempo são a ambientação no terreno, o GH e a análise tática do terreno. Além disso, existe a dependência das condições climáticas para seu estudo. Todo este processo pode demorar, em que pese o fato de o militar nunca ter ido antes à região, o que pode ser agravado devido às restrições de visibilidade como, por exemplo, chuvas ou neblinas.

Cabe destacar também que conduzir o pessoal militar para um estudo fora da área de aquartelamento exige uma grande quantidade de material e pessoal. Nestes deslocamentos, devemos considerar a disponibilidade de tempo, os custos ligados à logística, o pessoal envolvido em diversas funções (segurança, apoio, entre outros) e a preparação visando atender a todos os aspectos citados.

Planejar e preparar um treinamento ou uma missão envolvem vários detalhes. Entre eles, destacamos a necessidade de veículos para deslocar com as equipes (em alguns casos, requer a compra de passagens aéreas e terrestres), alojamento e alimentação para as equipes (em alguns casos em cidades onde não há Unidade Militar próxima). Para o EATT propriamente dito, cresce em importância o conhecimento sobre a região de estudo, o estabelecimento da segurança nas áreas a serem utilizadas, o reconhecimento das melhores posições com o foco na sequência das ações da operação, a rapidez das mudanças de posição e a preparação dos detalhes da instrução.

No caso de instruções e adestramentos, o tempo disponível para tais preparações é curto devido ao calendário escolar, o que é prejudicial para a melhor exploração do assunto. Outra dificuldade está na necessidade de se realizar grandes deslocamentos para mudança de PE, consumindo muito tempo para tal.

Outro fator importante se deve à inviabilidade da execução de adestramentos adicionais caso exista a sua necessidade para complementar o assunto. Neste sentido, como cada instrução sobre o EATT consome, em média, 10 dias, incluindo reconhecimento, preparação e treinamento, torna-se difícil a alocação de tempo extra.

Como exemplo, tomamos como base o Curso de Aperfeiçoamento Avançado do Corpo de Fuzileiros Navais (C-ApA-CFN) do ano de 2018, composto por 30 Oficiais-Alunos (OfA) e 22 militares da equipe de instrução. Para tornar um adestramento possível, existem custos referentes a deslocamento, hospeda-

gem e alimentação. A Tabela 2.1 exibe os gastos realizados com passagens e alimentação para os exercícios do respectivo ano.

LOCALIDADE	ALIMENTAÇÃO	DESLOCAMENTO	TOTAL
Seropédica / RJ	R\$ 11.520,00	R\$ 1035,00	R\$ 12.555,00
Itaóca / RJ	R\$ 5.670,00	R\$ 8.180,00	R\$ 13.850,00
Macaé / RJ	R\$ 29.205,00	R\$ 320,00	R\$ 29.525,00
Aracaju / SE	R\$ 29.205,00	R\$ 41.140,00	R\$ 70.345,00
Itajaí / SC	R\$ 37.260,00	R\$ 38.526,60	R\$ 75.786,60
Ladário / MS	R\$ 37.260,00	R\$ 49.898,00	R\$ 87.158,00

Tabela 2.1: Gastos necessários para conduzir um grupo contendo 30 OfA e 22 militares da equipe de instrução.

Para localidades mais próximas, como é o caso de Macaé e Seropédica, no estado do Rio de Janeiro, os custos foram menores pois os militares efetuaram o deslocamento por meios próprios. Porém, quando as distâncias envolvidas são muito grandes, como Aracaju, que dista cerca de 1.821Km da cidade do Rio de Janeiro, existe a necessidade de gastos com passagens aéreas, visto que o tempo de deslocamento comprometeria a execução do exercício⁴. Em que pese a grande diversidade de nosso país, esses deslocamentos se fazem mister para os OfA terem noção das peculiaridades das diferentes regiões do Brasil.

Por se tratar de uma região normalmente afastada de centros urbanos e em locais pouco povoados, onde pessoal uniformizado está envolvido, existem aspectos a serem observados como transporte do pessoal até os locais de estudo, além de segurança nas áreas de adestramento. Esta demanda de pessoal é relativamente grande, podendo influenciar na rotina das Unidades por se tratar de um período considerável de afastamento, assim como na ausência dos militares de seus lares.

Com a finalidade de se verificar possíveis alterações que, por ventura, tenham ocorrido na área desde o último exercício, 90% dos instrutores normalmente iniciam o deslocamento para a região cerca de 5 dias antes. Além disso, são definidos neste período os assuntos a serem abordados e realizado todo o estudo necessário para se preparar a aula ou a análise.

2.3

Questionário aplicado aos instrutores do C-ApA-CFN

De forma a relacionar os aspectos positivos do ET, relacionar as dificuldades em se realizar as instruções nos ET e dificuldades para se instruir os

⁴Em pesquisa realizada na internet, iniciando o deslocamento as 07h30 do dia 15 de março de 2025, a previsão de chegada seria as 14h30 do dia 17 de março. Considerando o tempo disponível para o adestramento de 5 dias, período em que os OfA dispõem no calendário escolar para tal, todo o exercício teria que ser feito em apenas um dia [35]

alunos sobre o estudo topotático do terreno, buscamos as opiniões de antigos e atuais instrutores do C-ApA-CFN. Esta pesquisa teve como objetivo fornecer subsídios para orientar nossas buscas por uma solução para problemas normalmente observados nesses tipos de exercícios.

Composição deste questionário foi de perguntas quantitativas e qualitativas. As perguntas quantitativas foram realizadas com o propósito de qualificar nosso público-alvo. Já as perguntas qualitativas foram criadas com o propósito de atender às demandas deste questionário, levantando aspectos positivos e dificuldades encontradas no ET.

Para a aplicação dos questionários, foi utilizado o Google Forms [36]. Um questionário piloto foi aplicado, sendo posteriormente descartado. Todos assinalaram dando consentimento para a realização do questionário através de um termo de consentimento, inserido na seção do formulário que antecedia o início das perguntas.

Ao todo, participaram deste questionário 16 participantes. Sobre os instrutores que responderam às perguntas, todos possuem entre 32 e 43 anos. Além disso, 13 dos 16 participantes estão no posto de Capitão de Corveta e possuem acima de 18 anos servindo à Marinha do Brasil, o que denota experiência e vivência no CFN. Finalizando, 13 dos 16 instrutores já ministraram instrução a pelo menos 2 turmas do C-ApA-CFN, e já participaram de mais de 6 ET relacionados ao EATT, o que denota experiência sobre o assunto.

Referente à pergunta “Em sua opinião, quais os pontos positivos em se realizar a análise do terreno no Exercício no Terreno (ET)” destacamos as seguintes respostas:

- Uniformizar o conhecimento do terreno;
- Os ET Provém ao aluno o conhecimento e real noção do planejamento. Podendo verificar se determinados ações, distâncias envolvidas, natureza da tropa são exequíveis no terreno;
- O estudo do terreno no ET é uma ferramenta muito importante para mostrar aos OfA se o que está sendo planejado pode ser levado em consideração ou necessitará de modificações. Isso ocorre porque durante o planejamento tudo é aceito no papel, mas quando se vai ao terreno pode-se verificar se é possível, pois pode-se achar um terreno desfavorável para algum tipo de progressão de viatura, por exemplo;
- Permite que o OfA aplique seus conhecimentos aprendidos em sala no terreno, de forma a saber efetivamente interpretar a carta com o terreno, onde efetivamente será realizada a operação planejada;
- Permite uma melhor análise do terreno, auxiliando na tomada de decisão;

- Permite o OfA a adaptar o seu planejamento às condições reais de emprego;
- Permitir que os OfA entendam as influências da AOp em sua manobra;
- Permite ao Aluno aplicar os conceitos teóricos na visualização do terreno, incluindo as possíveis alterações das condições topográficas em comparação com o que é representado tanto na carta quanto em imagens de Google Earth bidimensionais;

Baseado nas respostas listadas, entendemos que a visita ao terreno é importante para termos uma real noção das consequências de nosso planejamento, bem como padronizarmos o conhecimento a cerca da região onde se pretende conduzir as ações. O EATT cresce de importância à medida que, segundo as respostas obtidas, “o papel aceita tudo”, dado o fato de não existir nenhuma tecnologia no país aplicada à análise do terreno que auxilie na correção do planejamento escrito.

Referente à pergunta “Em sua opinião, durante o exercício, quais as dificuldades em se realizar as instruções nos ET” destacamos as seguintes respostas:

- Fazer com que os alunos observem o que está na carta com o que está no terreno. Principalmente quando o terreno está diferente da carta por diversos motivos evolutivos, como construção de casarios e outras obras de arte que acabam confundindo curvas de nível ou diferenças entre elevações;
- Normalmente, as grandes distâncias envolvidas;
- Questões logísticas e coordenação com os responsáveis pelos locais planejados como PO;
- Dificuldades inerentes a aspectos de ordem logística (escassez de pontos de apoio com estrutura adequada a execução dos exercícios, limitações de meios, etc);
- A dificuldade é mais administrativa. Arrumar viatura é conciliar datas com o calendário do curso. E até mesmo verba para que o exercício ocorra;
- A carta estar desatualizada por alterações recentes ocorridas no terreno;
- Normalmente são de caráter logístico (numero de vtr), condições de trafegabilidade até os locais de PO, e em alguns casos acesso negado aos mesmos PO ou outros locais que poderiam ser usados como bons locais de PO;

Nestas respostas foram relacionados, de maneira geral, a logística, as grandes distâncias, a escassez de pontos de apoio e o calendário do curso com poucas disponibilidades de tempo. Neste sentido, entendemos que tais fatores dificultam a execução de mais exercícios. Neste sentido, existe a necessidade de se obter uma ferramenta que permita aos OfA realizar os estudos mesmo na área de aquartelamento e de maneira prática e rápida.

Referente à pergunta “Em sua opinião, estando em um local de instrução (ocupando o PO com os alunos e instrutores), quais as dificuldades para se instruir os alunos sobre o estudo topotático do terreno” destacamos as seguintes respostas:

- Boa parte dos alunos não tinham tido essa instrução ou essa experiência durante a carreira, isso faz com que eles fiquem muito amarrados a carta e tenham dificuldade de entender o terreno. Uma consequência dessa dificuldade é fazer o aluno compreender que a manobra planejada deverá ser ajustada por conta das variáveis observadas no terreno;
- Manter todos os OfA atentos ao conteúdo apresentado mesmo com as dificuldades naturais de estar no PO;
- O aluno assimilar o que foi feito na carta no terreno. Muitos alunos após uma pergunta ficam olhando para a carta;
- Alguns tem dificuldade em expressar em palavras o que está vendo no terreno;
- A grande dificuldade é conseguir mostrar para o OfA como o terreno se comporta de acordo com a carta, de forma a propiciar o estudo válido para o planejamento. O estudo do terreno é uma arte;
- Falta de certeza se os alunos realmente identificaram o que se pretende mostrar. Dificuldade com o volume da voz, quando se têm muitos alunos;
- Às vezes o tempo limitava o aprofundamento da análise;
- As características que são “visualizadas” no terreno exigem do instrutor e alunos, quando estão descrevendo o terreno, a capacidade de bem detalhá-lo, o que nem sempre é claro;
- Transmitir o entendimento que as vezes não é fácil para quem não tem vivência.

Nesta terceira pergunta destacamos, entre as respostas recebidas, a dificuldade dos OfA em entender ou assimilar no terreno o planejamento realizado na carta. Outra resposta predominante foi a dificuldade em expressar em palavras o que se está vendo no terreno e detalhá-lo. Uma terceira resposta foi o fator tempo limitando o aprofundamento da análise. Estas

dificuldades denotam a necessidade de uma ferramenta que facilite a descrição e o detalhamento dos estudos e análises realizados no terreno.

Além disso, constatamos o fato de os instrutores possuírem dificuldades em passar alguns conceitos para os alunos devido à necessidade de imaginar as unidades militares ou medidas de coordenação no terreno real. Situações como essas geram como consequências vários erros conceituais, equívocos na escolha de locais adequados para o desenvolvimento de manobras militares, dificuldades em sanar dúvidas após o retorno às Unidades e ausência de um local semelhante para discussões sobre o assunto [12].

Um aspecto importante a ser considerado é o fato de nem sempre ser possível ter no quartel um lugar com características que se aproximem das que serão encontradas na área de treinamento, o que auxiliaria na preparação tanto dos instrutores quanto dos alunos para o exercício. Neste caso, a inserção dos alunos em um Ambiente de Realidade Virtual poderia trazer benefícios aos adestramentos.

Foram analisadas todas as respostas recebidas dos questionários. Em que pese as demais respostas do questionário estarem relacionadas no anexo A, obtivemos informações importantes que orientaram o trabalho de desenvolvimento da ferramenta de visualização do terreno virtual a ser proposta neste trabalho.

Cabe considerar ainda um aspecto destacado por Thijssen e Bosma, em “Recommendation System in an Integrated Digital Training Environment for the 5th Generation Air Force”. Este trabalho destaca o fato de que a tecnologia usada para encontrar uma solução deve se comportar de forma adequada à maneira como os alunos aprendem [13].

Baseado nos dados colhidos, esta ferramenta deve possibilitar a tradução do conteúdo inserido na CM para o terreno, de forma que seja possível auxiliar na expressão do que se está observando. Além disso, deve ser considerada a padronização dos conhecimentos, mostrando as dificuldades e problemas nos planos confeccionados. Outra característica seria a necessidade de rapidez nos deslocamentos entre os Pontos de Observação (PO) e a otimização do tempo para realizar o EATT. A próxima seção discorrerá sobre o uso de simulação hoje.

2.4

O uso da simulação nos treinamentos e planejamentos militares

O uso da simulação como ferramenta de aprendizagem teve seu auge, tanto na academia como em empresas, no final dos anos 90. Suas vantagens envolvem a possibilidade de se prover feedback imediato, a capacidade de se

efetuar correções e orientações em tempo real e em ambiente controlado, além de permitir a troca de informações importantes sobre as atividades realizadas [37].

De maneira geral, a simulação é qualquer sistema computacional capaz de executar um modelo para gerar seu comportamento, que pode ser por meio de um processador, um servidor, a mente humana, entre outros. Já o modelo é entendido como uma série de instruções para se gerar o comportamento pretendido [38]. Existem vários tipos de simulações, podendo ser empregadas na arte, engenharia, treinamentos, entre outros. Neste trabalho, focaremos nas simulações voltadas para treinamentos.

Na simulação com o foco em treinamento, observamos 3 categorias principais [39]. A primeira é simulação viva, quando pessoas reais participam em ambientes reais de uma simulação, e os resultados de suas interações ocorrem por meio de sistemas auxiliares. Neste primeiro tipo, o propósito está relacionado com a possibilidade de se obter uma experiência mais próxima da realidade.

Já a segunda é a simulação virtual, onde usuários reais participam de uma simulação em um ambiente simulado, onde o ambiente não é real. Neste tipo de treinamento, o propósito está mais relacionado com o aprimoramento de habilidades motoras. Este tipo de simulação utiliza, normalmente, computadores e óculos de realidade virtual. Como exemplo, temos os treinamentos que utilizam óculos de realidade virtual focados em treinamentos de procedimentos [40].

A terceira e última é a simulação construtiva, onde tanto jogadores como o ambiente são simulados. Neste caso, os usuários participam das simulações como decisores, aprimorando o processo de decisão e comunicação. Temos como exemplos nesta categoria os Jogos de Guerra (JG) [41].

JG são definidos como um modelo que envolve pessoas que tomam decisões em um ambiente sintético de competição ou conflito, no qual elas veem os efeitos de suas decisões nesse ambiente e, em seguida, reagem a essas mudanças [42]. Por ser uma simulação focada em treinamento, pode ser utilizada tanto no ramo militar quanto empresarial. No CFN, o SJD é um simulador categorizado neste tipo de simulação.

2.4.1

Sistema de Jogos Didáticos

O SJD é o simulador construtivo do CFN. Ele teve sua concepção em 1989 e foi desenvolvido em 1992, juntamente com a Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio), através do Instituto TecGraf. Ele é

utilizado até hoje, sendo atualizado e mantido pelos militares do Centro de Simulação do Corpo de Fuzileiros Navais (CSimCFN).

Assim como nos simuladores construtivos existentes, tanto as peças de manobra, que são as tropas dispostas no terreno, como o ambiente, são simulados. Podemos observar, na interface que está na fig. 2.5, a disposição destas peças durante a condução de um JG utilizando o sistema. Como possibilidades de visualização das tropas no terreno, destacamos a área de operações por meio de imagens de satélite, de imagem 3D, da carta militar e da CT.

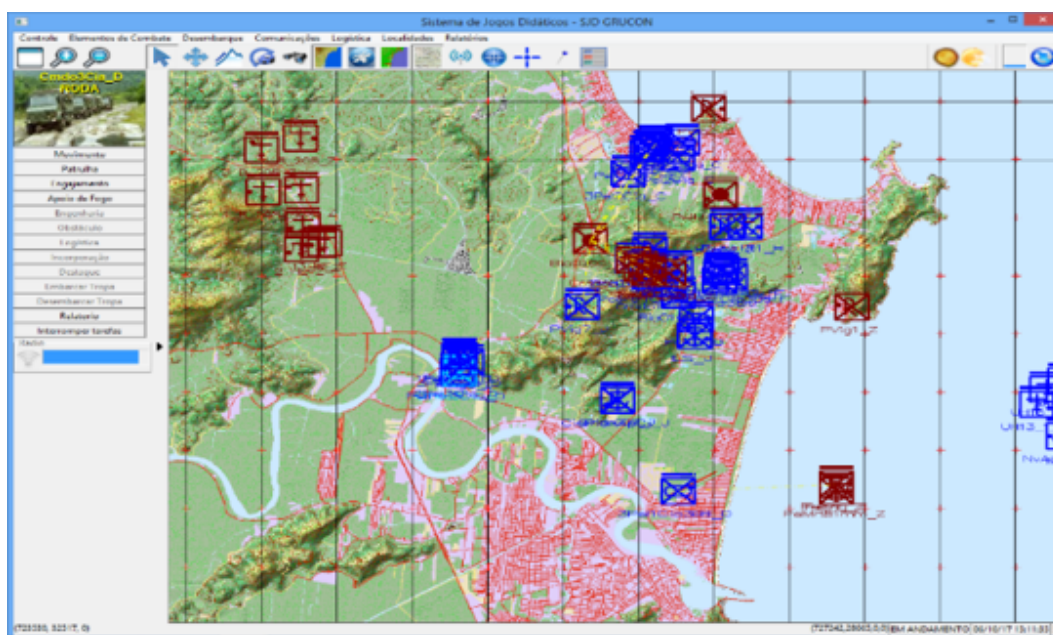


Figura 2.5: Sistema de Jogos Didáticos exibindo simbologias das forças amigas, em azul, e inimigas, em vermelho.

Seu propósito está em auxiliar os instrutores dos diversos cursos de carreira dos Oficiais do CFN na avaliação do processo de tomada de decisão dos OfA. Atualmente, vem sendo utilizado em JG dos Cursos de Aperfeiçoamento Avançado de Oficiais Fuzileiros Navais (C-ApA-CFN) e de Estado-Maior para Oficiais Intermediários (C-EMOI).

Além do fim didático descrito acima, ele é utilizado também no teste dos planejamentos dos exercícios realizados pelo setor operativo, a saber, a FFE (subseção 2.1). Este teste cresce em importância à medida que as atividades realizadas pela FFE envolvem posteriormente o emprego de tropas reais no terreno, de forma a cumprir as missões cujo planejamento foi testado no SJD antes. O parágrafo seguinte descreverá a sua utilização.

Antes do exercício propriamente dito, o EM recebe uma missão fictícia a ser cumprida em uma determinada região. Esta região normalmente é uma área de adestramento, cujo terreno já está mapeado e inserido nos simuladores.

Para o cumprimento desta missão, o EM é responsável pela confecção do planejamento, devendo seguir todas as etapas do PPM.

Quando este planejamento está pronto, ele é carregado no simulador, observando os locais corretos onde as tropas serão desembarcadas. Este desembarque pode ser por meio de lançamento de paraquedistas, por desembarque de helicóptero em pontos estratégicos, por desembarque em uma praia, entre outros.

Após o carregamento dos meios, os usuários executam este planejamento através de JG, colocando em prática o plano confeccionado, no ambiente simulado do SJD. Para a condução do exercício, os instrutores farão o papel de inimigos e avaliadores, explorando as falhas no planejamento, de forma a enriquecer a aprendizagem.

Para a condução correta do exercício, a CT do SJD é essencial. Como ela segue as conformações da região mapeada, é a responsável por limitar ou permitir os movimentos das forças amigas e inimigas. Sendo assim, os usuários perceberão as dificuldades atinentes ao uso correto do terreno no cumprimento da missão recebida. Como esta CT é elaborada analisando o terreno, ela pode ser obtida como um dos produtos do Estudo preliminar de Inteligência, mencionados na seção 1.1.

Neste contexto, uma possibilidade de uso deste simulador está no planejamento de operações reais, podendo ser utilizado na fase do Exame da Situação do Comandante (subseção 2.2), onde as LA elaboradas para o cumprimento da missão são testadas individualmente por meio do confronto. Porém, um fator dificultador para o uso do SJD no teste dos planejamentos da FFE com foco em missões reais está no tempo necessário para a confecção de uma carta de trafegabilidade para o uso no simulador.

Uma área de operações, a depender do contexto da missão, pode ter cerca de 60km² ou mais. Baseando-se nesta informação, uma carta de trafegabilidade deste tamanho, criada para o SJD, leva cerca de 2 a 3 meses para ficar pronta. Isto se deve à complexidade no mapeamento dos diferentes tipos de áreas e detalhamentos existentes. Um exemplo da carta de trafegabilidade utilizada no SJD está na fig. 2.6.

Verificando soluções para se obter estas regiões em tempo adequado, uma opção seria a utilização de imagens de drones, mas essas imagens demandam um longo tempo para aquisição, bem como grande capacidade computacional para se processar todas as imagens. Uma opção viável está no uso de imagens de satélites, obtidas por meio de sensoriamento remoto.

2.5

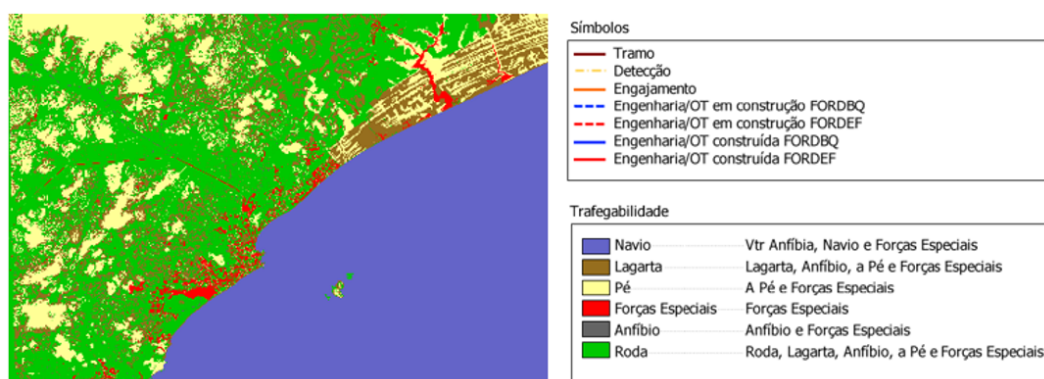


Figura 2.6: À esquerda, um exemplo de CT usada no SJD. À direita, em cima, legendas das ações executadas pelas peças de manobra. À direita, em baixo, as trafegabilidades existentes na região da carta.

Uso de sensoriamento remoto no estudo do terreno

Observamos que as imagens de satélite são uma fonte importante de informações. As imagens de satélite só ganharam atenção da comunidade recentemente, fruto do seu uso para composição de mapas, análise populacional, agricultura de precisão eficaz, tarefas de direção autônoma, entre outros [43].

As Técnicas de Sensoriamento Remoto (TSR) permitiram uma maior compreensão da complexa interação entre diferentes fenômenos geológicos e geomorfológicos. Temos como algumas TSR os o Sensoriamento Remoto Passivo (SRP) e o Radar de Abertura Sintética (RAS), que são utilizados em diversos tipos de situações. Entre elas destacamos a compilação de bancos de dados de deslizamentos de terra e para medir e quantificar a deformação do solo induzida pela ocorrência de deslizamentos de terra [18].

As Técnicas de SRP medem a energia que está naturalmente disponível. Normalmente, só pode ocorrer durante o período em que as fontes naturais de energia estão disponíveis e em quantidade grande o suficiente para ser registrada. Como exemplos, podemos citar o sol, quando este está iluminando a Terra, e o infravermelho térmico, que pode ser detectado de dia ou de noite [44].

Satélites com RAS, diferentemente do SRP, são sensores ativos e fornecem sua própria fonte de energia para iluminação. Como resultado, eles são capazes de obter medições a qualquer momento, independentemente da hora do dia ou da estação, usando bandas do domínio de micro-ondas.

Uma imagem RAS é composta de pixels correspondentes a uma área de solo, que por sua vez corresponde a uma metragem quadrada, chamada de resolução espacial. A resolução espacial das imagens RAS depende do sensor usado e de seu modo de aquisição. A resolução temporal depende do tempo de revisitação do satélite na mesma região, e varia de várias semanas (constelação

Radarsat) a alguns dias (constelação Sentinel-1).

A depender do tipo de satélite, o produto obtido engloba imagens geradas com base em subdivisões do espectro eletromagnético recebidas. Cada subdivisão obtém um tipo de resposta ao incidir em determinado tipo de objeto. O espectro eletromagnético se estende desde comprimentos de onda muito curtos associados aos raios cósmicos, passando por ondas de rádio de baixa frequência, até grandes comprimentos de onda [45]. Observamos na fig. 2.7 um desenho mostrando as regiões do espectro eletromagnético. Estas imagens podem nos fornecer informações baseadas nas combinações das regiões deste espectro, podendo destacar o nível de índices de vegetação, detecção de pragas entre diversos outros índices.



Figura 2.7: O espectro eletromagnético. Fonte: <http://imagers.gsfc.nasa.gov/ems/waves3.html>.

O uso destas imagens RAS é importante pois, a depender do tempo em que o satélite passa por uma determinada região, é possível obter imagens de poucos dias do evento ocorrido. Após obtidas as imagens, uma maneira de se reduzir o tempo de vetorização dessas regiões seria por meio da utilização de técnicas de Inteligência Artificial, a saber, segmentação semântica. Com isso, utilizando um modelo treinado e uma arquitetura eficiente, informações importantes para o planejamento de uma operação militar seriam extraídas rapidamente, com a consequente disponibilidade da CT, do terreno virtual e do terreno para o SJD.

2.6

Inteligência Artificial

Nesta seção versaremos um pouco sobre a história da Inteligência Artificial (IA) e seus conceitos básicos. Posteriormente, falaremos sobre os diferentes tipos de IA, conduzindo nossos estudos para as técnicas de Deep Learning (DL).

2.6.1

Surgimento da Inteligência Artificial

O primeiro pensamento sobre IA surgiu no século XVII através de conceitos do Filósofo e Matemático Gottfried Leibniz. Ele idealizou uma linguagem chamada *characteristica universalis*, capaz de explicar o pensamento racional de uma maneira tão precisa que uma máquina poderia ser criada para replicá-la. Esta máquina seria chamada de cálculo ratiocinator [46].

O nascimento da IA efetivamente aconteceu por meio de 2 eventos. O primeiro, tendo como idealizador Alan Turing, o pai da computação, em 1950. Ele criou o chamado teste de Turing para determinar se um computador pode ser considerado inteligente, e publicou como artigo seminal na revista *Mind* [47]. Seu teste é uma das referências amplamente usadas na IA.

Já o segundo evento considerado como sendo o nascimento da IA foi o Dartmouth Summer Research Project on Artificial Intelligence [48]. Esta conferência teve como conjectura a premissa de que os aspectos do aprendizado ou qualquer outro recurso da inteligência podem, em princípio, ser descritos com tanta precisão que uma máquina pode ser criada para simulá-los.

A primeira rede neural foi concebida por Warren McCulloch e Walter Pitts em 1943. Seu artigo introduziu a ideia de rede neural artificial, bem como elucidou algumas das definições que tomamos como certas hoje [49]. O próprio artigo abordou o problema das redes neurais como um problema lógico, partindo de definições, passando por lemas e teoremas.

Em seu trabalho, eles dividiram os neurônios em dois grupos. O primeiro, chamado hoje de “neurônios de entrada”, e o restante, neurônios de saída. Neste estudo, não foi introduzida a camada oculta, pois esta só foi empregada nas décadas de 1970 e 1980. Neste contexto, os neurônios poderiam estar em dois estados, disparando e não disparando, em um determinado momento [46].

Por volta de 1956, Frank Rosenblatt descobriu a regra de aprendizado do perceptron, uma regra que rege como atualizar os pesos das redes neurais. Além disso, entre os trabalhos publicados, explorou a ideia de redes de várias camadas semelhantes às redes convolucionais modernas, abrindo caminho para conceitos empregados hoje em Deep Learning.

Ao longo dos anos, existiram diversas outras descobertas na área, como o perceptron de Minsky e Papert em 1969. Em que pese o período de inverno da IA de 1974 a 1980 [50], uma série de evoluções ocorreram, tais como arquiteturas de redes neurais melhor elaboradas, estudos mais aprofundados a cerca dos tipos de redes neurais a serem empregadas, entre outros. Alguns desses conceitos veremos na subseção a seguir.

2.6.2

Conceitos básicos

O sistema nervoso humano pode ser visto como um sistema bastante complexo, e possui como centro o cérebro. Ele é representado pela rede neural, que continuamente recebe informações, percebe-as e toma as decisões rápida e eficientemente. Para se ter uma ideia, o cérebro possui cerca de 10 bilhões de neurônios, e cada neurônio possui entre 1.000 a 10.000 conexões. Neste sentido, cada pessoa pode dedicar 100.000 conexões para armazenar cada segundo de experiência, o que torna o cérebro com grande capacidade de paralelismo [51].

Já os neurônios são células do sistema nervoso consideradas como unidades básicas, e estão diretamente relacionadas com a propagação do impulso nervoso [52]. Ele é constituído pelo núcleo da célula, pelos dendritos, que são prolongamentos semelhantes a galhos de árvores, e pelos axônios, sendo estes revestidos pela bainha de mielina. Observamos entre os neurônios uma pequena distância, chamada de sinapse, onde o neurotransmissor é liberado, permitindo o estímulo e a passagem do impulso nervoso. Observamos uma representação do neurônio humano na fig. 2.8.

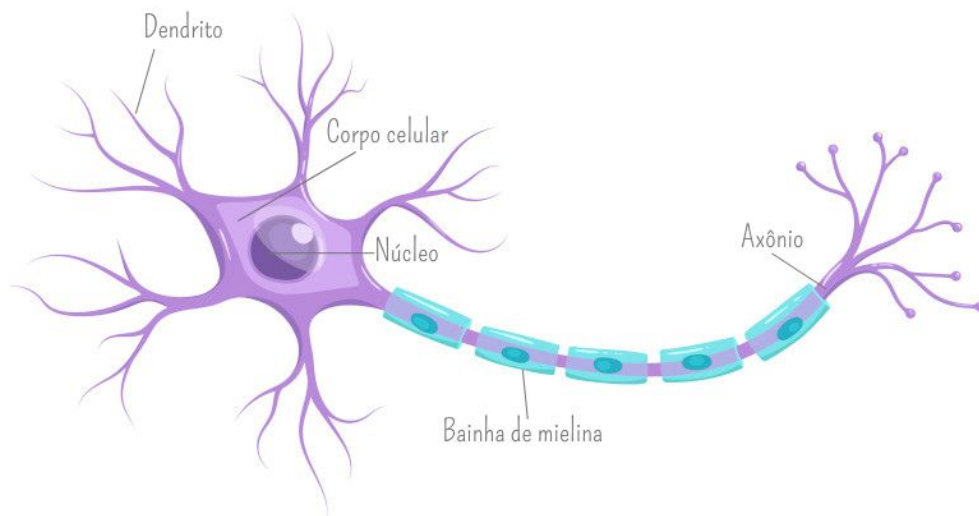


Figura 2.8: Neurônio encontrado no cérebro humano. Fonte: <https://escolakids.uol.com.br/ciencias/neuronios.htm>.

O ser humano, usando componentes eletrônicos ou software em um computador digital, desenvolveu a Rede Neural Artificial (RNA), de forma a tentar modelar a maneira pela qual o cérebro executa uma determinada tarefa ou função de interesse. Uma RNA é composta por neurônios artificiais, que armazenam conhecimento experimental e o disponibilizam para uso. O conhecimento adquirido ocorre também por meio do aprendizado [51], que

consiste nos ajustes dos pesos sinápticos. Observamos um neurônio artificial na fig. 2.9.

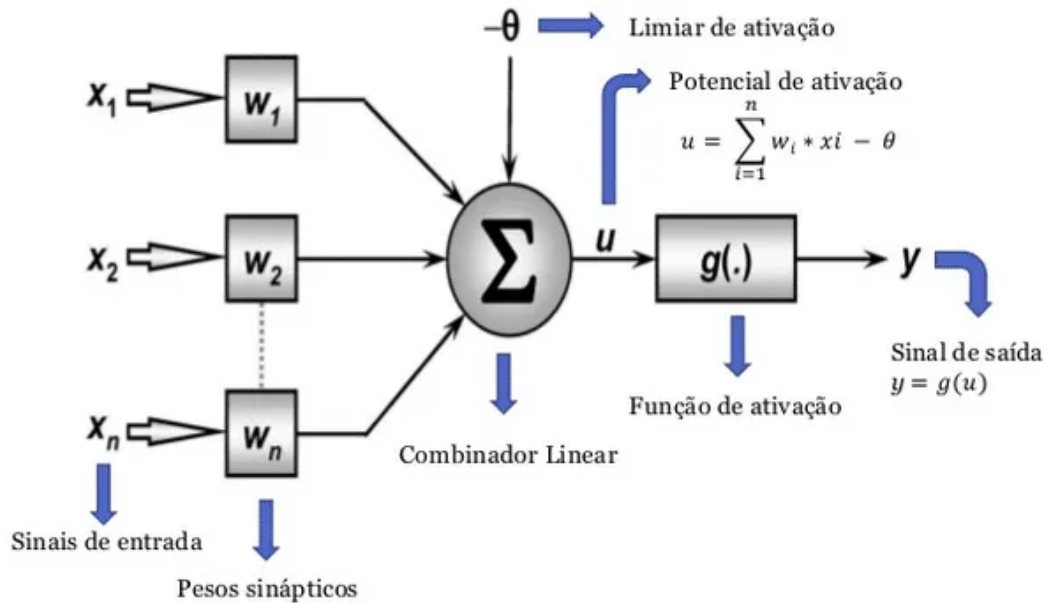


Figura 2.9: Neurônio artificial. Fonte: <https://medium.com/@avinicius.adorno/redes-neurais-artificiais-418a34ea1a39>.

Podemos observar na fig. 2.9 que o neurônio artificial possui sinais de entrada, pesos sinápticos, um combinador linear, um limiar de ativação ou bias, uma função de ativação e um sinal de saída. Todos esses elementos são combinados de forma similar às partes de um neurônio.

Os sinais de entrada (SE) são informações que o neurônio recebe para processar. A depender da origem das informações, podem vir diretamente do dataset, quando os neurônios estão na camada de entrada, ou de outros neurônios, quando são o sinal de saída de neurônios de camadas anteriores.

Os neurônios recebem os SE conforme observados na fig. 2.9, tendo origem em diversos outros neurônios. Neste sentido, os pesos sinápticos servem para regular a intensidade que o sinal recebido deve ser considerado para aquele neurônio específico. Os sinais são então combinados por meio do combinador linear, que consiste no somatório do produto entre os SE e os pesos sinápticos. A este somatório é adicionado um limiar de ativação ou bias, de forma a ajustar os valores iniciais do combinador linear.

Após a combinação linear realizada, uma função de ativação é aplicada no resultado. Esta função serve para limitar o sinal de saída do neurônio, impedindo valores extremamente altos que causarão desbalanceamento da RNA de maneira geral. Este estado final depende do tipo de função de ativação implementada e pode entregar valores inteiros ou reais entre 0 e 1, -1 e 1.

Existem vários tipos de funções de ativação que influenciam diretamente no tipo de sinal de saída que se deseja.

Os neurônios artificiais podem ser combinados de diversas maneiras, de forma a compor a RNA. Sua estrutura pode ser composta por 1 camada, sendo esta a camada de saída, composta por múltiplas camadas, sendo as camadas internas chamadas de camadas escondidas, e composta por uma rede recorrente, com capacidade de processar e obter informações sobre ajustes dos pesos sinápticos na mesma camada ou camadas anteriores (fig. 2.10).

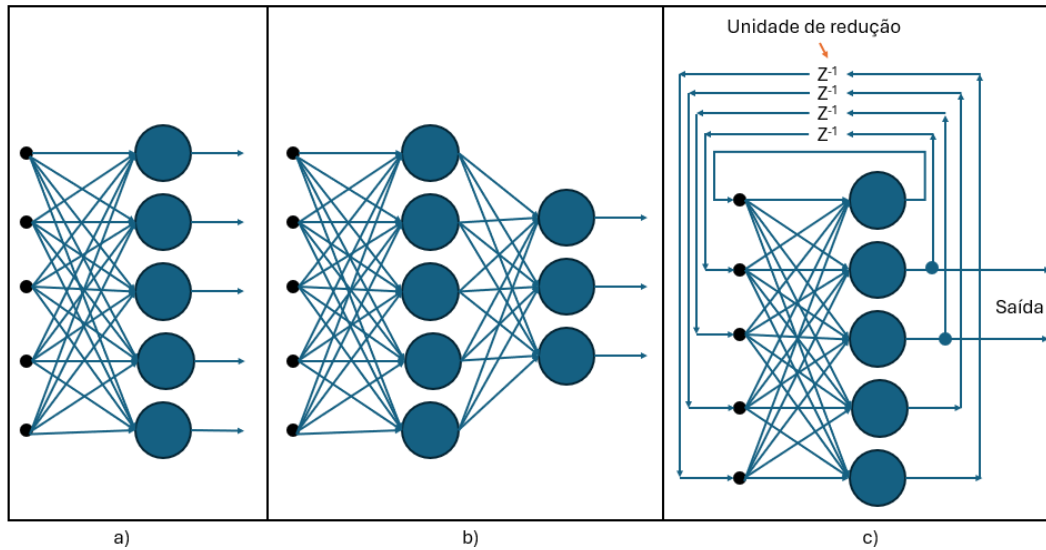


Figura 2.10: Estruturas de redes neurais artificiais. Em a) uma rede de 1 camada, em b) uma rede multi-camada e em c) uma rede recorrente [51].

Em uma estrutura de redes neurais podem existir vários neurônios artificiais. Nestes neurônios, uma das poucas variáveis passíveis de sofrer alteração são os pesos sinápticos, que influenciam diretamente no valor dos sinais de entrada recebidos. O procedimento usado para modificar os pesos sinápticos da rede de forma ordenada para atingir um objetivo de projeto desejado chama-se aprendizagem. A modificação dos pesos sinápticos é um dos métodos tradicionais para um projeto envolvendo redes neurais [51].

Após o período de aprendizagem, chamado de treinamento, os valores dos pesos obtidos são armazenados e utilizados para prever resultados em imagens diferentes das utilizadas anteriormente. A produção, pela rede neural, de resultados razoáveis para entradas não encontradas durante o treinamento chamamos de generalização.

No ramo da IA, existem várias abordagens para os inúmeros problemas existentes. Na programação clássica da IA, os humanos inserem regras, tais como valores dos pesos sinápticos vistos anteriormente, por exemplo, e dados a serem processados, obtendo respostas de acordo com as regras criadas.

A Aprendizagem de Máquina (do inglês Machine Learning - ML) começou a crescer na década de 1990. Ela se tornou o subcampo mais popular e mais bem-sucedido da IA, sendo impulsionada pela disponibilidade de hardware mais rápido e conjuntos de dados maiores [54]. Neste contexto, os dados de entrada e as respostas esperadas são inseridos, obtendo como resposta as regras necessárias. Sendo assim, uma rede neural é efetivamente treinada em vez de explicitamente programada. A subseção a seguir versará sobre uma das técnicas de ML que será um dos focos do nosso trabalho, chamada Deep Learning (DL).

2.6.3 Deep Learning

DL é uma estrutura matemática para aprender representações a partir de dados. Esta técnica de ML ganhou destaque no início da década de 2010. Nos poucos anos que se passaram desde então, ela conseguiu resultados notáveis em problemas de percepção, como visão e audição, que envolvem habilidades naturais e intuitivas para os seres humanos. Em particular, DL alcançou grandes avanços na classificação de imagens, reconhecimento de fala, transcrição de escrita à mão, tradução automática entre diversas outras pesquisas [54].

Este é um subcampo específico da ML que enfatiza a aprendizagem de camadas sucessivas e de representações cada vez mais significativas (fig. 2.11). O termo “profundo” representa a ideia de camadas sucessivas de representações. A quantidade de camadas que contribuem para um modelo de dados é chamada de profundidade do modelo [54].

As Redes Neurais Convolucionais (RNC) revolucionaram o campo da DL, mostrando-se particularmente eficazes em tarefas como a classificação de imagens. Sua aparição ocorreu em 2012 no ImageNet Large Scale Visual Recognition Challenge (ILSVRC) com uma arquitetura chamada de AlexNet [55]. Esta estrutura de modelo aprimora sua capacidade de aprender padrões complexos em imagens e, ao mesmo tempo, emprega técnicas de ativação e abandono de Rectified Linear Unit (ReLU) para evitar o ajuste excessivo [56]. Ao longo dos anos, diversos modelos foram criados, entre eles podemos destacar o VGGNet em 2014, a ResNet em 2015 e a Inception em 2015 [57] [58].

Os modelos básicos da visão computacional evoluíram continuamente. Neste contexto, o método baseado em RNC UNet foi um marco na segmentação semântica, que estabeleceu a arquitetura fundamental para diversos trabalhos subsequentes. Ele é composto por uma estrutura de codificador-decodificador com conexões de salto em forma de U, melhorando significativamente a precisão da segmentação semântica de imagens de sensoriamento remoto de

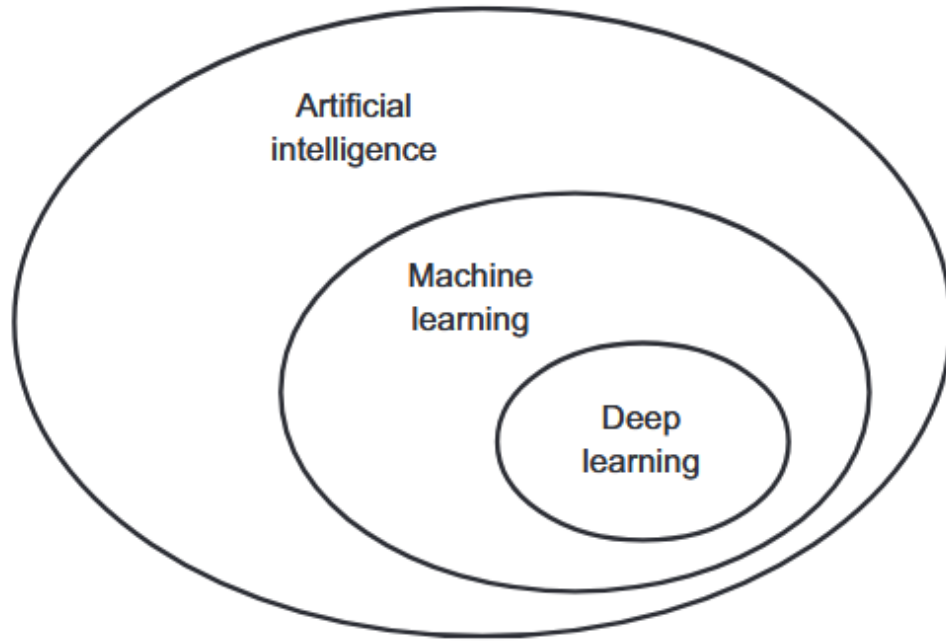


Figura 2.11: Deep Learning e Machine Learning são ramos específicos dentro da Inteligência Artificial [54].

alta resolução [59].

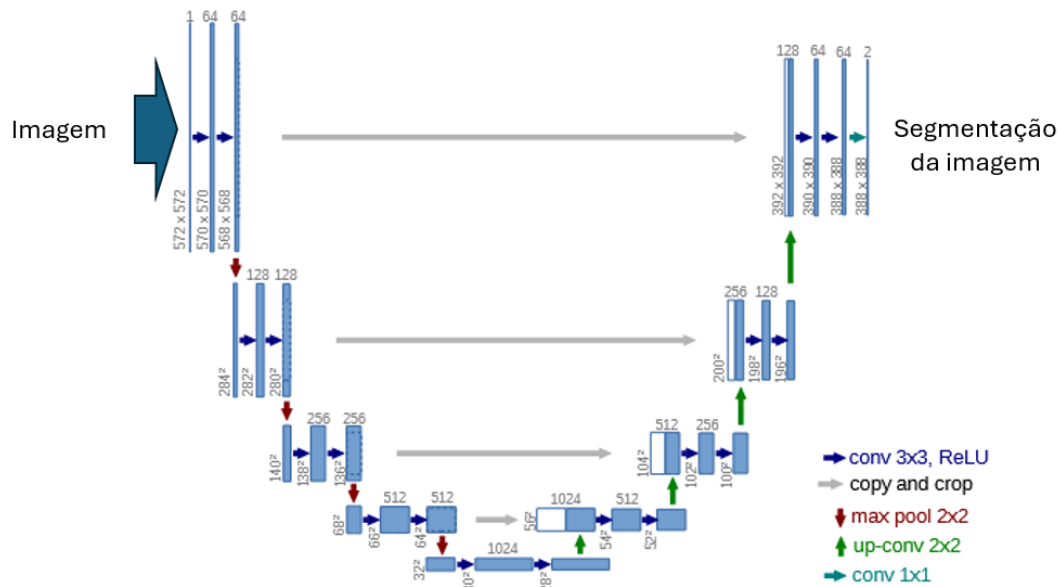


Figura 2.12: Arquitetura U-net (exemplo para 32x32 pixels na resolução mais baixa). Cada caixa azul corresponde a um mapa de recursos multicanal. O número de canais é indicado na parte superior da caixa. O tamanho x-y é fornecido na borda inferior esquerda da caixa. As caixas brancas representam mapas de recursos copiados. As setas indicam as diferentes operações [60].

Na fig. 2.12, podemos observar 3 tipos de camadas, chamadas de camada de convolução (conv 3x3, ReLU), camada max pooling (max pool 2x2) e

up-convolution (up-conv 2x2). A depender da arquitetura utilizada, tanto as camadas de compressão, que utilizam a convolução, como as camadas de descompressão, que utilizam as camadas de up-convolution, podem ser apresentadas de maneiras diferentes [61].

A RNC é um tipo de rede neural que não é recorrente, capaz de extrair recursos de dados com estruturas de convolução. A convolução é realizada por meio de filtros, conhecidos como kernels, que são aplicados à imagem de entrada, onde é possível extrair dados da imagem através da criação de mapas conhecidos como feature maps [58].

Neste contexto, as funções de perda e os otimizadores foram criados para obter os erros de previsão e acelerar as correções a serem realizadas, de forma que todo o sistema RNC aprenda o que se espera [58].

Além das camadas convolucionais, a camada de pooling aproveita o princípio da correlação local da imagem para reduzir a dimensionalidade da amostragem e reter informações úteis. Porém, dada a redução da imagem, ela também pode reduzir o número de parâmetros, removendo recursos que facilitem a aquisição de dados.

Os resultados da convolução podem ser chamados de mapas de recursos. Ao definir um núcleo de convolução com um determinado tamanho, perderemos informações na borda. Por isso, o preenchimento é introduzido para aumentar a entrada com valor zero, o que pode ajustar o tamanho indiretamente [58].

Além disso, para controlar a densidade da convolução, emprega-se o stride. Quanto maior o passo, menor a densidade. Após a convolução, os mapas de recursos consistem em um grande número de recursos que são propensos a causar problemas de sobreajuste. Como resultado, o pooling, também conhecido como down-sampling, é proposto para evitar a redundância, incluindo o pooling máximo e o pooling médio [58].

RNC podem ser utilizadas hoje em diferentes tipos de aplicações. Entre elas, destacamos a detecção de objetos, utilizadas para identificar objetos em uma cena sem obter sua posição detalhadamente. Uma outra aplicação está na medicina, através da segmentação semântica, onde é possível saber exatamente as dimensões e localização de um objeto em questão, que pode ser uma lesão, um tumor observado em uma ressonância magnética ou até mesmo uma célula em uma imagem de microscópio eletrônico [60].

Dadas as vantagens da segmentação semântica, seu emprego vai além da medicina. Hoje, observamos sua utilização em pesquisas relacionadas à observação do planeta Terra, utilizando imagens satélites, assunto a ser abordado na próxima seção.

2.7

Deep Learning no sensoriamento remoto

As técnicas de Sensoriamento Remoto (SR) estão se tornando cada vez mais importantes nas tarefas de coleta de dados. Serviços de localização, como Google Earth e Governos utilizam SR para uma variedade de serviços públicos. Estes serviços vão desde relatórios meteorológicos, monitoramento de tráfego, identificação de mudanças com o foco em áreas desmatadas entre diversas outras atividades [62].

As imagens de RS, em que pesem a resolução espectral e espacial, são reflexos da superfície terrestre. Sendo assim, uma propriedade importante é sua capacidade de registrar informações em várias escalas dentro de uma área. De acordo com o tipo de informação desejada, podem ser extraídos recursos baseados em pixels, como objetos ou estruturas. Porém, uma dificuldade que ainda existe é a falta de uma abordagem eficaz e universal para fundir esses recursos de forma ideal, sendo prejudicada pelas relações sutis entre os dados.

Em contrapartida, a tecnologia digital pode representar e organizar vários níveis de informações para expressar relações complexas entre os dados. De fato, as técnicas de DL podem mapear diferentes níveis de abstrações das imagens e combiná-las de baixo nível a alto nível.

Deve-se considerar ainda o fato de que o uso de dados de vários sensores de satélite pode incidir em resultados de pesquisa diferentes, já que eles geralmente têm resoluções espaciais diferentes e podem diferir em relação ao período do ano que sua imagem foi extraída. Portanto, é importante considerar a configuração espacial de qualquer padrão de paisagem, considerando as resoluções espaciais de imagens de satélite [63].

A resolução espacial tem sido outro foco dos estudos de SR. É necessário estimar a capacidade dos dados de sensoriamento remoto no mapeamento de paisagens, pois a aplicação do sensoriamento remoto pode ser limitada por sua resolução espacial. Neste sentido, imagens com resolução mais fina contêm maior quantidade de informações espaciais, o que, por sua vez, permite uma melhor caracterização de características menores.

Neste sentido, DL está em toda parte na análise de dados de SR. Esta técnica pode ser fortemente empregada desde os tópicos tradicionais de pré-processamento de imagens, classificação baseada em pixels, como segmentação semântica, e reconhecimento de alvos, até as recentes tarefas desafiadoras de extração de recursos e compreensão de cenas.

No entanto, a forma de utilizar essas informações ainda requer mais pesquisas. Ainda existem problemas nas imagens SR de alta resolução espacial, que têm apenas canais verde, vermelho e azul, o mesmo que os conjuntos de

dados de referência para DL. Na prática, há muito poucas amostras rotuladas disponíveis, o que pode dificultar a construção de uma rede pré-treinada [62].

As técnicas de DL podem mapear diferentes níveis de abstrações das imagens e combiná-las de baixo nível a alto nível. Por exemplo, considerando o reconhecimento de cenas, com a ajuda da DL, as cenas podem explorar as variações nos arranjos espaciais locais e nos padrões estruturais capturados, onde nenhum estágio de segmentação ou de extração de objetos individuais é necessário.

Outro emprego está no uso de técnicas de segmentação semântica para obter informações como, por exemplo, deslizamentos de terra em áreas atingidas por desastres naturais. Neste caso, a identificação rápida e precisa das faixas de impacto de desastres de deslizamentos de terra a partir de grandes quantidades de dados de imagens de sensoriamento remoto é essencial para orientar a prevenção e a mitigação de desastres [64].

2.7.1

Segmentação semântica no sensoriamento remoto

Esta seção discutirá sobre a abordagem adotada para se identificar os diferentes tipos de trafegabilidade em uma imagem satélite. Cabe destacar aqui que as elevações, normalmente obtidas através do MDE descrito na seção 6.1, não serão, ainda, consideradas, haja vista as diferentes maneiras de serem utilizadas após a obtenção das trafegabilidades descritas neste capítulo.



Figura 2.13: Comparação entre parte de uma carta de trafegabilidade e uma imagem satélite. As áreas urbanas, estradas e rios correspondem à trafegabilidades específicas segundo as cores estabelecidas.

Ao compararmos uma CT com uma imagem satélite, conforme mostrado na Figura 2.13, percebemos que cada tipo de área, seja ela uma vegetação mais densa, uma área urbana ou um rio sinuoso, corresponde a um tipo de trafegabilidade, sendo associada a uma cor. Neste trabalho, trataremos cada tipo de trafegabilidade como um determinado tipo de classe, haja vista o

tipo de área (floresta, plantação ou rio) restringir um determinado tipo de mobilidade.

Para tal, o modelo de IA deve identificar o posicionamento correto e a disposição das manchas das classes na imagem, de forma a evitar erros nos planejamentos. Esta demanda pode ser atendida pelo conceito de segmentação de imagens, muito utilizado na medicina para identificar regiões com câncer ou identificação de estruturas celulares [60].

Com isso, a intenção deste trabalho é a utilização de segmentação de imagem, mais especificamente segmentação semântica. Este tipo de abordagem considera a separação apenas dos tipos de classes, sem a necessidade de distinção específica dentro da classe.

A segmentação semântica consiste no particionamento da imagem em várias partes chamadas segmentos, evitando o processamento de toda a imagem de uma vez. O processamento de uma imagem inteira, sem o processo de segmentação, incorreria no processamento de regiões sem informação, consumindo tempo excessivo de processamento. Ao dividir a imagem, podemos usar os segmentos importantes para processá-la [85].

Um exemplo de arquitetura de um modelo de IA a ser utilizada na segmentação semântica é mostrado na Figura 2.12. Ela consiste inicialmente em um caminho de compressão ou encoding, seguido de um caminho de descompressão ou decoding.

O caminho de compressão segue a arquitetura típica de uma rede convolucional. Ele consiste na aplicação repetida de duas convoluções 3x3 (convoluções sem preenchimento), cada uma seguida por uma Unidade Linear Retificada (sigla em inglês, ReLU) e uma operação de pooling máximo 2x2, com passo 2 para downsampling. Ao final, obtemos um mapa de recursos detectados na imagem original.

Cada convolução envolve o uso de um filtro, que possui um determinado tamanho em pixels como, por exemplo, 3x3, 5x5 ou 7x7. Estes filtros servem para detectar recursos específicos em uma imagem, como bordas ou formas. Como observado na Figura 2.14, o filtro desliza sobre a imagem de entrada, realizando a multiplicação por elementos dos valores de pixel correspondentes. O tamanho da imagem de entrada é reduzido durante a convolução, o que é útil para um processamento mais rápido, mas também resulta em alguma perda de informação.

Na figura 2.15 observamos um exemplo de aplicação de um filtro horizontal na imagem de um gato, extraíndo seus contornos. A cada camada de convolução, mais características das imagens são extraídas, bem como mais filtros são aplicados, possibilitando obter o mapa de recursos pela rede.

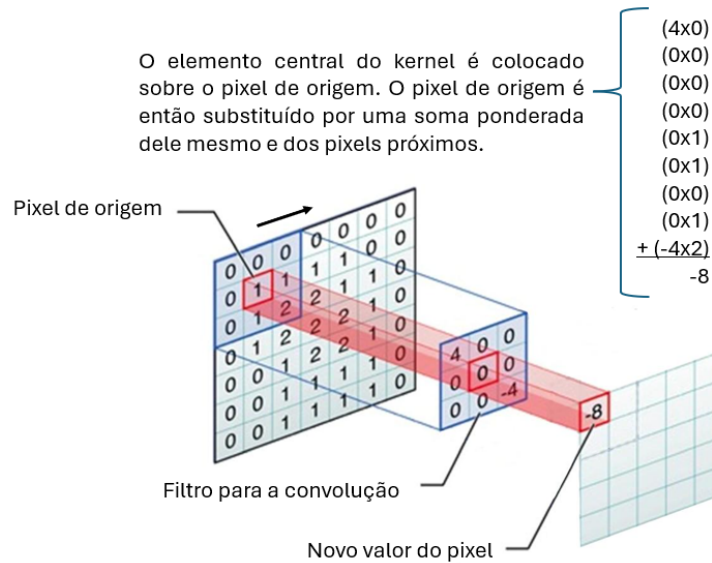


Figura 2.14: Processo de convolução, com filtro de 3x3, aplicado em uma matriz. Fonte: <https://www.kaggle.com/code/akshitsharma1/unet-architecture-explained-in-one-shot-tutorial>.

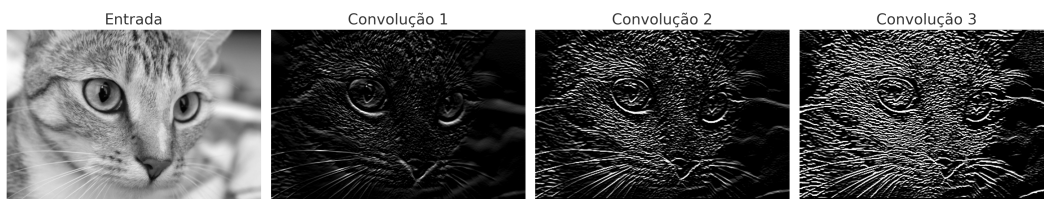


Figura 2.15: Processo de convolução em 3 camadas consecutivas, com filtro horizontal de 3x3, aplicado à imagem de um gato. Fonte: imagem gerada por Inteligência Artificial.

Cada camada de agrupamento ou pooling inserida na rede é responsável por condensar a imagem conforme a região estabelecida, diminuindo o tamanho desta e, com isso, o tempo computacional necessário. Os resultados são imagens mais abstratas, porém fornecendo maior contextualização para o modelo.

Já a etapa do decoding consiste em um upsampling do mapa de recursos seguido por uma convolução 2x2 (“up-convolution”), que reduz pela metade o número de canais, e uma concatenação com o mapa de recursos cortado correspondente do caminho de contração e duas convoluções 3x3, cada uma seguida por uma ReLU.

O corte é necessário devido à perda de pixels de borda em cada convolução. Na camada final, uma convolução 1x1 é usada para mapear cada vetor de característica, composto por 64 componentes, para o número desejado de classes. A classificação acontece pixel a pixel.

Neste capítulo, os assuntos abordados servem como auxílio para o entendimento do problema proposto. Considerando como público-alvo tanto milita-

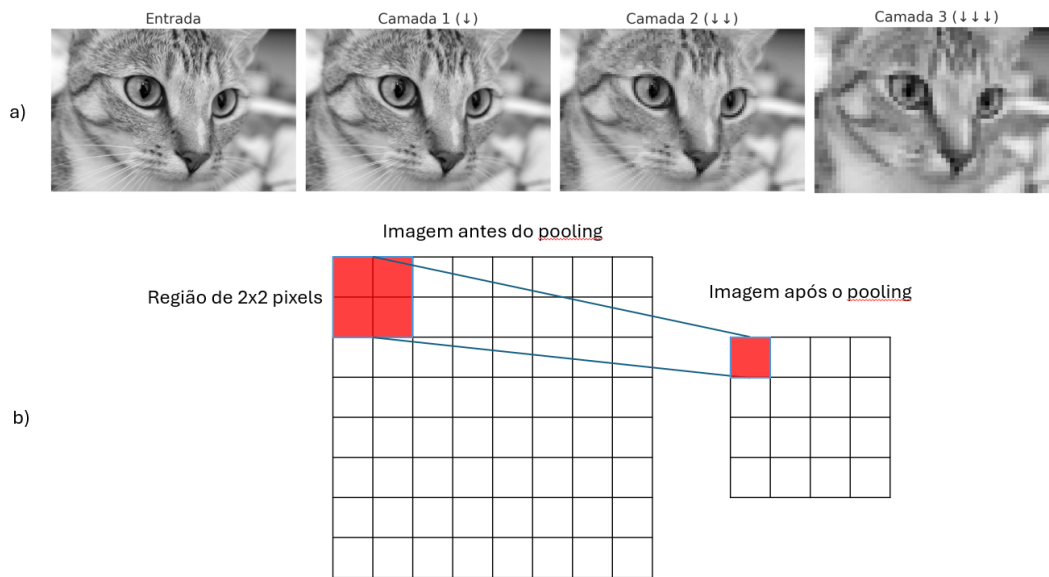


Figura 2.16: Em a), a imagem representa o processo de pooling em 3 camadas consecutivas aplicado à imagem de um gato. Nele, a resolução é reduzida (pooling) preservando os elementos importantes. Já em b), a imagem representa o processo de pooling usando uma região de 2x2 pixels. Fonte: imagens do gato geradas por Inteligência Artificial.

res como a academia, foram abordados assuntos de forma que suas definições conduzam para a solução proposta deste trabalho. Cabe relembrar aqui que não é o foco deste capítulo explicar toda a teoria, tendo em vista a complexidade dos assuntos aqui descritos. No próximo capítulo, veremos trabalhos relacionados que orientaram o desenvolvimento da solução para o problema proposto.

3

Trabalhos Relacionados

De forma a orientar nossas ações, buscamos na literatura trabalhos que pudessem prover um embasamento teórico e prático sobre segmentação semântica em sensoriamento remoto. Além disso, buscamos ainda maneiras de utilização do produto desta segmentação em simuladores, com vistas a facilitar o processo de análise do terreno. Sendo assim, os seguintes trabalhos acadêmicos nos forneceram subsídios importantes na busca de uma solução para o problema proposto.

O primeiro artigo, intitulado “High spatial resolution remote sensing image segmentation based on the multiclassification model and the binary classification model”, de autoria principal de Xiaoxiong Zheng e Tao Chen, apresenta um estudo sobre segmentação de imagens de sensoriamento remoto de alta resolução utilizando modelos baseados na arquitetura U-Net [65].

O trabalho faz uso do banco de dados Gaofen-2 [66], composto por 150 imagens multiespectrais com quatro bandas (azul, verde, vermelho e infravermelho próximo) e resolução espacial de $3.24m^2$. Foram categorizados 6 tipos de objetos no solo: built-up (áreas construídas), farmland (áreas agrícolas), water (corpos d’água), meadow (prados), forest (florestas) e Others (outros/sem classificação). Para treino e teste, as imagens foram divididas em blocos de 512×512 pixels, sendo 146 imagens para treinamento, totalizando 106.288 patches, e 4 imagens, totalizando 840 patches, para teste.

A metodologia compreende o treinamento de dois tipos de modelos U-Net: um para multiclassificação e outro para classificação binária (um modelo para cada classe). Aplicou-se um algoritmo de votação de vizinhança para resolver divergências de classificação pixel a pixel, melhorando a precisão ao reduzir ruídos. A avaliação utilizou métricas padrão da área como acurácia geral, coeficiente Kappa, média da interseção sobre união (mIoU), além de acurácia do usuário e do produtor.

Os resultados indicaram que o modelo de multiclassificação alcançou 82,27% de acurácia geral e coeficiente Kappa 0,7721 no conjunto de teste, enquanto o modelo de classificação binária obteve 79,75% de acurácia geral e Kappa 0,7377. A classe água foi a mais bem segmentada, com melhor consistência. O modelo multiclassificação mostrou melhor desempenho, porém com custo computacional superior, enquanto o modelo binário teve menor tempo de treinamento, sendo eficaz em diferentes contextos.

O segundo trabalho, "Semantic Segmentation of Remote Sensing Imagery

Using an Enhanced Encoder-Decoder Architecture", escrito por Aburaed et al. [67], apresenta uma abordagem para a segmentação semântica de imagens de sensoriamento remoto baseada em aprendizado profundo. O estudo propõe a adaptação do modelo Robust UNet (RUNet), originalmente concebido para super-resolução de imagens, para a tarefa de segmentação semântica, integrando um mecanismo de atenção utilizando o bloco Squeeze-and-Excitation (SE). Essa combinação resulta nas arquiteturas denominadas SE-RUNet e SE-UNet (quando o SE é aplicado ao UNet clássico).

Para avaliar o desempenho das arquiteturas propostas, os autores utilizaram o conjunto de dados "Semantic Segmentation of Aerial Imagery". Este banco de dados contém 72 imagens RGB da cidade de Dubai, com tamanhos variados e resolução espacial de $1m^2$. Suas características incluem cenários complexos com sombras, oclusões e estrutura urbana densa. As imagens foram divididas em patches de 160×160 pixels para facilitar o procedimento de treinamento do modelo proposto, totalizando 3483 imagens após o patching. As métricas para avaliação incluíram acurácia, precisão, recall, F-score e média do índice de interseção sobre união (mIoU).

Os resultados experimentais mostrados na Tabela 3.1 demonstraram que as versões com atenção (SE-RUNet e SE-UNet) superaram consistentemente suas contrapartes originais, com destaque para o SE-UNet, que obteve o melhor desempenho geral em todas as métricas. A análise visual dos mapas de segmentação confirmou essa superioridade, exibindo maior fidelidade na delimitação das classes, especialmente em regiões com objetos menores e bordas menos definidas.

Model	Accuracy	Recall	Precision	F-score	mIoU
UNet	0.8806	0.8769	0.8854	0.8812	0.4559
RUNet	0.8668	0.8642	0.8705	0.8673	0.4842
SAE-UNet	0.8849	0.8819	0.8891	0.8855	0.4900
SAE-RUNet	0.8706	0.8685	0.8736	0.8711	0.4781

Tabela 3.1: Resultados dos modelos considerando Accuracy, Precision, Recall, F-score, e mIoU.

Os autores concluíram que a incorporação do mecanismo de atenção SE eleva significativamente a capacidade de segmentação semântica das arquiteturas UNet e RUNet em imagens de sensoriamento remoto, apontando para o potencial dessas redes aprimoradas em aplicações complexas, como monitoramento ambiental, planejamento urbano e gestão de recursos.

O trabalho "What Data are needed for Semantic Segmentation in Earth Observation?", de Castillo-Navarro et al. [68], investiga os requisitos de dados para o aprendizado supervisionado em segmentação semântica aplicada a

imagens de sensoriamento remoto. O estudo aborda duas questões principais: a robustez das estratégias existentes de aprendizado supervisionado em relação ao volume e diversidade dos dados, e quais propriedades são desejáveis em um grande conjunto de dados para técnicas de observação da terra.

No contexto da crescente disponibilidade de imagens de sensoriamento remoto, os autores destacam a importância de mapas semânticos precisos para várias aplicações, como ecologia e planejamento urbano, e enfatizam o desafio da necessidade de grandes bases anotadas para treinar redes neurais profundas. O trabalho apresenta uma análise experimental que mostra que, para pequenos conjuntos de dados homogêneos, como o banco de dados Vaihingen, é possível obter bons resultados mesmo com poucos dados anotados, especialmente quando se utiliza transferência de aprendizado. Contudo, essa robustez reduz-se significativamente em conjuntos de dados mais diversos e realistas.

Para avançar na análise em larga escala, os autores introduzem o banco de dados MiniFrance, que é 2719 vezes maior que Vaihingen e reúne imagens aéreas de 16 cidades francesas com anotações de uso do solo em 15 classes, enfatizando a importância da diversidade geográfica e da generalização dos modelos.

Experimentos com MiniFrance mostram que aumentar a diversidade do banco de dados é crucial para obter modelos com maior capacidade de generalização, embora os desafios de anotações aproximadas e classes semânticas abstratas sejam evidentes.

Em conclusão, o estudo ressalta que, enquanto redes profundas são robustas a reduzir drasticamente o volume de dados em contextos homogêneos, o sucesso em larga escala depende da diversidade, qualidade e representatividade das anotações, propondo MiniFrance como um importante recurso para avançar a pesquisa em segmentação semântica em observação da terra. O trabalho ainda fornece uma importante base experimental e discute desafios futuros para algoritmos de segmentação em grandes e variados conjuntos de dados de sensoriamento remoto.

O terceiro trabalho, escrito por Castillo-Navarro et al., intitulado “What Data are needed for Semantic Segmentation in Earth Observation?”, investiga a quantidade e a qualidade dos dados necessários para a segmentação semântica supervisionada em imagens de sensoriamento remoto [68].

Neste contexto, os autores analisam a robustez dos modelos de aprendizagem profunda quando treinados com diferentes tamanhos e diversidades de conjuntos de dados. Apresentam o MiniFrance, um conjunto de dados projetado para promover o desenvolvimento de modelos de segmentação semântica mais generalizados e eficientes para observação da Terra, avançando em direção

a um recurso para aplicações de sensoriamento remoto.

Os autores explicam que, no contexto do sensoriamento remoto, a adaptação de domínio é essencial quando modelos treinados em dados de um local precisam ser aplicados a diferentes regiões ou sensores. Essa adaptação pode levar à redução de desempenho devido a diferenças ambientais, características do sensor ou outros fatores.

Dessa forma, o objetivo principal foi verificar quantos dados são necessários para treinar uma rede neural supervisionada voltada à segmentação semântica na observação da Terra. As contribuições envolvem uma análise experimental da quantidade e variabilidade de dados necessárias para o aprendizado supervisionado eficaz. Também destacam a constituição de um conjunto de dados em larga escala, de diversas imagens de sensoriamento remoto e anotações de múltiplas fontes, superando limitações presentes em conjuntos de dados padrão.

Para todos os experimentos, utilizou-se o SegNet como linha de base [69], apresentando uma arquitetura de codificador-decodificador, com o codificador equivalente ao VGG16 [70] e inicialização possível com pesos pré-treinados. O treinamento utiliza o algoritmo Stochastic Gradient Descent e, como função de perda, a standard cross-entropy aplicada aos pixels.

Em um primeiro teste, foi realizada uma análise do aprendizado supervisionado em conjuntos de dados contendo escala pequena, usando o banco de dados Semantic Labeling Vaihingen [71]. Este consiste em 33 blocos de infravermelho-vermelho-verde com uma resolução espacial de 9 cm/pixel e um tamanho médio de 2000×1500 pixels. As anotações foram separadas em 16 blocos para 6 classes de interesse: superfícies impermeáveis, edifícios, vegetação rasteira, árvores, carros e desordem.

No primeiro teste, os autores realizaram análises supervisionadas em conjuntos de dados de pequena escala, empregando o banco de dados Semantic Labeling Vaihingen [71]. Estes dados consistem em 33 blocos infravermelho-vermelho-verde, com resolução de 9 cm/pixel e tamanho médio de 2000×1500 pixels. As anotações foram separadas em 16 blocos para seis classes de interesse: superfícies impermeáveis, edifícios, vegetação rasteira, árvores, carros e desconhecido.

Foram realizados 2 experimentos, sendo o primeiro usando 12 blocos para treinamento e 4 blocos para validação, e o segundo utilizando apenas 1 bloco, enquanto as imagens de validação permaneceram inalteradas.

Duas configurações de treinamento foram comparadas: a primeira com pesos pré-treinados do VGG-16 no ImageNet; em seguida, repetiu-se com pesos inicializados aleatoriamente, considerando política de redução ou ampliação

exponencial das magnitudes dos sinais de entrada [72], permitindo verificar transferência de aprendizado.

Os resultados deste experimento são exibidos na Tabela 3.2. Observou-se uma mudança de 90% para 80% em OA e de 65% para 50% em mIoU ao reduzir o número de blocos de treinamento de 12 para 1.

Tabela 3.2: Desempenho da classificação com relação à quantidade de dados em pequena escala.

Train set	OA	mIoU
12 blocos	90.00	65.00
1 bloco	80.00	50.00

No segundo experimento, foi analisado o aprendizado supervisionado em grandes volumes de dados, utilizando o MiniFrance, que visa acrescentar diversidade aos bancos de dados de segmentação semântica.

O MiniFrance contém imagens aéreas de 16 cidades ou aglomerações urbanas de diferentes regiões da França, criadas com ortofotos de domínio público da Agência Nacional de Mapeamento da França, adquiridas entre 2012 e 2014, e do Copernicus Urban Atlas 2012. Foram consideradas quinze classes de uso da terra, incluindo áreas urbanas, indústrias, transporte, terras aráveis, pastagens, florestas, espaços abertos, áreas aquáticas e água.

Para a distribuição, 8 cidades foram utilizadas para treinamento e 8 para teste, mantendo a diversidade em termos de arquitetura e design urbano em ambos os subconjuntos. O total constitui 2121 imagens, cada uma delas com tamanho de 10.000×10.000 pixels e resolução espacial de 50 cm/pixel.

De forma similar aos experimentos com dados em pequena escala, foram efetuados treinamentos com todo o banco de dados e, em seguida, apenas com 10% das imagens deste, tendo sempre atenção à conservação de sua diversidade. Por fim, foi utilizada apenas a cidade de Caen para o treinamento, que representa 12,5% do total de imagens. O conjunto de teste permaneceu o mesmo. Seguindo os resultados de experimentos anteriores, os pesos são inicializados como pesos pré-treinados para a rede VGG-16.

Os resultados são mostrados na Tabela 3.3. Os desempenhos não estão atingindo o mesmo nível do experimento com dados de pequena escala. Entretanto, considerando o problema proposto pelos autores, eles destacam que o treinamento com todos os dados e com 10% levou a pontuações semelhantes, tanto em precisão quanto em IoU.

Ao escolher as imagens de amostra de 10% em todo o conjunto de dados, a diversidade do conjunto de treinamento foi preservada e os resultados não

foram degradados. O treinamento com um único local implicou uma perda de 10% na precisão e 5% menos de mIoU.

Os autores destacam ainda que o desempenho da rede variou notavelmente entre algumas aglomerações urbanas, revelando diferenças entre as cidades e uma falta de capacidade de generalização do modelo. Sendo assim, os autores afirmam que o banco de dados MiniFrance é um conjunto de dados muito mais diversificado do que muitos outros, e oferece desafios interessantes a serem superados.

Tabela 3.3: Os desempenhos em dados de pequena escala

Train set	OA	mIoU
100%	52.40	15.79
10%	50.14	15.25
Caen only (12.5%)	42.09	10.05

Como estudo de número quatro, temos a pesquisa realizada por Kumar et al., intitulada “A Comprehensive Survey on Ensemble Methods”. Este estudo explica que o método de Comitê de modelos são baseados no princípio de decisões coletivas, feitas por grupos de classificadores individuais, de forma a prever o output mais provável [73]. Ele usa a independência, a descentralização, a diversidade de opiniões e a agregação como critérios, baseados na teoria sobre “Wisdom of Crowd” [74].

O Comitê de modelos é uma combinação de vários classificadores que treinam com os mesmos dados de treinamento e, em seguida, os resultados são combinados com base em métodos como média ponderada, média simples, votação ou probabilidade. Podemos observar seu framework na Figura 3.1.

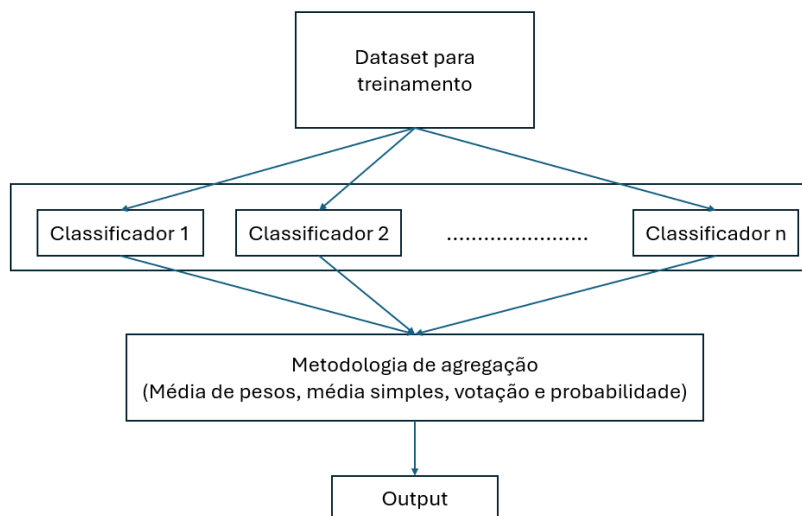


Figura 3.1: Framework referente ao método de Comitê de modelos segundo Kumar et al. [73].

Existem várias abordagens para se realizar a combinação dos classificadores. Entre elas, podemos destacar a manipulação dos dados, criando subconjuntos, que podem ser de dados randômicos ou não. Outra abordagem está na alteração dos algoritmos dos classificadores, onde cada classificador utilizado possui um algoritmo diferente. Uma terceira abordagem está na manipulação do output, onde são utilizados vários classificadores binários no lugar de um único multiclassificador.

Segundo Mishra [75], dentre as vantagens encontradas nesta implementação, destacamos:

- Redução do overfitting, que ocorre quando um modelo é muito complexo e se ajusta muito aos dados de treinamento, resultando em uma generalização ruim para novos dados. Neste caso, o aprendizado em conjunto pode ajudar a reduzir o excesso de ajuste combinando vários modelos que foram treinados em diferentes subconjuntos de dados ou com diferentes algoritmos;
- Melhora de acurácia onde, através da combinação de vários modelos, podemos tirar proveito de seus pontos fortes e fracos individuais;
- Tratamento de dados contendo ruídos, onde estes podem introduzir erros e vieses no modelo. Neste contexto, combinando vários modelos que foram treinados em diferentes subconjuntos de dados ou com diferentes algoritmos, é possível reduzir o impacto dos ruídos nos treinamentos;
- Tratamento de bancos de dados desbalanceados, que ocorre quando dados de uma classe são muito mais predominantes do que as demais. Neste caso, combinando vários modelos que foram treinados em diferentes subconjuntos de dados ou com diferentes algoritmos, o viés passível de acontecer pode ser amenizado; e
- Tratamento de bancos de dados grandes, cujo processamento pode ser extremamente caro computacionalmente. Neste caso, vários modelos podem ser treinados em paralelo, em diferentes subconjuntos de dados ou com diferentes algoritmos.

O artigo finaliza elucidando as limitações de emprego. Primeiramente, os autores falam sobre o tempo de computação ser longo devido à agregação de vários resultados e, em segundo, o fato da interpretação dos resultados ser complexa devido aos vários modelos.

Desai e Ghose, em “Active Learning for Improved Semi-Supervised Semantic Segmentation in Satellite Images”, propuseram o uso de Active Learning para selecionar um conjunto representativo de exemplos rotulados

para segmentação semântica, com aprendizagem semi-supervisionada, para classificação de bancos de dados focados em Land Cover & Land Use (LCLU) [89].

Para isso, eles utilizaram uma Conditional Generative Adversarial Networks (CGAN), que recebe um pequeno número de exemplos rotulados e um grande conjunto de dados não rotulados. Essa CGAN gera, então, “pseudo-rótulos” com base em exemplos rotulados, limitados para aumentar o conjunto rotulado.

Os autores demonstraram a eficácia do método em dois conjuntos de dados de segmentação semântica existentes contendo imagens de satélite: os bancos de dados UC Merced Land Use Classification [90] e DeepGlobe Land Cover Classification [43]. Eles utilizaram Active Learning para segmentação semântica [91] produzindo patches da imagem de entrada dada que são mais informativos. Em seguida, obtiveram anotações, a nível de pixel, somente para as amostras ativamente selecionadas, o que forneceu diversos dados de treinamento rotulados para segmentação semântica semi-supervisionada. Finalizando, eles propuseram o método de seleção de amostras como um processo de orientação para a criação de conjuntos de dados em grande escala, exigindo a coleta de anotações densas em nível de pixel.

Para a implementação, os autores utilizaram a ResNet 101 e a ResNet 50 como redes de classificação de imagens, que foram treinadas usando Cross-Entropy como função de perda. A rede foi treinada usando o otimizador Stochastic Gradient Descent (SGD), com uma taxa de aprendizagem básica de 0,001 e momentum de 0,9.

Para o treinamento, foi utilizado batch size de tamanho igual a 4, e treinaram por 50 épocas após cada consulta. Como hardware, os autores destacam que realizaram o treinamento em uma única GPU NVIDIA GTX-2080TI.

Como resultados, podemos observar na Tabela 3.4 uma comparação quantitativa do método proposto com a linha de base para o conjunto de dados de classificação de cobertura de terra do DeepGlobe [43]. Os autores registraram uma melhoria máxima de mIoU de cerca de 27% com apenas 2% de dados rotulados, uma melhoria máxima de cerca de 6% em relação à linha de base ao treinar com 5% de dados rotulados e uma melhoria de aproximadamente 8% com 12,5% de dados rotulados nas duas estratégias de aprendizagem ativa.

Proposto por Tian et al., “Dense Fusion Classmate Network for Land Cover Classification” adota uma rede chamada de Dense Fusion Classmate Network (DFCNet) na classificação da cobertura do solo [93].

Tabela 3.4: Resultados de mIoU para o conjunto de dados de classificação LCLU com o banco de dados DeepGlobe.

Labeled Ratio	2%	5%	12.5%
s4GAN [92] (Baseline)	0.403 ± 0.035	0.486 ± 0.018	0.511 ± 0.007
s4GAN + Entropy (Desai e Ghose)	0.469	0.513	0.554
s4GAN + Margin (Desai e Ghose)	0.511	0.513	0.529

Neste trabalho, a estratégia “Classmate”, como o aprendizado multitarefa, foi adotada. Ela combina dois conjuntos de dados de tarefas aparentemente não relacionadas e obtém melhorias na tarefa especificada. Neste caso, foram utilizados dados de estradas auxiliares e terrenos para treinamento. Apesar da segmentação do terreno não ser estruturada por não conter informações explícitas de limite, a estrada pode ajudar a compensar a falta de informações de cobertura do solo.

Além disso, um módulo de fusão densa também foi integrado, o que garante a discriminação precisa de pixels confusos e beneficia a otimização da rede a partir do zero. Este módulo resulta em vantagens do fluxo de gradiente, refinamento de recursos e fusão em várias escalas. Os dados de treinamento da extração de estradas utilizados foram os do Desafio de Satélite DeepGlobe CVPR de 2018, e têm um cenário semelhante ao da cobertura do solo. Estes dados foram considerados como dados auxiliares para o treinamento da rede, ajudando a gerar informações sobre estradas.

Para o treinamento, foi utilizado o Caffe [94] como estrutura de aprendizagem profunda. Todos os modelos usaram a FC-DenseNet como backbone. Como hardware foram utilizadas 8 GPUs, com o mini batch size de 32, e 80 mil iterações. Foi definida a taxa de aprendizado de 0,001, weight decay de 0,0005, momentum de 0,9 e a estratégia de otimização ploy.

Foram feitos uma série de experimentos com tamanhos de corte de 513 a 1025, e os autores obtiveram resultados diferentes no conjunto de dados de teste cortado por eles a partir do conjunto de dados de treinamento, incluindo 203 imagens. Isso mostrou que o tamanho de corte relativamente maior pode obter um desempenho melhor. Ao final, foi escolhido o tamanho de corte de 1025 x 1025.

Como resultados, o DFCNet obteve uma pontuação de 54,13 no conjunto de dados de teste. Para reduzir a aleatoriedade do processo de previsão e reduzir o ruído causado pelos próprios dados, foram feitas previsões em várias escalas e fundidas na previsão final, além de alguns modelos baseados no

DFCNet para prever o teste. Ao fazer isso, os autores obtiveram uma pontuação de 55,59.

Já no sétimo trabalho acadêmico, Taylor e Clayton, em sua pesquisa intitulada “Form from Function: Applying Flow Driven Experiential Learning to the Integration of Immersive Technology in Formal Military Aviation Training Programs”, realizaram um estudo, aplicando uma estrutura baseada na teoria da aprendizagem experimental e nos princípios da teoria do fluxo, para o Design da Experiência de Aprendizagem (DEA) [76].

O trabalho propõe o “Flow Driven Learning Experience Design” (FDLXD), uma estrutura fundamentada na teoria da aprendizagem experimental e orientada pelo conceito de alcançar e manter o fluxo, que também atende a vários critérios usados para desenvolver a aprendizagem em um ambiente de aprendizagem virtual ou imersivo.

Segundo os autores, a integração de ambientes de aprendizagem imersivos utilizados como método de entrega em programas de treinamento militar deve considerar 2 aspectos: o objetivo de aprendizagem desejado e o ponto em que ele ocorre no programa de treinamento geral. Neste sentido, o uso intencional de tecnologias imersivas tem como finalidade evitar a perda de eficácia do treinamento e maximizar os possíveis benefícios do aumento do desempenho do aluno e da eficiência do programa.

Eles afirmam ainda que os resultados podem ser alcançados pelo uso de Ambientes Imersivos. Segundo este conceito, a Realidade Aumentada pode ser utilizada com base em um ambiente real, aprimorando-o com objetos virtuais. Já o ambiente virtual é acessado por meio da Realidade Virtual (RV). Finalizando, a Realidade Mista (RM) descreve o intervalo entre os ambientes real e virtual no qual o usuário interage simultaneamente com ambos.

Durante o trabalho, são explorados alguns conceitos como a teoria do fluxo. A fig. 3.2 mostra o gráfico relacionado. As duas dimensões existentes, a saber, desafios e habilidades, variam ao longo dos eixos vertical e horizontal. Para os fins desta discussão, a letra A no diagrama representa qualquer aluno envolvido em um evento de treinamento técnico e o número representa um estado deste aluno, e observamos em vermelho a região do Canal de Fluxo.

Neste contexto, quando o aluno começa o evento de treinamento (A1), o ideal é que ele seja desafiado em um nível adequado à sua habilidade, marcado pela seta em azul na fig. 3.2, colocando-o em um estado de fluxo. Porém, à medida que o evento avança, os desafios apresentados podem aumentar em um ritmo abaixo do nível de habilidade crescente do aluno (deslocando a seta para A2) ou aumentar em um ritmo acima do nível de habilidade atual do aluno (deslocando a seta para A3). Em ambos os casos, a experiência será

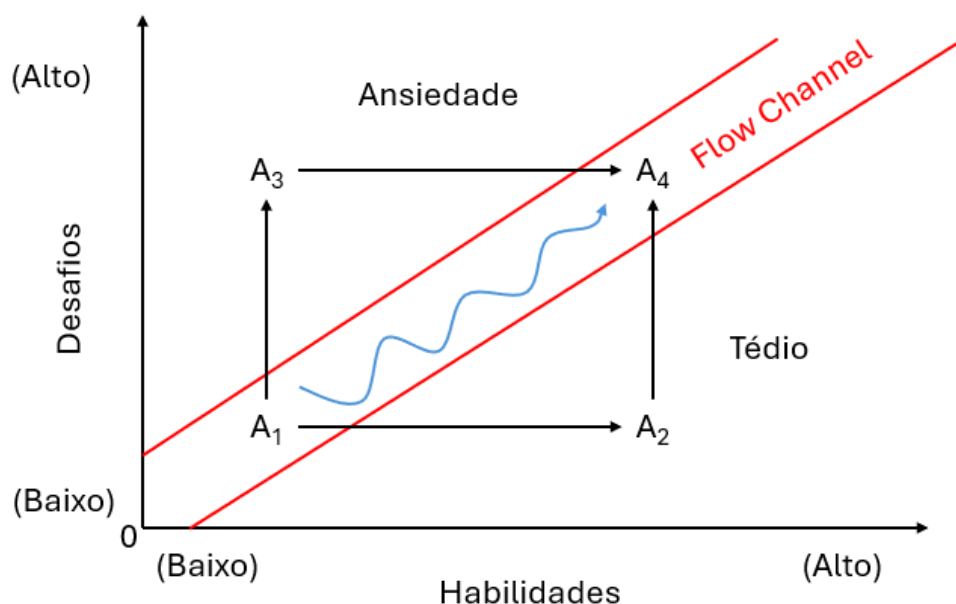


Figura 3.2: Gráfico relativo à teoria do fluxo.

prejudicada. Sendo assim, para retornar ao estado de fluxo, os desafios e habilidades devem ser ajustados.

Outros conceitos utilizados no FDLXD são a Teoria da Distância Transacional e os quatro critérios de Ehrlich (2002) para a aprendizagem virtual [77]:

1. Aluno para Interface, que consiste na capacidade de o aluno interagir com a tecnologia imersiva;
2. Aluno para Conteúdo, e a capacidade de o currículo ser aplicado no ambiente de aprendizagem imersiva;
3. Aluno para Aluno, e a capacidade de o aluno aprender com outros alunos; e
4. Aluno para Instrutor, e a capacidade de o aluno aprender com o instrutor

Para utilizar a estrutura, um evento de treinamento deve ser construído de forma que o princípio do fluxo de definição de metas claras com expectativas sustente a etapa de Experiência Concreta. O aluno deve, então, levar essa experiência para a segunda etapa de observação reflexiva e para a terceira etapa de conceitualização abstrata. Tanto na segunda quanto na terceira etapa, para estabelecer as condições para o fluxo, o aluno deve receber feedback dos instrutores e de outros alunos.

Depois de formular a sua experiência através de conceitos e teoria, o aluno passa para a experimentação ativa. Nesta etapa, o aluno é levado a um

evento que assimila os desafios percebidos da tarefa e a capacidade percebida pelo próprio aluno de resolvê-la.

O estudo concluiu que a ergonomia cognitiva da Realidade Aumentada e da Realidade Virtual em uma tarefa complexa de tomada de decisão baseada em atividades reais envolvendo Operações Navais foi importante como ferramenta de treinamento para os militares. Além disso, a realidade virtual, a realidade aumentada e o currículo baseado em conteúdo de vídeo foram abordagens de aprendizagem comprovadamente eficazes para a aprendizagem imersiva [76].

O oitavo estudo, escrito por Michael Kozhevnikov e Maria Kozhevnikov, intitulado “The Effect of Environment Immersivity on Spatial Perspective-Taking Task Performance”, examinaram o efeito de diferentes ambientes de visualização no desempenho de uma tarefa de tomada de perspectiva espacial. Esta tarefa consiste na capacidade de se orientar mentalmente para um objeto externo e indicar sua localização em relação a si mesmo [78].

Para esse experimento, 24 estudantes de graduação em psicologia (10 homens e 14 mulheres) participaram do estudo. Nele, todos os participantes tinham visão normal ou corrigida para normal. Os participantes receberam uma tarefa de tomada de perspectiva apresentada em quatro visualizações diferentes: área de trabalho do computador (2D), tela grande com o uso de óculos estereoscópicos combinados com a projeção de estímulos em uma parede (3DP), tela grande com o uso de óculos estereoscópicos combinados com a projeção de estímulos em uma parede e atualização em tempo real (3DPT) e ambiente virtual totalmente imersivo usando um Head-Mounted Display (3DVR).

Cada participante concluiu uma tarefa de tomada de perspectiva espacial em quatro ambientes de visualizações (AV) diferentes (em ordem crescente de imersividade do ambiente): 2D, 3DP, 3DPT e 3DVR.

A condição 2D foi usada como linha de base para mostrar o desempenho dos participantes na tarefa de PTA convencional e, posteriormente, foi usada como referência para comparação de qualquer possível desvio dos padrões de resposta da linha de base em 3DP, 3DPT e 3DVR.

Tanto o 3DP quanto o 3DPT têm um tamanho equivalente da tela do projetor, que é maior do que a do 2D. O sistema de rastreamento da cabeça foi adicionado na condição 3DPT para simular a atualização espacial, ajustando a cena da tela de exibição percebida de acordo com o movimento da cabeça dos indivíduos.

Como resultados obtidos, as quatro visualizações apresentaram um padrão semelhante de respostas às tarefas de tomada de perspectiva no mundo

real. Os autores sugeriram a dependência do sistema egocêntrico durante a codificação e a manipulação de estímulos nas tarefas em todos os quatro ambientes.

Os autores sugeriram ainda que isso implica que o teste poderia ser usado como uma ferramenta de diagnóstico definitiva para traçar o perfil da aptidão profissional para as profissões que exigem altas habilidades de navegação, como astronáutica, pilotagem aérea, UAV e outras teleoperações, robótica, cirurgia etc.

Além disso, segundo os autores, os resultados dos experimentos revelaram que o próprio ambiente imersivo é um fator crucial para medir com precisão a codificação e o processamento egocêntricos. Uma análise comparativa das respostas dos sujeitos em todas as quatro visualizações sugere que o 3DVR exige o maior envolvimento da codificação espacial egocêntrica.

Em particular, embora não tenha havido diferença nos erros adjacentes entre os diferentes ambientes, o número de erros de reflexão para apontar para trás foi significativamente mais proeminente do que nos outros três ambientes, indicando que o 3DVR, de fato, dependia mais de um quadro de referência centrado no corpo.

Os resultados deste estudo sugerem ainda que, embora o 3DP e o 3DPT sejam melhores do que a tela de computador 2D padrão, apenas ambientes totalmente imersivos (como o 3DVR) parecem garantir que os indivíduos codifiquem imagens em relação ao quadro de referência espacial egocêntrico (baseado no visualizador).

Os autores explicam suas descobertas pelo “efeito de moldura” existente em ambientes 2D, 3DP e 3DPT. Neste aspecto, de acordo com os autores, quando os estímulos são apresentados em uma tela de computador 2D ou até mesmo em uma tela de projeção grande, mas as bordas da tela são visíveis (formando uma “moldura”), os objetos-alvo são incorporados em uma “moldura fixa”.

Este efeito sugere que sempre que uma pessoa está observando uma cena de “fora” (por exemplo, uma cena é definida pela moldura ou pela tela do computador), ela depende mais do uso da codificação alocêntrica (baseada na cena).

Um nono trabalho, escrito por Rahimi et al., intitulado “Generative AI Meets Virtual Reality: A Comprehensive Survey on Applications, Challenges, and Future Direction”, apresenta uma pesquisa sobre IA Generativa abordando aplicações, percepções e desafios relativos ao seu emprego em Realidade Virtual (RV) [95].

Para isso, os autores exploraram como a IA generativa aprimora os recur-

sos de RV em domínios como educação, saúde, jogos, simulação e treinamentos. O artigo analisa componentes como interação multimodal, adaptação dinâmica de conteúdo, storytelling, treinamentos, simulação e IA com consciência emocional.

Sua abordagem envolve como esses componentes redefinem coletivamente as experiências do usuário, possibilitando ambientes adaptáveis, personalizados e orientados pelo contexto. Ele cita que a IA generativa possibilita suporte a aplicativos de RV colaborativos, permitindo ambientes multiusuários em que os participantes podem criar e se envolver em tempo real.

Outro aspecto está na adaptabilidade ao conteúdo dinamicamente, com base no feedback do grupo, promovendo o trabalho em equipe e a criatividade. Como exemplo, temos as aplicações ambientais, capazes de aproveitar os modelos generativos para simular os efeitos da mudança climática, as interações do ecossistema e os cenários de planejamento urbano. Tais alterações fornecem aos tomadores de decisão percepções valiosas para estratégias de sustentabilidade.

Outros empregos envolvem o turismo virtual e a preservação do patrimônio. Neste sentido, a IA poderia reconstruir locais históricos, simular civilizações perdidas ou criar experiências de viagem virtuais personalizadas, permitindo que os usuários explorem marcos culturais de uma forma que transcenda as limitações físicas e financeiras.

Já em Internet of Things (IoT) industrial e na infraestrutura inteligente, a IA generativa poderia permitir a visualização em tempo real de fluxos de trabalho de fábricas, automação de depósitos e operações da cadeia de suprimentos.

Porém, à medida que os usuários navegam em ambientes virtuais imersivos, eles precisam processar grandes quantidades de informações sensoriais e cognitivas, o que muitas vezes leva à sobrecarga cognitiva. Outro aspecto, o design ético da IA generativa em RV, concentra-se em garantir a inclusão, a justiça e o respeito aos direitos do usuário. Ainda neste sentido, dados confidenciais apresentam riscos à privacidade se forem protegidos de forma inadequada ou usados sem consentimento.

Esta preocupação existe especialmente em sistemas que geram automaticamente conteúdo, avatares ou narrativas. Um exemplo está nos vieses em conjuntos de dados de treinamento de IA, que podem resultar em comportamentos excludentes ou culturalmente insensíveis.

Além disso, muitos modelos de IA não têm a capacidade de interpretar o espaço tridimensional com precisão, levando a posicionamentos ruins de objetos, comportamentos inconsistentes de NPCs e adaptabilidade limitada às ações do usuário. Essas deficiências prejudicam o realismo e a aplicabilidade

prática dos sistemas de RV, especialmente em domínios como projeto arquitetônico, planejamento urbano e simulações de resposta a emergências.

A integração de diversas modalidades como entrada, fala, gestos, olhar ou movimentos corporais é fundamental para as interações imersivas em RV. No entanto, essa integração é complexa devido à variabilidade no comportamento do usuário, condições ambientais, etc. Os autores destacam também as limitações de hardware, como potência de processamento, duração da bateria e gerenciamento térmico dos dispositivos de RV autônomos.

Os autores destacam ainda que, para enfrentar esses desafios, serão necessários avanços na eficiência do hardware, na arquitetura dimensionável de borda de nuvem e nas iniciativas de educação direcionadas do setor para promover a prontidão e derrubar as barreiras à adoção comercial.

Coletivamente, os desafios descritos acima - latência abrangente, restrições de hardware, complexidade multimodal, escalabilidade, riscos éticos, confiança do usuário e barreiras comerciais - representam limitações fundamentais que atualmente impedem o desenvolvimento e a implantação generalizados de sistemas de RV generativa.

Os autores destacam ainda que os problemas citados não apenas restringem o desempenho técnico, mas também restringem a adoção pelo usuário, o investimento do setor e a integração social. Sem soluções direcionadas para melhorar a eficiência computacional, a interpretabilidade, a escalabilidade do sistema e as proteções éticas, a RV generativa permanecerá limitada a aplicações experimentais ou de nicho, em vez de atingir todo o seu potencial transformador.

Finalizando, o décimo trabalho, escrito por Bolkas et al. intitulado “Creating a virtual reality environment with a fusion of sUAS and TLS point-clouds”, propõe um pipeline para a criação de ambientes virtuais a partir de nuvens de pontos 3D [96].

A motivação e o objetivo geral deste estudo são criar ambientes virtuais que serão usados em laboratórios de levantamento topográfico virtual para avaliar a função da realidade virtual imersiva e interativa no ensino de engenharia topográfica.

A ênfase está na fusão de nuvens de pontos derivadas de escaneamento a laser, Sistemas Aéreos não Tripulados (SANT) e texturização. Esta fusão pode levar a nuvens de pontos mais completas que podem ser usadas para aprimorar a criação de ambientes de realidade virtual. Os objetos artificiais como prédios e escadas são modelados manualmente usando primitivos padrão (por exemplo, caixas, esferas, cilindros) e ajustando esses primitivos na nuvem de pontos fundidas.

Como resultados, o algoritmo de fusão, o terreno e a modelagem de edifícios permitiram aos autores manter a fidelidade geométrica do mundo real na realidade virtual. As texturas personalizadas se assemelham com fidelidade suficiente ao ambiente físico, criando assim a ilusão de que o usuário está no ambiente do campus.

No artigo, os autores destacaram também os desafios encontrados. Em termos de criação da nuvem de pontos e do modelo de terreno mesclados, as tecnologias de levantamento topográfico forneceram informações geométricas em 3D que não foram facilmente transferíveis para mecanismos de desenvolvimento de jogos.

Outro aspecto estava na modelagem de objetos 3D e o fato de ser trabalhosa quando feita manualmente. Já o problema da reconstrução urbana e de superfície é um tópico ativo de pesquisa com vários desafios em termos de automação total, limitações de algoritmo, perda de qualidade e limitações de dados.

Com relação à texturização, os autores identificaram que as informações de cores visíveis que acompanham as nuvens de pontos 3D geralmente são insuficientes para a texturização de ambientes virtuais. Além disso, para replicar um ambiente físico, muitas vezes foi necessário criar texturas personalizadas, consumindo muito tempo. Além disso, as texturas criadas muitas vezes parecem “falsas” ao serem aproximadas. Sendo assim, texturas mais detalhadas podem criar uma sensação de naturalismo, mas a renderização de texturas mais complexas pode ser computacionalmente caras e difíceis de criar.

Ao longo deste capítulo, obtivemos informações importantes que servem como referência. Nele detalhamos 6 (seis) trabalhos relacionados ao uso de segmentação semântica, sendo 5 envolvendo sensoriamento remoto. Além disso, foram relacionados 3 trabalhos que versam sobre o emprego de Simulação no processo de aprendizagem e interação, e 1 trabalho que versa sobre a importância da criação de terrenos virtuais para estudos. Estas informações foram resumidas na Tabela 3.5.

Cabe destacar aqui que apenas 2 dos bancos de dados dos trabalhos relacionados são iguais ao proposto por nossa abordagem, e serão abordados posteriormente no Capítulo 4.

Em nossa abordagem, desenvolvemos um simulador capaz de utilizar a região virtual criada. Ele permite o estudo da região utilizando diversas interfaces e recursos, todas interagindo entre si. Dessa forma, seria possível aumentar o conhecimento adquirido por alunos e planejadores antes mesmo do término da confecção de seus planos ou da ida à região real.

Os capítulos seguintes versarão sobre a criação e o treinamento do modelo

Tabela 3.5: Resumo dos trabalhos relacionados neste capítulo

Trabalho Publicação	Descrição	Metodologia	Banco de dados Utilizado	Resultados Relevância
Zheng e Chen (2023)	Segmentação de imagens de alta resolução; modelos binário e multiclassifica- ção	Arquitetura U- Net; comparação entre modelos de classificação bi- nária e multi- classificação	Gaofeng-2	Resultados com- parativos, desta- cando eficiência
Castillo- Navarro et al. (2019)	Avaliação da quantidade e qualidade dos dados para segmentação supervisionada	Redes neurais profundas; Seg- Net com VGG16 pré-treinado; Stochastic Gra- dient Descent	MiniFrance, Semantic Labeling Vaihingen	Importância da diversidade nos dados e transfe- rência de apren- dizado
Propostas Open-Set Recognition (2024)	Cenários com classes desco- nhecidas nos dados de teste	Open Gaus- sian Mixture of Models (OpenGMM), Conditional Reconstruction (CoReSeg), superpixels (FuSC)	ISPRS Vaihingen e Potsdam	Melhora na con- sistência semân- tica e reconheci- mento de classes desconhecidas
Análise Com- parativa LiDAR	Compressão dinâmica de nu- vens de pontos para veículos autônomos	Autoencoder convolucional; segmentação semântica para compressão	Dados sintéticos e reais LiDAR	Alta taxa de compressão mantendo infor- mações críticas
Experimentos MiniFrance (Castillo- Navarro et al.)	Aprendizado su- pervisionado em grande escala	Treinamento com frações do banco de dados; avaliação de adaptação ao domínio	MiniFrance	Treinamento com poucos dados preserva desempenho; variação entre cidades

de IA para a segmentação semântica, o desenvolvimento de um visualizador para o terreno virtual criado fruto deste treinamento, e a criação da CT e do terreno para o SJD, todos em capítulos separados. As propostas detalhadas serão então concatenadas na conclusão.

4

Segmentação Semântica em imagens de Sensoriamento Remoto

Este capítulo busca detalhar a primeira parte da nossa abordagem, a saber, a segmentação semântica de maneira geral. Para isso, versaremos sobre as metodologias empregadas para a realização do treinamento com o modelo de IA.

O nosso estado final desejado será composto por imagens RGB em formato “.png”. Estas imagens serão utilizadas como base para a criação do terreno virtual e criação da carta de trafegabilidade. Além disso, as imagens obtidas também serão utilizadas como dados de entrada para a confecção do terreno do SJD.

Nossa abordagem busca comprovar a seguinte hipótese básica estabelecida no Capítulo 1: O método de Comitê de modelos (do inglês, “Ensemble methodology”), para segmentação semântica binária, apresenta desempenho superior à abordagem multiclasse tradicional, permitindo identificar com maior precisão as classes relacionadas à trafegabilidade do terreno em imagens de sensoriamento remoto.

Cabe destacar aqui que as hipóteses básicas serão tratadas no decorrer das abordagens dos demais capítulos. No capítulo 7, todas as hipóteses básicas serão revisitadas de forma a concatenar os resultados obtidos.

De forma a atingirmos o nosso estado final desejado, uma série de experimentos e testes empíricos foram realizados. Destes, 4 experimentos são citados neste trabalho por possuírem características marcantes, tais como o uso de datasets com combinações de índices normalizados. Estes testes são descritos nas seções seguintes.

Para tal, versaremos primeiro sobre as dificuldades na obtenção dos datasets, com a finalidade de delimitarmos nosso espaço de busca pelos dados de entrada a serem utilizados. Em seguida, descreveremos os experimentos realizados, detalhando os datasets e as arquiteturas utilizadas, exibindo seus resultados e análises.

4.1

Sobre a busca pelo dataset

Conforme já descrito na seção 2.5, satélites que obtêm imagens SAR utilizam sua própria fonte de iluminação por meio de bandas do domínio de micro-ondas, sendo capazes de obter medições a qualquer momento. Além

disso, a resolução espacial das imagens SAR depende do sensor usado e de seu modo de aquisição.

O uso destas imagens SAR é importante pois possibilita obter imagens de boa qualidade e em tempo reduzido se comparado com outras fontes abertas. A depender da constelação de satélites utilizada, o tempo de recebimento das imagens pode ser de dias ou até semanas, sendo assim um processo mais rápido que o manual. Porém, uma dificuldade está no desenvolvimento de uma maneira para extrair as características dos diferentes tipos de áreas existentes.

Neste contexto, utilizaremos técnicas de segmentação semântica, descritas na subseção 2.6.3. Nossa abordagem envolverá, para o primeiro experimento, a identificação das áreas existentes em uma imagem SAR e, nos demais experimentos, a utilização de imagens RGB óptico, e a sua separação em classes. Essas classes serão então utilizadas como inputs para os próximos capítulos deste trabalho.

Como abordagem, ficou estabelecido como parâmetro que as imagens a serem utilizadas deverão possuir anotações correspondentes às classes usualmente utilizadas nas cartas de trafegabilidade e nos simuladores. Além disso, foram consideradas na pesquisa imagens com resolução acima de 1024×1024 , devido à possibilidade de cortes em patches menores e a consequente possibilidade de treinamento englobando áreas grandes, e classes que se aproximassem das necessárias à demanda deste trabalho.

Neste sentido, foi realizada busca por informações junto ao setor operativo do Corpo de Fuzileiros Navais em busca de CT confeccionadas pelas equipes de planejamento. Esta aproximação visava trabalhar em conformidade com o padrão exigido por equipes inseridas neste contexto e obter informações mais próximas das necessárias. Porém, dado o caráter sigiloso dos documentos, não obtivemos sucesso.

Sendo assim, os primeiros experimentos buscaram imagens que se aproximassem dos terrenos já utilizados pelo SJD. Já nas tentativas seguintes, levamos em consideração dados que utilizassem imagens de sensoriamento remoto, contendo uma resolução espacial mais próxima possível de $0.3m^2$ [81].

Durante o andamento da pesquisa por datasets sem grau de sigilo, algumas dificuldades foram encontradas. Entre elas, destacamos os poucos datasets existentes que atendessem às demandas por anotações contendo as classes necessárias. Outros casos encontrados demandavam custos, sendo muito caro realizar sua aquisição considerando a quantidade de imagens e o tamanho destas.

Além disso, boa parte dos bancos de dados encontrados eram mal anotados ou possuíam apenas coordenadas dos limites das classes, sendo necessária

montagem utilizando bibliotecas específicas em Python. Desta forma, não seria possível detalhar, com precisão, as regiões de estradas, vegetações, entre outras áreas anotadas [82]. Outros problemas envolviam, também, a ausência de anotação de certas classes, ainda que existentes na imagem. Essas distorções prejudicariam gravemente o treinamento da Rede Neural, levando a resultados aquém do esperado.

Outro aspecto importante observado foi a escassez, ou até inexistência, de um dataset contendo as diferentes regiões do Brasil, com terrenos anotados que atendessem à nossa demanda. Porém, dado o tempo necessário para a criação e anotação das imagens, a criação destas imagens não foi possível.

De forma a conduzirmos os segundo, terceiro e quarto experimentos, utilizamos o banco de dados DeepGlobe 2018 Satellite Image Understanding Challenge [43] com algumas variações. Este dataset possui 3 diferentes tipos de divisões, sendo uma divisão voltada para diferentes tipos de regiões, chamada de Land Cover Classification Challenge, uma segunda divisão voltada para identificação de estradas e rodovias, chamada de Road Extraction Challenge, e uma terceira voltada para identificação de prédios, chamada de Building Detection Challenge.

O Land Cover Classification Challenge contém 1.146 imagens de satélite de tamanho 2.448×2.448 pixels no total, divididas em conjuntos de treinamento, validação e teste, contendo 803, 171 e 172 imagens, respectivamente. O tamanho total da área do conjunto de dados é equivalente a $1.716,9 \text{ km}^2$.

Já o Road Extraction Challenge contém 8570 imagens de satélite de tamanho 2.448×2.448 pixels no total, divididas em conjuntos de treinamento, validação e teste, contendo 6.226, 1.243 e 1.101 imagens, respectivamente. O tamanho total da área do conjunto de dados é equivalente a 2.220 km^2 .

Finalizando, o Building Detection Challenge contém 8570 imagens de satélite de tamanho 2.448×2.448 pixels no total, divididas em conjuntos de treinamento, validação e teste, contendo 6.226, 1.243 e 1.101 imagens, respectivamente. O tamanho total da área do conjunto de dados é equivalente a $4.917,2 \text{ km}^2$.

Todas as imagens contêm dados RGB, com uma resolução espacial de 50 cm, coletados do conjunto de dados DigitalGlobe Vivid+. As anotações são compostas por imagens RGB com 7 classes, onde:

- Urban land: Áreas construídas e feitas pelo homem com artefatos humanos;
- Agriculture land: Qualquer plantação regular, terras de cultivo, pomares, vinhedos, viveiros e áreas de horticultura ornamental;

- Rangeland: Qualquer área não florestal, não agrícola, terra verde, grama.
- Forest land: Qualquer terra com pelo menos 20% de densidade de copa de árvore mais cortes rasos;
- Water: Rios, oceanos, lagos, zonas úmidas, lagoas;
- Barren land: Montanha, rocha, sobremesa, praia, terra sem vegetação; e
- Unknown: Nuvens e outros.

Considerando as dificuldades encontradas acima, buscamos datasets que atendessem à solução do problema proposto da melhor maneira possível. Em alguns casos, tentamos criar datasets que servissem como solução, porém com resultados aquém do esperado. As seções 4.2 a 4.5 fornecerão maiores detalhes sobre o assunto.

4.2

Experimento 01 - Usando Feature Pyramid Network em imagens satelitais Sentinel 2

Para este primeiro experimento, foi utilizado como IDE o Google Colaboratory [83]. Seu hardware disponibilizado foi o processador Intel Xeon, contendo 16GB de RAM na CPU. A GPU utilizada foi a TESLA k80, com 12GB de RAM. A linguagem de programação foi utilizada Python, com bibliotecas PyTorch inicialmente e, posteriormente, Tensorflow, Pillow, Numpy, Matplotlib, Segmentation Models entre outras. A troca do PyTorch por Tensorflow se deve à maior familiaridade, seu suporte a diferentes tipos de hardware (CPU, GPU e TPU) e seu ecossistema [84]. Já o uso das demais bibliotecas se deve à grande comunidade existente, a suas capacidades no trato das imagens e realização de treinamentos com modelos de IA de maneira geral.

Como dataset para este experimento, foram utilizados 6 terrenos virtuais criados anteriormente para o SJD pelos militares do CSimCFN. A intenção foi aproveitar as CT já criadas, fruto do desenvolvimento dos terrenos para jogos de guerra, utilizando-as como máscaras para os treinamentos.

Estas máscaras possuem, como base, imagens de satélites obtidas da constelação Sentinel-2, com resolução espacial de $10m^2$. Além disso, as imagens possuem diferentes resoluções, sendo a maior resolução de tamanho 7800 x 5100 pixels.

A imagem a ser utilizada como entrada para o modelo de IA foi criada por este autor para este experimento, concatenando índices por diferença normalizados, com a intenção de aproveitar o máximo de recursos disponibilizados pela constelação Sentinel-2. Estes índices são obtidos através de cálculos pixel a pixel realizados entre as bandas do espectro eletromagnético obtidas dos satélites.

Para este experimento, foram combinados o Índice de Vegetação por Diferença Normalizada (NDVI), o Índice de Água por Diferença Normalizada (NDWI) e o MDE, todos obtidos em [116]. Esta imagem foi criada buscando utilizar fatores que normalmente influenciam na criação da trafegabilidade, a saber, vegetação, hidrografia e relevo.

Os cálculos do NDVI e do NDWI são mostrados através das equações abaixo. Elas consideram as bandas do espectro eletromagnético visíveis a olho nu, correspondentes às cores vermelho (Red) e verde (Green), além da banda invisível de Infravermelho Próximo (NIR). Cada equação envolve o uso da reconstrução espectral usando a imagem tirada com os filtros Red e Green, respectivamente, a partir das leituras do sensor.

$$NDVI = (NIR - Red)/(NIR + Red) + 1 \quad (4-1)$$

$$NDWI = (Green - NIR)/(Green + NIR) \quad (4-2)$$

A equação (4-1) aborda a obtenção do NDVI, que usa a diferença entre o NIR e o Red, dividida pela soma do NIR e do Red. Já a equação (4-2) mostra o cálculo do NDWI, que usa a diferença entre o Green e o NIR, dividida pela soma do NIR e do Green.

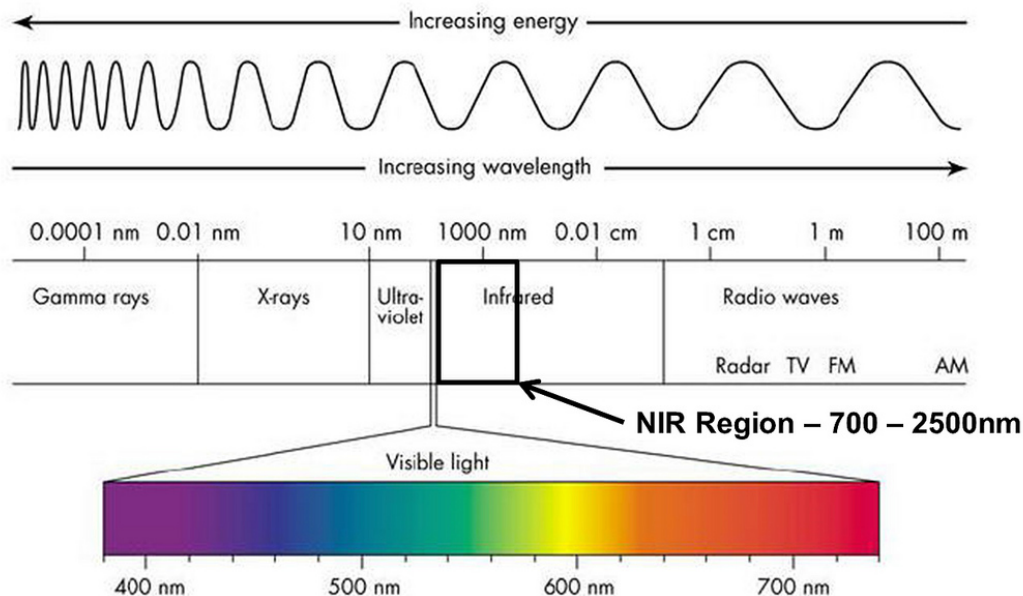


Figura 4.1: Faixas do espectro eletromagnético exibindo os respectivos cumprimentos de ondas. A região identificada como “visible light” identifica a faixa contendo as cores visíveis. Já a região identificada como “NIR region”, descreve a região contendo a faixa do NIR. Fonte: <https://bbpsales.com/near-infrared-measurements-how-do-they-work/>.

Observamos na imagem 4.1 as faixas do espectro eletromagnético. Utilizamos a cor vermelha para o cálculo do NDVI pois, em relação ao espectro eletromagnético, a clorofila absorve os comprimentos de onda vermelho e azul, mas reflete o verde e o infravermelho [86]. Neste sentido, a quantidade de clorofila nas plantas está diretamente relacionada à sua capacidade de produzir alimentos para si mesmas, o que também se manifesta na saúde das culturas. Já referente ao NDWI, segundo Pereira et al. [87], a água reflete mais luz na região do verde do espectro eletromagnético,

Sendo assim, o NDVI, o NDWI e o MDE foram concatenados em 3 canais de uma imagem, nesta ordem. Posteriormente, as imagens foram cortadas em patches de tamanho (128, 128, 3), de forma a se obter maior detalhamento das características das regiões.

Ao todo, 1698 imagens foram obtidas, sendo divididas entre 1398 imagens para treinamento e 300 imagens para teste. O uso de patches para os treinamentos se deve à necessidade de se obter mais detalhes dentro da imagem e evitar mudar o seu tamanho por ocasião de seu uso na entrada do modelo, sob risco de perda de detalhamento.

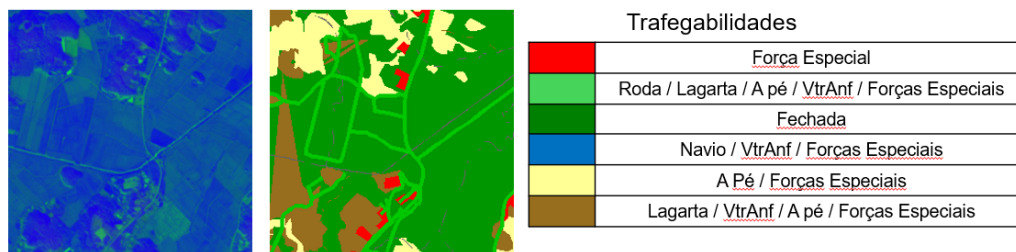


Figura 4.2: À esquerda, imagem concatenada para ser usada como dado de entrada no modelo de IA. À direita, a anotação referente à esta imagem.

Observamos na Figura 4.2 o exemplo de imagem de entrada na RNA e sua respectiva máscara. A imagem utilizada como máscara é composta pelas anotações das 7 classes existentes, a saber: navio, lagarta, pé, forças especiais, anfíbio, roda e unknown. Observamos as classes utilizadas na Figura 2.6.

Para o treinamento, a abordagem foi baseada na arquitetura Feature Pyramid Network, observando em [88]. Neste modelo, existe a necessidade de tiles com medidas em múltiplos de 32. O backbone utilizado foi o modelo ResNet50, a função de ativação foi a softmax, a função de perda usada foi a Categorical Cross Entropy e o otimizador foi o Adam.

Com as configurações citadas acima, foram realizados treinamentos com combinações de 3 classes. Porém, dada a pouca quantidade de imagens obtidas

e a predominância de imagens serem de regiões de mar, obtivemos como resultado máximo o mIoU 0.2232.

Outros fatos observados durante a análise das imagens foram:

- A imagem da trafegabilidade não se ajustou perfeitamente à imagem satélite obtida;
- Houve dificuldade na anotação da máscara, de forma a estar padronizada com a trafegabilidade esperada pelos meios. Como exemplo, citamos o fato do CLAnf poder navegar no mar e no rio, em que pese as dificuldades existentes no rio como a corrente, tipo de solo e profundidade, que podem influenciar de forma negativa no seu deslocamento se comparado com o mar; e
- Dificuldades em se obter imagens correspondentes à região com qualidade desejada. Muitas das imagens existentes possuem coberturas de nuvens ou, até mesmo, pixels sem dados; e
- A confecção da imagem de trafegabilidade não foi feita ajustando exatamente à carta militar nem às imagens satélites do terreno.

Dadas as dificuldades citadas anteriormente, podemos acrescentar ainda as dificuldades em obter mais imagens para treinamento e teste. Além disso, considerando também o pouco detalhamento das regiões fruto de resoluções espaciais altas quando comparadas a outros tipos de satélites, a próxima abordagem buscou melhores resoluções de imagens e novas arquiteturas.

4.3

Experimento 02 - Usando UNet em 3 classes do banco de dados Deepglobe

O segundo experimento utilizou o banco de dados DeepGlobe Land Cover Classification challenge [43] com algumas modificações. Nesta abordagem, de forma a obter informações mais generalizadas, buscou-se a separação entre água, vegetação e background apenas. Para isso, Agriculture land, Rangeland e Forest land foram classificadas como vegetação de maneira geral. Da mesma forma, a classe água envolveu todas as conformações relativas à classe, já detalhadas acima.

Foram realizados testes com a resolução de 299x299 por ser a resolução de entrada na rede, contendo 3 canais para as imagens de satélites e máscaras. Para isso, as imagens foram separadas em tiles com a resolução citada anteriormente, com interseção de 30% entre elas.

Ao todo, 8.518 imagens foram obtidas, sendo divididas entre 5.452 imagens para treinamento, 1.363 imagens para validação e 1.703 imagens de

teste. Nesta separação de dados, obedecemos às porcentagens de 80% para treinamento e 20% para teste, sendo retirada do treinamento 20% das imagens para validação. Além disso, foi empregado o processo de one-hot Encode.

A IDE empregada foi o Google Colaboratory [83], com a linguagem de programação em python. Neste contexto, foi disponibilizado para hardware o processador Intel(R) Xeon(R) CPU @ 2.20GHz, com cerca de 86 GB RAM e 12 núcleos. Como GPU, foi fornecida a NVIDIA Tesla T4, contendo 16GB de RAM.

Para os treinamentos, compilamos uma arquitetura UNet com backbone em inceptionresnetv2 [99]. Como função de perda, foi utilizada a categorical cross-entropy e, como otimizador, selecionamos o Adam optimizer. As imagens foram separadas com batch size de tamanho 8 e, ao todo, foram realizadas 100 épocas.

A Inception-ResNet-v2 é uma rede neural convolucional treinada no banco de dados ImageNet. A rede tem 164 camadas de profundidade e pode classificar imagens em 1.000 categorias de objetos, como teclado, mouse, lápis e muitos animais. A rede tem um tamanho de entrada de imagem de (299, 299, 3) [97].

Já a função de perda categorical cross-entropy é usada para a classificação multiclasse. Ela serve para quantificar a dissimilaridade entre as probabilidades previstas e os rótulos categóricos verdadeiros [98].

Para isso, devemos realizar uma conversão dos pixels das imagens para one-hot encode. Ela é composta por uma ativação Softmax mais uma perda de cross-entropy. Se usarmos essa função de perda, treinaremos uma CNN para gerar uma probabilidade sobre a quantidade de classes necessárias para cada imagem. Observamos sua equação abaixo:

$$L = - \sum_{i=1}^C y_i \cdot \log(\hat{y}_i) \quad (4-3)$$

Nela temos o número de classes como C , o rótulo verdadeiro y_i na forma one-hot (ou probabilidade real da classe i) e $\hat{y}_i$ como a probabilidade prevista pelo modelo para a classe i .

O otimizador Adaptive Moment Estimation (Adam) foi desenvolvido para otimização baseada em gradiente de funções objetivas estocásticas, com base em estimativas adaptativas. Ele foi projetado para melhorar as velocidades de treinamento em redes neurais profundas e alcançar a convergência mais rapidamente [100].

Para isso, ele personaliza a taxa de aprendizagem de cada parâmetro com base em seu histórico de gradiente, e esse ajuste ajuda o modelo a aprender de forma eficiente como um todo.

Realizando o treinamento com estes parâmetros, obtivemos resultados aquém do esperado, com valor máximo de mIoU de 0.3448 na parte da validação. Sendo assim, não foram realizados testes com predições.

4.4

Experimento 03 - Usando UNet em 6 classes do banco de dados Deepglobe

Para este treinamento, foram utilizadas 6 classes do banco de dados DeepGlobe Land Cover Classification challenge [43]. Como mencionado na seção 4.1, este banco contém 803 imagens de satélite com suas respectivas máscaras, de tamanho 2.448×2.448 pixels cada.

O propósito da utilização de 6 das 7 classes disponibilizadas está em detectar efetivamente as trafegabilidades necessárias para sua utilização na criação do terreno virtual e da CT. As classes “rangeland”, “agriculture land” foram agrupadas na classe “soft vegetation” por se tratarem de vegetações com características muito parecidas.

Para tal, foi compilada uma arquitetura UNet com modelo ResNet50. Este modelo foi escolhido pois obteve o melhor desempenho durante os treinamentos com a configuração destacada neste experimento. A configuração desta arquitetura possui como input o shape (384,384, canais) e pesos pré-treinados com o dataset Imagenet.

Já a função de ativação utilizada foi a Softmax e o otimizador escolhido foi o Adam. Como função de perda, foi escolhida a categorical focal jaccard loss. Um batch size de valor 8 e o total de 200 épocas foram determinados para este treinamento. Foi inserido também um early stopping de 30 épocas.

A função de perda categorical focal jaccard é calculada pela soma da perda focal categórica e do índice Jaccard [101]. Esta função é uma métrica usada na avaliação de modelos de segmentação semântica que enfatiza a importância dos pixels facilmente classificáveis e prioriza os pixels difíceis, mais frequentemente rotulados incorretamente pelo modelo. A função de perda categorical focal jaccard pode ser expressa pela seguinte equação:

$$\mathcal{L}_{focal} = - \sum_{c=1}^C GT_c \cdot \alpha (1 - P_c)^\gamma \cdot \log(P_c) \quad (4-4)$$

Nesta equação GT_c representa o rótulo verdadeiro da classe c (one-hot encoded), P_c é a probabilidade prevista para a classe c (saída do softmax), α é um fator de balanceamento entre classes γ é o parâmetro de foco, que controla a ênfase em amostras difíceis.

Como métrica, usamos IoU, descrita na equação ???. Esta é a métrica mais comumente usada para comparar a semelhança entre duas formas arbitrárias.

O IoU codifica as propriedades de forma dos objetos em comparação com suas medidas e locais delimitadores e, em seguida, calcula uma medida normalizada que se concentra em suas áreas (ou volumes). Essa propriedade torna o IoU invariável à escala do problema em consideração. Devido a essa propriedade atraente, todas as medidas de desempenho usadas para avaliar a segmentação, a detecção de objetos e o rastreamento utilizam muito esta métrica [102].

Foram realizados treinamentos com as imagens do dataset, obtendo mIoU na validação de 0.4206, além dos resultados obtidos na predição, exibidos na Tabela 4.1, na Seção 4.7. Buscamos, então, outra abordagem para o nosso problema proposto. A seção a seguir descreverá a abordagem efetivamente utilizada para esta Tese.

4.5

Experimento 04: o experimento proposto por este trabalho

Considerando os resultados obtidos na seção anterior, buscamos uma forma de identificar as dificuldades existentes em cada classe, propondo uma abordagem binária. Desta forma, utilizamos a abordagem de Comitê de modelos, onde cada classe será treinada individualmente, e as decisões para a segmentação serão tomadas baseadas em decisões coletivas [73]. Esta metodologia foi explicada no Capítulo 3.

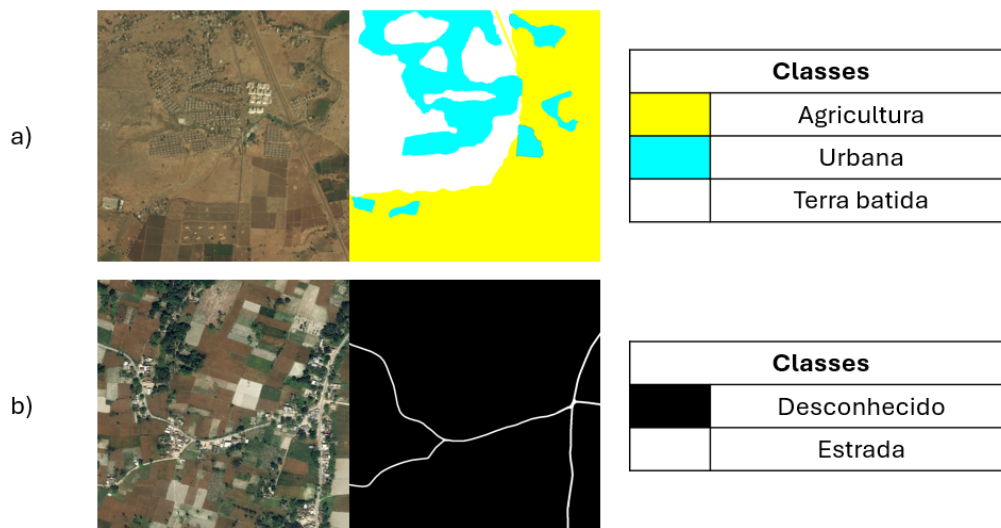


Figura 4.3: Em a), observamos uma imagem e sua respectiva máscara do banco de dados Land Cover Classification Challenge. Em b) observamos a imagem e a máscara do banco Road Extraction Challenge.

Sendo assim, os bancos de dados escolhidos foram o Land Cover Classification Challenge e o Road Extraction Challenge. Podemos observar exemplos

dos bancos de dados na Figura 4.3. Neste contexto, chamaremos de classe cada tipo de região dentro da imagem, sejam elas estradas, áreas urbanas, etc.

Para este treinamento, foram aproveitadas apenas as 803 imagens de treinamento do primeiro banco e as 6.226 imagens de treinamento do segundo banco, devido às imagens conterem anotações. Além disso, as classes “Agriculture land” e “rangeland” foram suprimidas pela classe “soft vegetation”.

O hardware utilizado foi um desktop com processador AMD Ryzen 5 5600 6-Core (3.5GHz), com 64 GB de RAM, e uma GPU NVIDIA GEFORCE RTX 4070 (12GB de RAM na GPU). Esta escolha se deve a limitações de equipamentos existentes para uso.

Como linguagem de programação, utilizamos Python. Anaconda foi utilizada como plataforma [79] e Visual Studio como IDE [80]. Como principais bibliotecas, utilizamos TensorFlow, NumPy, SciPy, Segmentation Models, Matplotlib, Open Neural Network Exchange (ONNX) entre outros. Estas bibliotecas foram escolhidas dada a grande comunidade existente e sua consolidação no mercado, de forma a permitir a continuidade nos trabalhos.

Para o modelo de IA, nossa abordagem focou nas classes do dataset de maneira binária. Nela, cada classe foi separada e categorizada utilizando a numeração de 0 para background e 1 para a classe respectiva.

Utilizamos, para a RN de cada classe, a arquitetura UNet com backbone em ResNet50. Nela, as imagens de entrada possuem shape (384,384,3), os pesos foram inicializados com “ImageNet” e a função de ativação utilizada foi “sigmoid”.

Como função de perda, foi utilizado o “binary focal dice loss” e Adam como otimizador. Já como métricas, utilizamos o IoU score. O dataset Road Extraction Challenge, por ser binário, não precisou ser separado e categorizado, bastando apenas o treinamento com o seu respectivo modelo.

A função de perda Binary Focal Dice Loss é observada na equação (4-7). Ela é composta pela soma da função de perda Binary Focal e da função de perda Dice. A função de perda Binary Focal (Eq. (4-5)) permite que os exemplos difíceis de classificar sejam penalizados mais fortemente em relação aos exemplos fáceis de classificar. Já a função de perda Dice prioriza a sobreposição entre previsões e alvos (equação (4-6)).

$$L_1(G, P) = -G \alpha (1 - P)^\gamma \log(P) - (1 - G) \alpha P^\gamma \log(1 - P) \quad (4-5)$$

$$L_2(p, r) = 1 - \left(1 + \beta^2\right) \frac{p \cdot r}{\beta^2 \cdot p + r} \quad (4-6)$$

$$L_{total} = L_1(G, P) + L_2(p, r) \quad (4-7)$$

Nas equações acima, G é o rótulo verdadeiro (ground truth), P é a probabilidade prevista, α é o fator de balanceamento entre classes, γ é o parâmetro de foco da Focal Loss, p e r representam *precision* e *recall* e β controla o peso relativo entre precisão e revocação.

Por ocasião do treinamento, foram estabelecidas 200 épocas como limite, usando early stopping em 30 épocas e batch size 8. As imagens foram divididas entre treinamento, validação e teste, seguindo a proporção 0.2 para teste, 0.8 para treinamento e validação e, dentro desta proporção, 0.8 para treinamento e 0.2 para validação.

Antes de serem usadas no treinamento propriamente dito, as imagens eram cortadas em patches menores de shape (384, 384, 3) valendo-se de duas formas. A primeira forma era sequencial, com stride de 0.75, onde o último patch era ajustado de forma a terminar no limite da imagem tanto no comprimento como na altura. A segunda forma de corte era randômica, de maneira a se obter patches diferentes dos padronizados, obtendo, assim, imagens complementares às sequenciais.

O tamanho escolhido para o patch corresponde ao tamanho de entrada da RN. Cabe destacar que o patch era selecionado para o treinamento apenas se ele estivesse com uma porcentagem da classe escolhida entre 0.2 e 0.8.

4.6

Metodologia para o treinamento do experimento 4

Conforme descrito na seção anterior, as imagens utilizadas foram as imagens de treinamento dos datasets dada a existência de máscaras para referência. Inicialmente, a imagem a ser usada como input em nosso modelo será cortada em patches menores, com shape (384, 384, 3). Os cortes foram realizados de maneira tanto sequencial como randômica, com o objetivo de alcançar imagens adequadas para o treinamento com redes neurais.

O treinamento com os modelos será realizado por classes individualmente. Neste contexto, modelos atuarão de forma binária, identificando um tipo de classe por vez, de forma a ser capaz de realizar predições apenas na sua classe.

Ao final do treinamento, as classes foram reunidas em uma única imagem. Para isso, 3 maneiras diferentes foram empregadas, de forma que geramos 3 imagens contendo todas as classes, cada uma com uma técnica diferente. A primeira considerava um “empilhamento” das classes, inseridas em ordem empírica, na sequência “bare_land” (terra batida), “soft_vegetation” (gramíneas e plantações), “arbor_vegetation” (florestas), “water” (água), “urban” (área urbana) e “road” (Road Extraction Challenge dataset).

Já na segunda tentativa, chamada de “probabilidade”, as imagens foram

montadas considerando o maior valor da probabilidade do pixel entre todas as classes. Neste caso, a classe contendo o pixel com a maior probabilidade por ocasião da predição preencheria o pixel correspondente.

A terceira maneira, chamada de “híbrida”, considerou as probabilidades apenas entre as classes do Land Cover Classification Challenge dataset. Neste caso, a probabilidade era aceita apenas se ela fosse superior a 0.1, sendo classificada como background caso não fosse aceita. Após esta etapa, é inserida a classe “road” do Road Extraction Challenge por cima da imagem formada. Podemos observar um fluxograma de treinamento da rede na figura 4.4

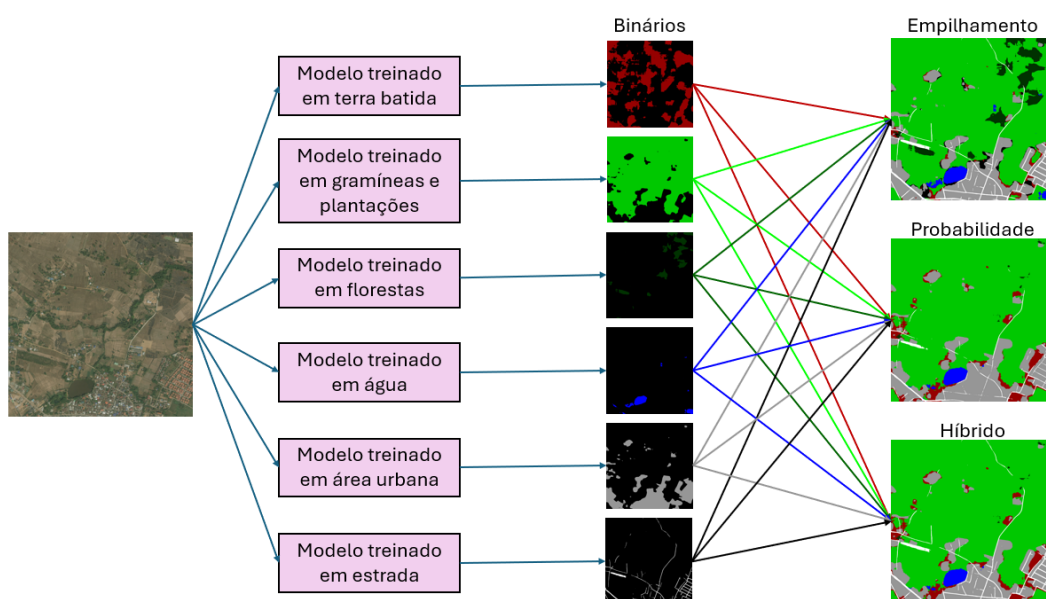


Figura 4.4: Fluxograma para treinamento da segmentação.

Em todas as maneiras de reunir as classes, o produto final é uma imagem em 3 canais com a extensão “.png”. Porém, considerando que o input para o ArcGIS, software utilizado na criação dos terrenos, é um arquivo shapefile, existe a necessidade da conversão destas imagens “.png” para o formato “.shp” (shapefile). A importância da utilização do arquivo shapefile se deve à necessidade de se terminar o tratamento das imagens para então exportar em extensão “.shp” para o blender e gerar um terreno virtual a ser utilizado no SVETT, bem como converter as classes existentes nos tipos de trafegabilidade e utilizá-los para o SJD. A seção a seguir analisará os resultados obtidos.

4.7

Resultados obtidos com o treinamento do modelo de IA

Considerando os treinamentos e predições realizados com os datasets DeepGlobe 2018 Satellite Image Understanding Challenge, onde destacamos o

Land Cover Classification Challenge e o Road Extraction Challenge, esta seção elucidará os resultados obtidos relativos a cada modelo treinado e empregado.

Nossa abordagem envolve primeiramente a exibição dos resultados tanto quantitativos como qualitativos, seguida das análises considerando os questionamentos realizados pela hipótese básicas no início deste trabalho, em que pese todo o conteúdo explicitado no capítulo 2.

Os resultados serão comparados de duas maneiras dentro deste mesmo experimento. A primeira maneira, realizada dentro do experimento 4, no caso do dataset Land Cover Classification Challenge, as predições serão comparadas às máscaras da imagem satélite por meio de treinamento supervisionado. A partir deste ponto, teremos métricas como IoU, F1-Score, Accuracy, Recall e Precision.

A segunda maneira ocorrerá fruto da classe atinente à estrada, cujo modelo foi treinado no Road Extraction Challenge e testado no Land Cover Classification Challenge dataset. Devido à ausência da classe estrada neste último dataset, não haverá valores relativos ao IoU por ocasião da predição. Sendo assim, serão realizadas comparações entre as imagens das máscaras e as imagens satélites.

Além das duas maneiras acima, realizaremos uma comparação com a abordagem anterior, que contém 6 classes, chamada de experimento 3 neste trabalho (seção 4.4). Para isso compararemos métricas como IoU, F1-Score, Accuracy, Recall e Precision.

Uma última comparação ocorrerá entre os experimentos 3 e 4 com trabalhos que utilizem o mesmo dataset empregado nesta Tese. Desta forma, seremos capazes de verificar nossa hipótese básica elencada no início deste capítulo. Cabe destacar que, em que pese o fato de termos realizado treinamentos de maneiras diferentes e, sendo assim, não termos como comparar os resultados, temos indícios de resultados satisfatórios.

Como resultados obtidos observamos, na tabela 4.1, os valores de precision, recall, accuracy, f1-Score e IoU do experimento 3, e suas respectivas médias. Além disso observamos, na tabela 4.2, as métricas precision, recall, accuracy, f1-Score, IoU por classe e mIoU do experimento 4. Todos os valores citados neste parágrafo foram obtidos por ocasião da predição.

Observamos ainda, na tabela 4.3, resultados de mIoU do trabalhos acadêmicos feito por [89], que constam no capítulo 3.

Observamos ainda, na Figura 4.6, exemplos das imagens obtidas considerando as três combinações de concatenação após a passagem pelo modelo do experimento 4. Em seguida observamos, na Figura 4.5, imagens da predição obtidas da mesma região do experimento 3, contendo 6 classes.

Tabela 4.1: Resultados das predições realizadas com o modelo do experimento 3 contendo 6 classes.

Metric	Precision	Recall	Accuracy	F1-Score	IoU
background	0.0001	0.0016	0.9959	0.0002	0.0126
water	0.363	0.319	0.9892	0.3097	0.3877
urban	0.5168	0.5282	0.9717	0.4975	0.5138
soft_vegetation	0.8392	0.7885	0.9026	0.7834	0.7281
arbor_vegetation	0.2213	0.2283	0.9576	0.2144	0.6094
bare_land	0.1967	0.1402	0.9503	0.1298	0.2719
Mean	0.3562	0.3343	0.9612	0.3225	0.4206

Tabela 4.2: Resultados obtidos nas diversas classes do dataset DeepGlobe através da abordagem da metodologia de Comitê de modelos, utilizando segmentações binárias.

Metric	Precision	Recall	Accuracy	F1-Score	IoU classe	mIoU
water	0.0564	0.0096	0.9649	0.0104	0.2306	0.5977
urban	0.4384	0.6871	0.9565	0.4774	0.4089	0.6736
soft vegetation	0.8048	0.836	0.8738	0.7968	0.7259	0.6412
arbor vegetation	0.2207	0.2494	0.9164	0.2253	0.2137	0.5365
bare land	0.1689	0.3112	0.8292	0.1913	0.1444	0.4765

Tabela 4.3: Resultados de mIoU alcançados por Desai e Ghose, em 2022, com treinamentos realizados com o Dataset DeepGlobe Land Cover.

mIoU	Labeled Ratio	2%	5%	12.50%
	s4GAN [XX1] (Baseline)	0.403 $\pm$ 0.035	0.486 $\pm$ 0.018	0.511 $\pm$ 0.007
	s4GAN + Entropy (Desai e Ghose)	0.469	0.513	0.554
	s4GAN + Margin (Desai e Ghose)	0.511	0.513	0.529

Buscando responder à primeira hipótese básica, que afirma que “os resultados obtidos fruto dos treinamentos dos modelos de redes neurais deste trabalho obtiveram resultados melhores que os já disponibilizados em trabalhos acadêmicos”, efetuaremos uma comparação com o trabalho escrito por Desai e Ghose, que consta no capítulo 3.

Observamos na tabela 4.3 que os mIoU obtidos pelo modelo s4GAN + Entropy superam em sua totalidade a abordagem do experimento 3 (Tabela 4.1). Porém, observando as predições sobre mIoU do experimento 4 (Tabela 4.2), a utilização de 2% dos dados anotados é superada por todas as predições

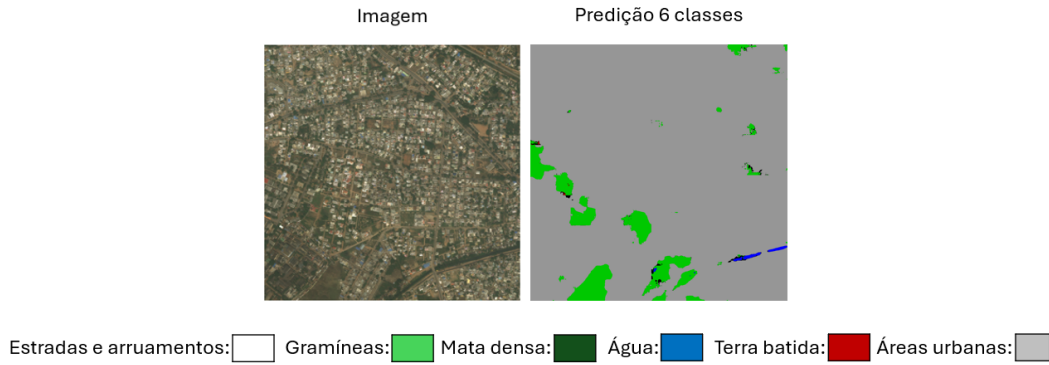


Figura 4.5: Predição realizada com a imagem do Experimento 3 contendo 06 classes.

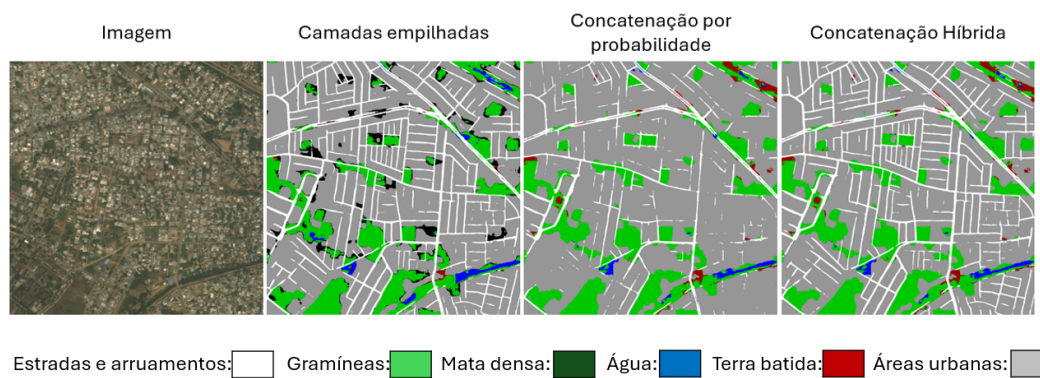


Figura 4.6: Comparação de uma mesma imagem prevista pelos 3 tipos diferentes de concatenação existentes neste trabalho.

realizadas. Com a fixação de 5% de dados anotados, apenas a classe “bare land” perde para as classificações realizadas. Já na utilização de 12.5% dos dados anotados, as predições das classes “bare land” e arbor vegetation possuem valores de mIoU abaixo do trabalho comparado.

Valendo-se agora da abordagem com o modelo s4GAN + Margin, apenas a classe “bare land” do experimento 4 não superou o trabalho comparado. Cabe destacar que, durante as buscas por trabalhos que utilizassem o mesmo dataset, o trabalho de Desai e Ghose foi o mais atual encontrado.

Realizaremos uma comparação agora da Tabela 4.1, do experimento 3, com a Tabela 4.2, do experimento 4. Nela podemos observar que as classes, no experimento 3, superam em todas as métricas as classes do experimento 4. Porém, quando consideramos a mIoU, todas as médias do experimento 4 superam as médias do experimento 3. Neste contexto, podemos observar os valores baixos do “background” do experimento 3 influenciando negativamente no cálculo da média.

Efetuando agora uma comparação visual entre a Figura 4.5, do experimento 3, e a Figura 4.6, do experimento 4, podemos observar que a predição

realizada pelo experimento 4, neste caso, foi melhor que a realizada pelo experimento 3.

Levando em consideração as informações acima, podemos comprovar que a hipótese básica em questão efetivamente foi comprovada, à medida que os modelos treinados pelo experimento 4 superaram, em sua maioria, os modelos do trabalho utilizado como referência.

Já referente à segunda hipótese básica, que afirma que “O terreno criado pode ser utilizado tanto para análise do terreno como em JG através do SJD.”, parte do processo envolve a criação de camadas vetorizadas no software ArcGIS, onde cada camada é modelada de forma independente. Considerando o fato de que nossa proposta entrega imagens contendo as classes em imagens separadas, em que pese o fato de sua conversão para imagens vetorizadas ser possível, podemos dizer que o resultado obtido neste capítulo contribui para a afirmação da abordagem proposta. Esta hipótese será confrontada novamente no capítulo 6, quando abordaremos o uso do SJD.

Buscando responder à terceira hipótese, que versa “O método de Comitê de modelos (do inglês, “Ensemble methodology”), para segmentação semântica binária, apresenta desempenho superior à abordagem multiclases tradicional, permitindo identificar com maior precisão as classes relacionadas à trafegabilidade do terreno em imagens de sensoriamento remoto”, vamos nos valer do fato das diferentes classes serem geradas de forma independente e, posteriormente, serem concatenadas. Neste sentido, obtivemos uma imagem que demarca os diferentes tipos de áreas existentes, que podem ser utilizadas para distinguir os diferentes tipos de trafegabilidade. Neste sentido, podemos afirmar, em parte, que nossa proposta atesta a hipótese básica proposta. Porém, abordaremos esta hipótese novamente no capítulo 6, de forma a comprovar se o material produzido serve efetivamente como CT.

Como possibilidade do seu uso no Corpo de Fuzileiros Navais (CFN), citamos a possibilidade de emprego do Simulador Virtual para Estudo Topográfico do Terreno (SVETT) [12] na análise da região afetada, de forma a permitir um estudo mais detalhado. Para estudos valendo-se de JG, existe a possibilidade de utilização do Sistema de Jogos Didáticos (SJD), simulador construtivo desenvolvido para treinamentos dos Estados-Maiores na condução das ações, já existente e utilizado abertamente no CFN.

Ambos os simuladores carecem da necessidade de criar o terreno virtualmente para o seu uso, fator dificultador devido principalmente ao tempo necessário para sua criação. Mais informações sobre o Sistema de Jogos Didáticos podem ser encontradas na seção 2.

Na próxima seção, falaremos sobre a ferramenta que utilizará o produto

gerado pelo estado final deste capítulo e a metodologia utilizada para o avaliar. Este software foi desenvolvido também com o foco neste trabalho, e foi chamado de Simulador Virtual para Estudo Topotático do Terreno (SVETT).

5

Simulador Virtual para Estudo Topotático do terreno - SVETT

Neste capítulo, descreveremos a solução proposta para a visualização do terreno virtual, criado no Capítulo 4, chamada de SVETT. Esta solução é composta tanto pela criação do terreno virtual como pelo Simulador, e seus recursos desenvolvidos e inseridos.

Além disso, versaremos, também, sobre os experimentos realizados com o SVETT junto aos Aspirantes da Escola Naval no processo de ensino sobre o giro do horizonte. Desta forma, seremos capazes de afirmar as seguintes hipóteses básicas estabelecidas no Capítulo 1:

- A integração de ambientes virtuais para estudo topotático do terreno, combinando recursos didáticos tradicionais com simulação virtual, potencializa o aprendizado e a confiança dos usuários (alunos e planejadores), sendo eficiente no treinamento antes do deslocamento ao campo real para análise de missão; e
- O pipeline desenvolvido, que abrange desde a segmentação semântica até a geração de terrenos virtuais para simuladores e jogos de guerra, é capaz de fornecer dados suficientes e condizentes com situações reais, possibilitando a realização de testes efetivos de planejamento militar em ambientes simulados.

O processo de criação do terreno virtual pode ser feito de várias maneiras, seja por meio do Blender ou utilizando asset da Unity chamado de Real World Terrain [108]. Para a inserção das camadas, utilizamos o asset Vegetation studio Pro [109], em que pese o fato deste último não estar mais disponível.

Após a criação do terreno, utilizamos o recurso do Vegetation Studio Pro para inserir a vegetação baseado em máscaras, permitindo inserir os diferentes tipos de vegetação no terreno. Este processo carece ainda de um recurso para inserir as residências em suas posições exatas, haja vista identificarmos apenas a área contendo construções. Estudos futuros visando a identificação positiva das construções podem ser realizados de forma a complementar este trabalho.

O SVETT foi desenvolvido com o objetivo de facilitar o entendimento atinente ao estudo topotático do terreno e padronizar os conhecimentos adquiridos. Nele, através da utilização do conceito de um Ambiente de Aprendizagem, buscamos oferecer recursos educacionais para entender melhor a análise

do terreno antes mesmo da presença em campo, sendo essa uma das hipóteses básicas levantadas no início de nosso trabalho.

Como mencionado na seção 1, para conduzir operações militares, os planejadores devem fazer uma análise abrangente do ambiente operacional, incluindo seus componentes, atores e suas relações [103]. Esses fatores influenciarão o processo de planejamento da operação, que normalmente envolve aspectos relacionados ao relevo, ao inimigo, à vegetação e ao clima, conhecidos em conjunto como aspectos topotáticos do terreno [14].

Para ensinar e treinar militares sobre esse tipo de atividade, os alunos costumam ser levados a diferentes regiões em ET para entender o que deve ser considerado para seu estudo. Mesmo participando desses ET, é bastante difícil para o instrutor explicar aos alunos o assunto, dada a abstração necessária para visualizar as medidas de coordenação para uma operação em uma região.

Com base na tendência mundial em se buscar a simulação para treinamento Tropa [20], o primeiro contato com o terreno poderia ser feito já em sala de aula por meio da imersão em uma Área de Treinamento Sintética. Neste contexto, se os alunos já chegassem ao local de treinamento tendo um contato prévio com a região, as dúvidas mais básicas já seriam resolvidas dentro de uma sala de aula. Desta forma, as atividades mais complexas ou peculiares seriam realizadas no ET.

Ainda neste contexto, com a disponibilidade de recursos que o terreno real não permite, como desenhar no terreno para uma explicação e inserção de simbologias, os resultados dessa interação trariam grandes benefícios do ponto de vista educacional, criando condições para um melhor estudo.

Nesse sentido, desenvolvemos o SVETT. Todos os recursos estão disponíveis para instrutores e alunos, para serem usados por meio de três interfaces, não apenas durante as aulas, mas também em interações individuais ou em grupo, antes ou depois de qualquer estudo do terreno. Estas interfaces serão explicadas mais à frente, na seção 5.1.

Para o seu desenvolvimento, usamos o Unity [104] como mecanismo de desenvolvimento, com assets 3D e objetos gráficos criados usando o Blender [105]. Além disso, a camada de comunicação foi possível usando Mirror [106] e o armazenamento dos dados através do PostgreSQL [107]. Para fins de desenvolvimento e experimento, usamos um desktop com uma RTX3070Ti conectada a três TVs de 82", um tablet Galaxy Tab A8 e dois OCULUS QUEST 2 Head Mounted Display (HMD). Esses dispositivos foram conectados uns aos outros por meio de um roteador, mas não houve testes com foco na quantidade de conexões devido a limitações orçamentárias.

Como o objetivo do simulador é ajudar no estudo dos aspectos topográficos

ficos e táticos de qualquer região, a área a ser usada com essa finalidade pode ser criada de acordo com as necessidades dos instrutores ou obtida de qualquer fonte aberta. A área escolhida geralmente observa os locais que serão usados em ET, devido ao seu potencial para aprender e explorar os assuntos abordados em sala de aula. Os objetos na cena (árvores, casas, instalações, etc.) são todos estáticos e não podem ser movidos.

Para facilitar as condições de ensino dos instrutores e a interação entre os alunos, o SVETT foi desenvolvido explorando recursos normalmente vistos em sala de aula, como uma caneta ou uma simbologia, fornecendo-os para serem usados na cena principal. Nela, todos os usuários conectados são inseridos na mesma cena, e podem fazer todas as suas anotações e inserir seus objetos em tempo real, sendo observados pelos demais, possibilitando estudos mais dinâmicos.

Alguns recursos incluem desenhar o terreno de estudo para ajudar a marcar medidas importantes. Entre elas, citamos os limites entre pelotões, marcar locais como rios, inserir objetos estáticos. Outro recurso importante inserido está na inserção de simbologias militares.

5.1

Interfaces do SVETT

A interação no cenário principal ocorre por meio de três interfaces (Figura 5.1), criadas com o objetivo de aumentar a acessibilidade para o estudo e a exploração dos aspectos do terreno. Ela pode ser utilizada tanto pelo instrutor quanto pelos alunos, que também podem observar e interagir entre si no Ambiente Virtual. Todas se diferenciam pelos recursos disponíveis, pelo modo de visualização e pelo hardware.

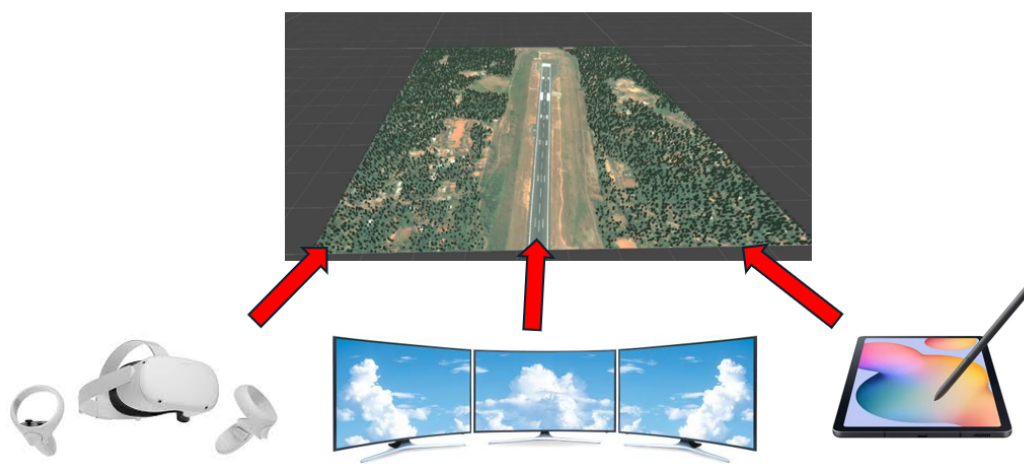


Figura 5.1: Interfaces do SVETT. Nela, os 3 tipos de equipamentos interagem dentro do terreno virtualmente criado

A primeira interface é chamada de Interface de projeção (Figura 5.2). Ela pode conectar até 8 telas simultaneamente, e utilizar uma TV ou mesmo um projetor para isso. Essa interface é capaz de observar a cena principal e mudar sua posição, voando para o novo local.



Figura 5.2: Interface de projeção mostrando a mesma representação esboçada na Interface Construtiva na Figura 5.3.

A principal vantagem dessa interface é a possibilidade de interagir com uma quantidade maior de usuários ao mesmo tempo. Ao combinar com a interface construtiva, é possível aumentar a interação fazendo anotações e observando-a em ambos os tipos de visualização. Um aspecto importante é que, ao dar instruções aos alunos, o instrutor pode observar seus rostos para ver suas expressões e detectar se há algum problema no processo de aprendizagem. Outras vantagens incluem observar as anotações feitas pelas outras interfaces em tempo real.

A segunda interface é chamada de interface construtiva (Figura 5.3). Podemos utilizá-la com um tablet ou um laptop, onde ambos são conectados sem fio ao roteador. Nela, a área de estudo pode ser vista como uma imagem satélite ou como um mapa militar. Em ambos os casos, essa observação é feita a partir da vista superior, e com uma câmera ortográfica apontada para baixo. Além disso, podemos observar a posição e a direção das demais interfaces, com um padrão de simbologia exclusivo.

Essa interface é capaz de desenhar o terreno, chamar a atenção das outras interfaces para um local específico e inserir objetos que são observados por todas as outras interfaces. Além disso, ela pode teletransportar as outras

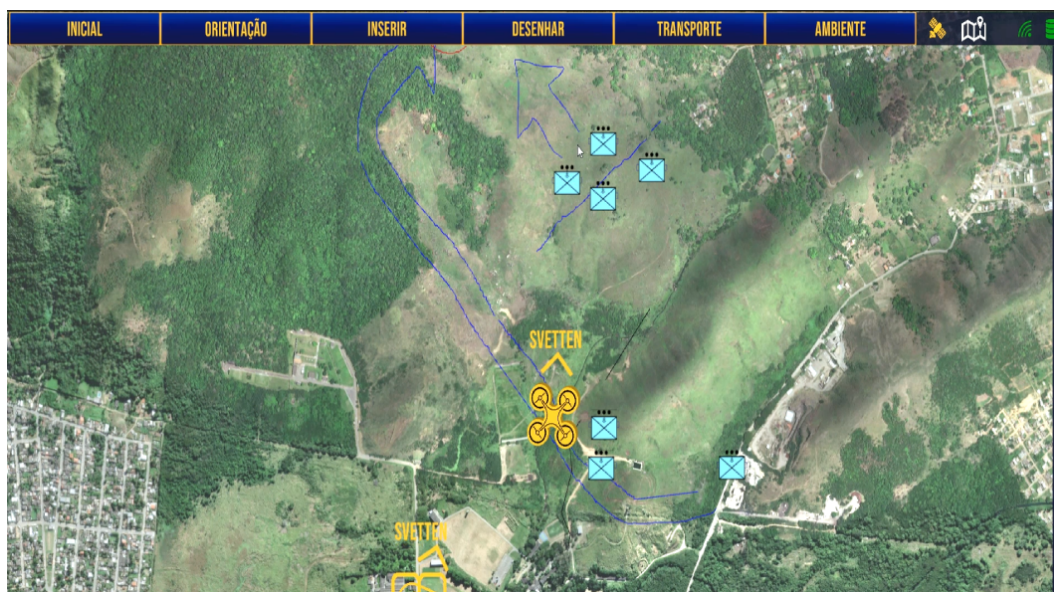


Figura 5.3: Interface construtiva mostrando uma representação da interface do projetor em amarelo no centro da imagem e alguns esboços no terreno.

interfaces conectadas para qualquer local no terreno e controlar a interface de projeção.

A vantagem da Interface Construtiva é que podemos observar o estudo de cima para baixo, escrever todas as medidas planejadas anteriormente e inserir objetos como rascunhos em um mapa militar. Essas medidas e objetos inseridos serão vistos pelos outros usuários por meio de suas interfaces. Além disso, é possível armazenar os estudos realizados para posterior utilização.

A interface de realidade virtual é a última e usa um óculos de realidade virtual sem fio para imersão na área de estudo. Essa interface interage com as outras no terreno com um avatar usando um uniforme militar. O usuário será posicionado acima de uma plataforma que pode se mover para qualquer lugar na área, controlando a plataforma com o controle do óculos.

Essa interface pode desenhar no terreno, inserir simbologias, medir distâncias, inserir uma bússola magnética, inserir objetos e grades como os mapas militares, que são observados pelas outras interfaces conectadas.

A principal vantagem dessa interface é a possibilidade de aplicar um estudo mais detalhado devido à sua imersão e à noção espacial. Outra vantagem está na possibilidade de ver todos os objetos e anotações em qualquer ponto do terreno como se estivesse em um ET e, com isso, ter uma noção melhor do planejamento. Uma visão da Interface de Realidade Virtual é mostrada na Figura 5.4.

De forma a respondermos às hipóteses básicas elaboradas, desenvolvemos o SVETT e o avaliamos com 54 alunos, segundo a metodologia mostrada na seção a seguir.

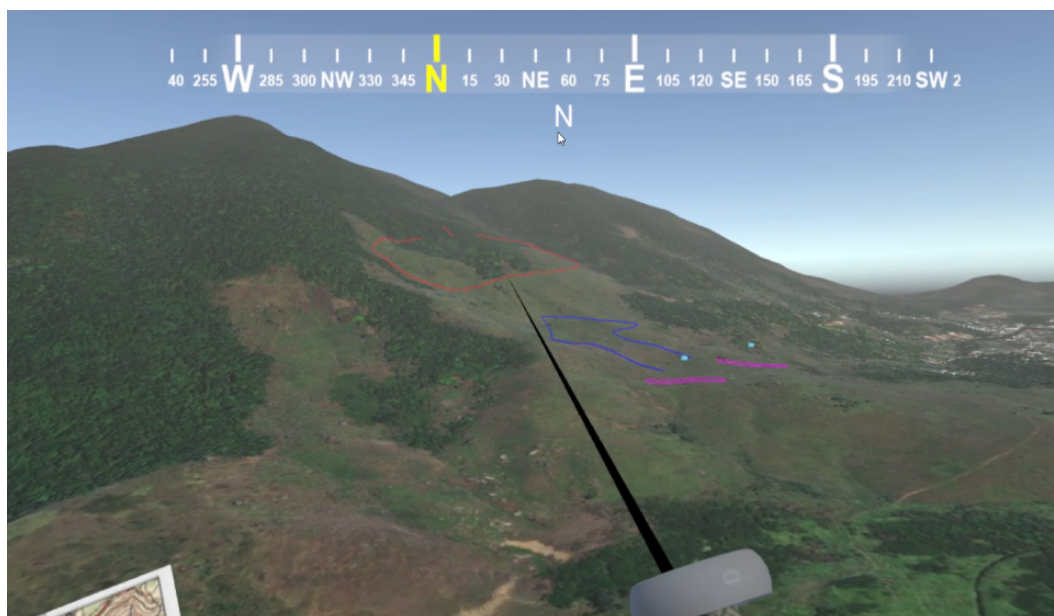


Figura 5.4: Interface de realidade virtual apontando para uma posição inimiga representada no terreno.

5.2

Metodologia de experimento com o SVETT

De forma a respondermos nossas hipóteses básicas, desenvolvemos o SVETT e o avaliamos com 54 alunos. A intenção é obter informações de forma a comprovar que, usando o SVETT entre a teoria e a ida ao ET, os alunos o acharão melhor quando participarem do Exercício devido ao seu primeiro contato com a região ainda em sala de aula.

Esta seção tem o objetivo de explicar o conjunto de testes realizados para avaliar o SVETT junto aos Aspirantes da Escola Naval, detalhando os cenários, bem como suas respectivas cinemáticas. Todas as interações foram gravadas em áudio e vídeo, onde foram observadas dificuldades tanto por parte dos instrutores quanto por parte dos alunos.

O propósito foi verificar se é possível que a concatenação do conceito de um Ambiente de Aprendizagem (LE) e dos recursos normalmente vistos em sala de aula possa oferecer recursos educacionais para entender melhor a análise do terreno antes mesmo da presença em campo.

Realizamos uma avaliação do SVETT com alunos da Escola Naval para identificar se o uso dos diferentes recursos disponíveis, em apoio a uma instrução sobre o estudo topológico e tático do terreno, responde às hipóteses estabelecidas na introdução deste trabalho.

Para cada tipo de avaliação, foi realizado um teste piloto para verificar possíveis falhas e melhorias, que posteriormente foi descartado. Como se tratava de um momento onde a instrução teórica sobre o assunto já havia

acontecido e eles ainda iriam ter o primeiro contato com o terreno real, os alunos não foram avaliados com notas pelos instrutores.

Para isso, foram criados dois cenários para os alunos. No primeiro cenário, chamado de cenário “sala de aula”, eles foram inseridos em uma aula sobre o estudo topotático do terreno usando o simulador. No segundo cenário, chamado de cenário “exercício no terreno”, o estudo ocorreu em um ET na mesma área utilizada virtualmente no simulador.

Em ambos os casos, cada cenário durou cerca de 50 minutos. No cenário “sala de aula”, observado na Figura 5.5, o instrutor interagiu com 54 alunos utilizando o SVETT, o que levou à utilização dos recursos instrucionais disponíveis para transmitir conhecimento ou sanar dúvidas. Nesse contexto, foi possível observar a interação de todos com o aplicativo. Durante o cenário, o instrutor utilizou um tablet e os alunos observaram as interações através de televisores. Testes com os óculos de RV foram realizados em outras oportunidades.



Figura 5.5: Cenário “sala de aula”.

No cenário “exercício no terreno” não houve uso de simulador, sendo utilizados apenas os recursos usuais do instrutor, que são a carta militar e a bússola. Neste cenário, apenas 30 dos 54 alunos que participaram do cenário “sala de aula” estiveram presentes. Esta falta de participantes se deve ao fato de os alunos serem de turmas diferentes, sendo obrigados a cumprir uma grade curricular em local diferente neste dia.

O objetivo dos cenários criados está relacionado à utilidade do simulador como ferramenta para introduzir os alunos ao estudo básico do terreno, antes mesmo de sua chegada ao local. Deve-se observar que, após o uso do simulador

para transmitir o conteúdo básico, o assunto abordado no simulador, em teoria, não seria mais visto. Os assuntos a serem abordados nos exercícios em locais reais focariam em um conteúdo mais avançado, para obter um nível mais alto de aprendizado e, conseqüentemente, aumentar a qualidade do ET.

Em ambos os cenários, todos os alunos participaram juntos em uma única aula. Cada participante, durante os cenários, ao chegar ao local do experimento, recebeu instruções iniciais sobre a cinemática e foi convidado a assinar um termo de consentimento. Após a assinatura, os usuários foram posicionados no local previamente estabelecido para o experimento. Antes de iniciar as atividades, eles se familiarizaram com os gestos e comandos possíveis. Em ambos os cenários, todos os participantes receberam instruções do mesmo instrutor, que seguiu um currículo previamente estabelecido.

Depois de concluir a instrução do primeiro cenário, os alunos foram solicitados a preencher dois questionários. O primeiro tinha como objetivo obter conhecimento sobre sua experiência como aluno, sua experiência no estudo do terreno e experiências com simuladores que utilizam Realidade Virtual .

O segundo questionário buscou avaliar as primeiras impressões da instrução com o sistema. Para este segundo questionário do "cenário de sala de aula", observamos cinco perguntas com uma escala Likert de 1 a 7, em que 1 e 2 são valores negativos, 3, 4 e 5 representam valores neutros e 6 e 7 representam valores positivos. Depois disso, observamos quatro perguntas com respostas dissertativas [110].

Ao final da instrução do segundo cenário, os alunos preencheram dois questionários. O primeiro foi o “Student Satisfaction and Self-Confidence in Learning”, traduzido para o português por Almeida et al. [111]. Para esse questionário, a palavra “médico-cirúrgico” na questão 7 e a palavra “clínico” na questão 8 foram omitidas para ajustar as perguntas ao contexto correto. O segundo questionário foi criado para avaliar o conceito da instrução após a conclusão dos dois cenários. Para os questionários no “cenário de exercício de campo”, observamos uma graduação em escala Likert de 1 a 5, em que 1 e 2 representam valores negativos, 3 representa valor neutro, e 4 e 5 representam valores positivos. A próxima seção mostra alguns dos resultados preliminares obtidos e uma discussão.

5.3

Resultados obtidos com o SVETT

Ao longo desta seção, mostraremos alguns dos resultados inicialmente obtidos, analisando conforme a busca por afirmações às nossas hipóteses

básicas. Além disso, trabalhos futuros obtidos a partir da discussão nesta seção serão propostos no capítulo 7.

Em relação ao questionário 1, que qualifica os participantes, foi identificado que os alunos possuem idade entre 20 e 24 anos, com gênero de 53 homens e 3 mulheres. Além disso, eles estão fazendo o mesmo curso, no qual identificamos o padrão do público-alvo que utilizará o simulador para treinamento. Portanto, as avaliações não incluem pessoas de idade avançada ou crianças, pois elas não são empregadas em atividades relacionadas ao curso. Se houver oportunidade e necessidade, serão realizados novos testes com um público-alvo diferente.

Outro detalhe é o fato de todos estarem abordando inicialmente a disciplina sobre o estudo do terreno, na qual há predominância da parte teórica, possibilitando o uso do simulador no tempo correto, antecedendo o ET. Outro aspecto é o fato de a maioria dos alunos ter passado por duas ou nenhuma região de estudo, o que nos leva a inferir que os alunos têm pouca ou nenhuma experiência no estudo do terreno. Concluindo o primeiro questionário, 32 dos participantes usaram óculos de realidade virtual entre uma vez por ano e uma vez por mês e os demais nunca os usaram, o que denota pouco ou nenhum conhecimento ou familiaridade com simuladores de RV.

Já perguntas do questionário 2 apresentadas na tabela 5.1 e as respostas apresentadas na Figura 5.6, possuem uma escala Likert de 0 a 7 pontos e avaliam a utilização da ferramenta na instrução. Sobre a pergunta “A condução do giro do horizonte foi...”, esta obteve uma média de 4,28. A observação do terreno teve uma média de 5,69. Entretanto, devido à falta de observação prévia do local e à pouca experiência, inferimos que essa graduação se limita à observação de um terreno em um mundo virtual, não sendo possível compará-la com a realidade do local. Considerando esses dois valores, podemos inferir também que houve uma avaliação positiva, apesar da existência de alguma dificuldade na observação do terreno. As três últimas questões obtiveram valores acima de 6,37, de onde se pode afirmar uma aceitação muito positiva para o uso da ferramenta.

Sobre as questões dissertativas, podemos destacar como aspectos positivos a possibilidade de estudar a área antes de estar nela e a facilidade de estudar o terreno com o simulador. Como aspectos negativos, destacamos a baixa qualidade da resolução da tela, o pouco detalhamento da área e o pouco tempo de aprendizado no simulador. Comparando os aspectos negativos destacados com as respostas anteriores, podemos inferir que a baixa resolução da tela atrapalhou a observação do terreno.

Os questionários seguintes foram aplicados no cenário “exercício no ter-

Tabela 5.1: Perguntas usadas no questionário 2.

QUESTIONÁRIO 2	
NÚMERO	PERGUNTA
1	A condução do giro do horizonte foi...
2	A observação do terreno foi...
3	De modo geral em achei a ferramenta de instrução...
4	Receber instrução com essa ferramenta foi...
5	Com relação às respostas que deste neste questionário, eu estou...

reno”. O questionário “Student Satisfaction and Self-Confidence in Learning” tem uma escala Likert de 5 pontos e é dividido em dois subitem: “Satisfação com o aprendizado atual” e “Autoconfiança no aprendizado”. Podemos observar as perguntas na tabela 5.2. Para o primeiro subitem mostrado na Figura 5.7, a pergunta “Os métodos de ensino usados nesta simulação foram úteis e eficazes.” recebeu uma nota de 4,47, a partir da qual podemos afirmar que a combinação do simulador com o estudo do exercício de campo foi positiva.

Tabela 5.2: Questões do subitem “Satisfação com a aprendizagem atual” do questionário “Student Satisfaction and Self-Confidence in Learning”.

SATISFAÇÃO COM A APRENDIZAGEM ATUAL	
NÚMERO	QUESTÃO
1	Os métodos de ensino utilizados nesta simulação foram úteis e eficazes.
2	A simulação forneceu-me uma variedade de materiais didáticos e atividades para promover a minha aprendizagem do currículo.
3	Eu gostei do modo como meu professor ensinou através da simulação.
4	Os materiais didáticos utilizados nesta simulação foram motivadores e ajudaram-me a aprender.
5	A forma como o meu professor ensinou através da simulação foi adequada para a forma como eu aprendo.

A pergunta “A simulação me forneceu uma variedade de materiais e atividades de aprendizagem para promover meu aprendizado do currículo” recebeu uma nota de 4,20. Combinando essa resposta com a resposta do questionário 2 “Em geral, achei a ferramenta de instrução...” (6,50 em uma escala Likert de 7 pontos), podemos inferir que as ferramentas desenvolvidas para apoiar a instrução ajudaram os alunos a aprender a matéria.

Com relação à nota da questão 3: “Gostei da maneira como meu instrutor ensinou a simulação” (4,67), combinada com a questão 5 “A maneira como meu(s) instrutor(es) ensinou(aram) a simulação foi adequada à minha maneira

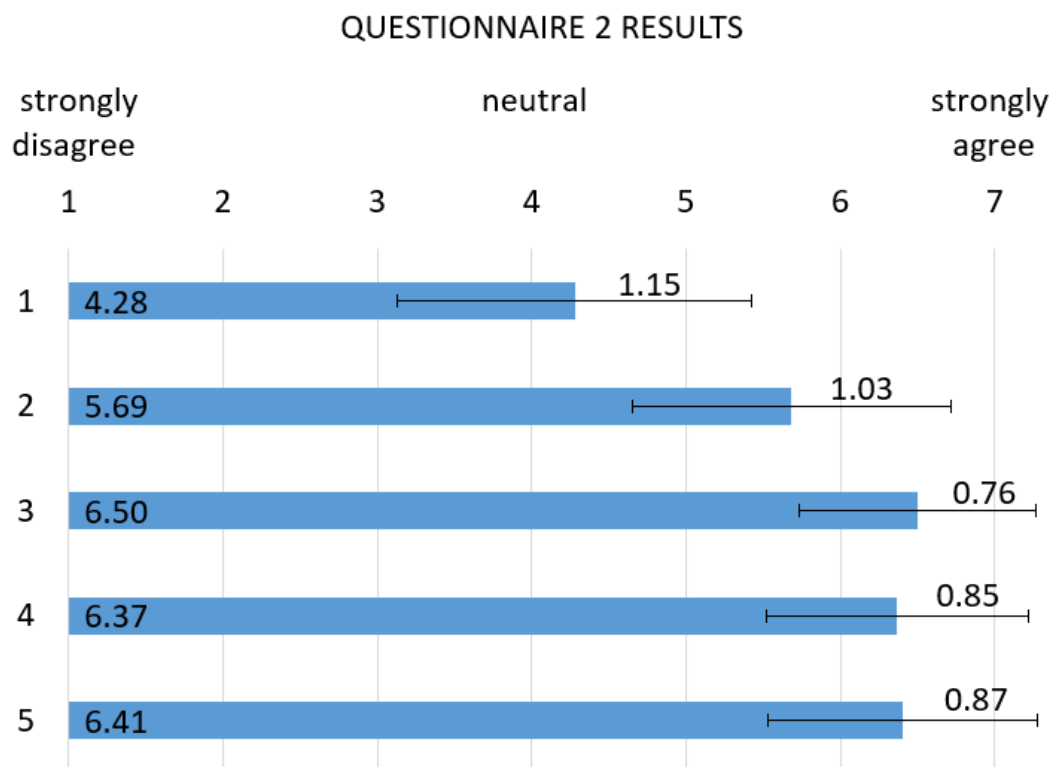


Figura 5.6: Questionário 2, utilizado para avaliar o uso da ferramenta nas instruções.

de aprender” (4,67), podemos concluir que o instrutor, de alguma forma, aprendeu a explorar as ferramentas e chamou a atenção dos alunos.

Sobre a resposta 4: “Os materiais didáticos usados nessa simulação foram motivadores e me ajudaram a aprender”, obtivemos uma nota de 4,43. Considerando todas as respostas desse subitem e a nota mínima obtida (4,20), já que esse questionário foi aplicado após os exercícios de campo, podemos afirmar um resultado positivo no subitem “Satisfação com o aprendizado atual”. Desta forma, combinando as questões desta primeira parte, em que pese estarmos tratando primeiramente do SVETT, podemos afirmar nossa hipótese básica, que diz que “o pipeline desenvolvido, que abrange desde a segmentação semântica até a geração de terrenos virtuais para simuladores e jogos de guerra, é capaz de fornecer dados suficientes e condizentes com situações reais, possibilitando a realização de testes efetivos de planejamento militar em ambientes simulados”.

Sobre o subitem “Autoconfiança no aprendizado”, mostrado na Tabela 5.3 e na Figura 5.8, a maioria das notas foi maior que 4. Porém, a pergunta “Estou confiante de que estou dominando o conteúdo da atividade de simulação que meus instrutores me apresentaram” recebeu nota 3.57, a pergunta “Eu sei como usar as atividades de simulação para aprender aspectos críticos dessas habilidades” recebeu nota 3,7 e a pergunta “É responsabilidade do



Figura 5.7: Média e desvio padrão das respostas do subitem “Satisfação com a aprendizagem atual” do questionário “Student Satisfaction and Self-Confidence in Learning” com escala Likert de 5 pontos.

instrutor me dizer o que eu preciso aprender sobre o conteúdo da atividade de simulação durante a aula...” recebeu nota 3,97. Analisando as respostas e, considerando o fato de o exercício de campo ter sido a segunda instrução, considerando ainda que a segunda maior nota (4,50) foi “Estou confiante de que estou desenvolvendo as habilidades e obtendo o conhecimento necessário com esta simulação para realizar as tarefas necessárias em um ambiente”, em que pese um dos aspectos negativos do questionário 2 ter sido “pouco tempo de aprendizado no simulador”, podemos inferir que o tempo no simulador não foi suficiente para que os alunos se sentissem mais confiantes. Por outro lado, considerando a nota mais alta (4,70) ter sido dada à pergunta “Meus instrutores usaram recursos úteis para ensinar a simulação” (com o menor desvio padrão, 0,46), nos permite afirmar que os recursos desenvolvidos no simulador ajudaram os alunos no processo de ensino e aprendizagem.

Sobre o questionário 3 mostrado na Tabela 5.4 e na Figura 5.9, que avalia a combinação de estudo em sala de aula e estudo de exercício em campo, a pergunta “Eu me senti mais seguro quando recebi a instrução em campo depois de ter recebido a instrução por meio do simulador” (nota 4,69), combinada com a pergunta “O conceito de instrução que consiste no uso do simulador seguido de exercício em campo na mesma região do simulador facilitou meu

Tabela 5.3: Questões do subitem “autoconfiança na aprendizagem” do questionário “Student Satisfaction and Self-Confidence in Learning”.

AUTOCONFIANÇA NA APRENDIZAGEM	
NÚMERO	QUESTÃO
6	Estou confiante de que domino o conteúdo da atividade de simulação que meu professor me apresentou.
7	Estou confiante que esta simulação incluiu o conteúdo necessário para o domínio do currículo.
8	Estou confiante de que estou desenvolvendo habilidades e obtendo os conhecimentos necessários a partir desta simulação para executar os procedimentos necessários em um ambiente.
9	O meu professor utilizou recursos úteis para ensinar a simulação.
10	É minha responsabilidade como o aluno aprender o que eu preciso saber através da atividade de simulação.
11	Eu sei como obter ajuda quando eu não entender os conceitos abordados na simulação.
12	Eu sei como usar atividades de simulação para aprender habilidades.
13	É responsabilidade do professor dizer-me o que eu preciso aprender na temática desenvolvida na simulação durante a aula.

aprendizado.” (nota 4,66 com o segundo menor desvio padrão, 0,47) nos ajuda a afirmar a nossa quarta hipótese básica: “A integração de ambientes virtuais para estudo topotático do terreno, combinando recursos didáticos tradicionais com simulação virtual, potencializa o aprendizado e a confiança dos usuários (alunos e planejadores), sendo eficiente no treinamento antes do deslocamento ao campo real para análise de missão’.

O próximo capítulo detalhará o processo de criação do terreno virtual para o SJD e a criação da carta de trafegabilidade, analisando resultados de pesquisas e questionários.

Tabela 5.4: Perguntas do questionário 3.

QUESTIONÁRIO 3	
NÚMERO	PERGUNTA
1	A condução do giro do horizonte foi muito fácil.
2	O entendimento do estudo do terreno foi muito fácil.
3	A observação do terreno foi muito boa.
4	Receber instrução no terreno real após ter passado pelo SVETT foi muito divertida.
5	Me senti mais seguro quando recebi a instrução no terreno depois de ter recebido a instrução pelo simulador.
6	O conceito de instrução composto pelo uso do simulador seguido de Exercício no terreno na mesma região do simulador facilitou meu aprendizado.
7	Com relação às respostas que deste neste questionário, eu estou muito seguro.

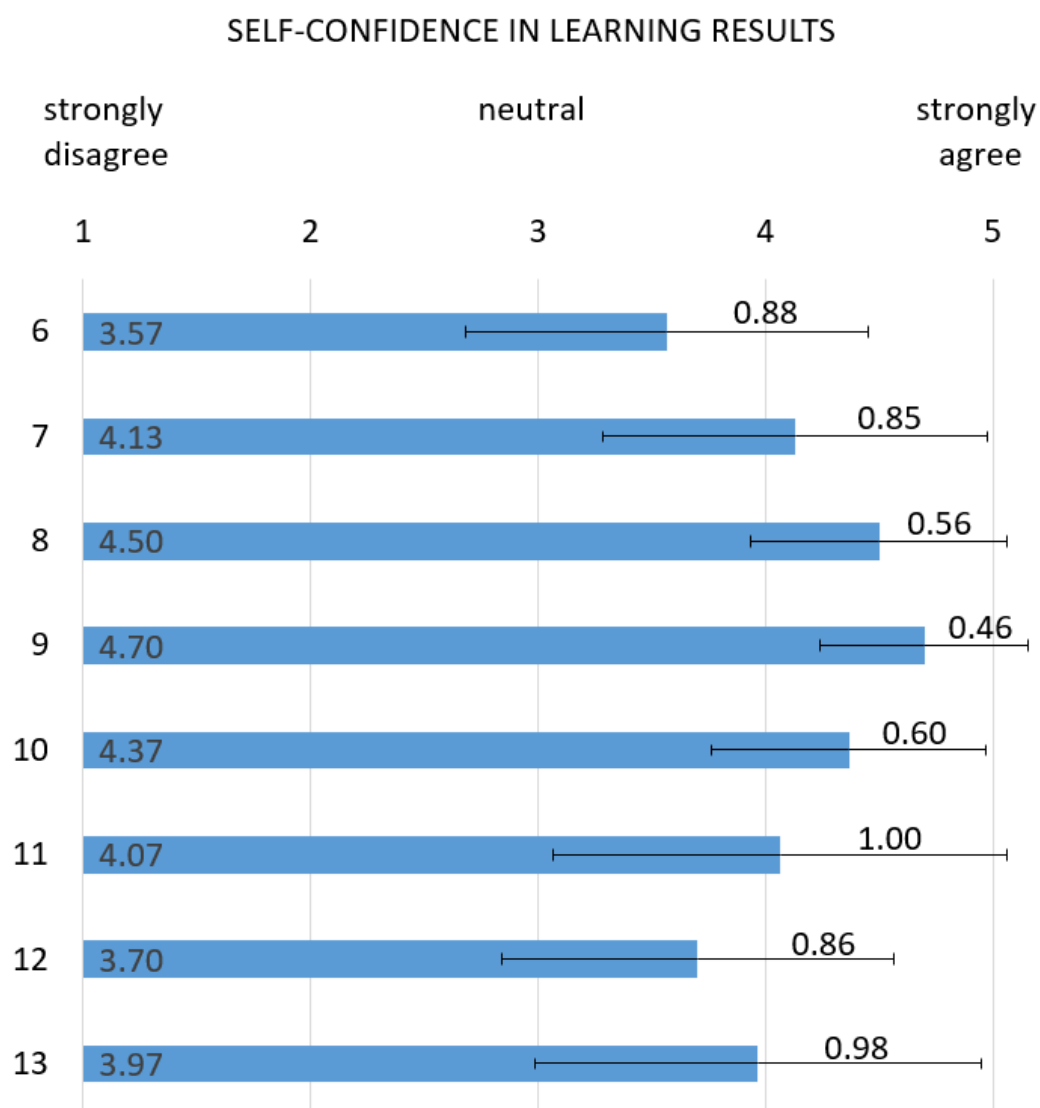


Figura 5.8: Média e desvio padrão das respostas do subitem “autoconfiança na aprendizagem” do questionário “Student Satisfaction and Self-Confidence in Learning” com escala Likert de 5 pontos.

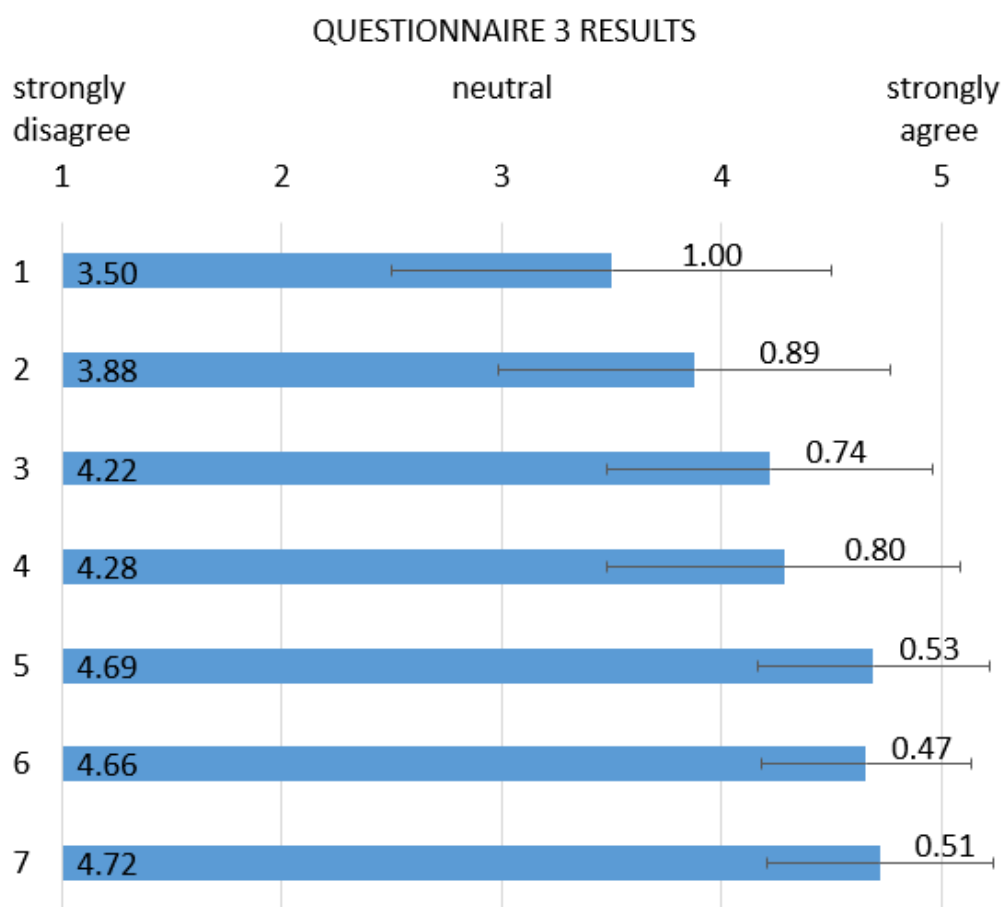


Figura 5.9: Questionário 3, que avalia o processo de aprendizagem iniciando em sala com a teoria, efetuando o uso do SVETT e participando do ET.

6

Criação da Carta de Trafegabilidade e terreno para o SJD

Neste capítulo descreveremos os passos para as imagens obtidas no Capítulo 4 serem utilizadas no processo de criação do terreno do SJD. Além disso, exibiremos os resultados de uma pesquisa realizada com planejadores militares sobre a usabilidade das 3 imagens concatenadas obtidas na seção 4.7.

Desta forma, buscaremos responder às seguintes hipóteses básicas estabelecidas no Capítulo 1:

- O pipeline desenvolvido, que abrange desde a segmentação semântica até a geração de terrenos virtuais para simuladores e jogos de guerra, é capaz de fornecer dados suficientes e condizentes com situações reais, possibilitando a realização de testes efetivos de planejamento militar em ambientes simulados.
- Os produtos gerados pela segmentação semântica – particularmente as classes identificadas – podem ser utilizados para criar cartas de trafegabilidade que atendem aos requisitos operacionais das Operações militares, acelerando o processo de análise e planejamento do terreno; e

Primeiramente falaremos sobre a criação dos terrenos utilizados no sistema de jogos didáticos.

6.1

Processo manual de criação do terreno para o SJD e a utilização da solução proposta

Tanto a elaboração da CT como a criação dos terrenos do SJD são realizadas de forma manual. Este processo segue uma série de passos trabalhosos que serão resumidos nesta seção. A intenção é mostrar a dificuldade em se criar este tipo de recurso, sendo um fator impeditivo para a utilização de simulação no planejamento de operações reais.

A base para a criação da CT são as elevações de uma região. Elas servem como referência por não sofrerem praticamente nenhuma alteração ao longo dos anos. Estas elevações podem ser obtidas por meio de CM ou através de imagens de satélites especificamente utilizadas para se criar o Modelo Digital de Elevação (MDE). O MDE pode ser obtido em fontes como o 3D Elevation Program (3DEP) [112].

De posse do relevo da região, o desenvolvedor criará as curvas de nível da área, que são linhas que unem todos os pontos com a mesma altitude, de

forma a se obter as elevações presentes na região. Após a criação do relevo, são inseridas as vegetações, hidrografia e áreas urbanas.

Para ajudar no detalhamento, são utilizadas imagens de fontes abertas como “Google Earth” [113] e “Open Street Maps” [114], além de fotografias da região de interesse. Todo o processo é complementado com conhecimentos existentes nas Unidades especializadas.

Cada mancha contendo um tipo de vegetação (mata densa, média ou gramínea), hidrografia, área urbana e rodovias é contornada, marcando assim que toda a área dentro é relativa a este tipo específico. A este processo chamamos de vetorização. No final, estas manchas são utilizadas como input em um algoritmo criado no CSimCFN, que utiliza o MDE como coeficiente, criando assim as trafegabilidades. Observamos na Tabela 6.1 as cores utilizadas conforme as mobilidades permitidas.

Tabela 6.1: Trafegabilidades existentes no Sistema de Jogos Didáticos (SJD).

DESCRIÇÃO	RGB		
Força Especial	255	0	0
A Pé / Forças Especiais	255	255	150
Roda / Lagarta / A pé / VtrAnf / Forças Especiais	0	200	0
Lagarta / VtrAnf / A pé Forças Especiais	150	110	30
Navio / VtrAnf / Forças Especiais	100	100	200
Anfibio / Forças Especiais	100	100	100
Qualquer Terreno	0	120	0

O software utilizado para a criação do terreno virtual é o ArcGIS [115]. Nele, são vetorizados em camadas os tipos de regiões existentes, como, por exemplo, vegetação densa (florestas), vegetação rala (gramíneas), água, áreas urbanas, entre outras. Cada camada, depois de vetorizada, recebe uma classificação por característica onde, por exemplo, se identifica uma estrada pavimentada ou terra batida, se estabelece a altura média dos quarteirões em áreas urbanas, etc. Cabe destacar que esta vetorização deve ser feita detalhadamente, o que torna o processo de criação da região extremamente demorado e trabalhoso.

Depois da vetorização e classificação, é feita a junção dessas camadas e a combinação com o Modelo Digital de Elevação (MDE), que são as representações matemáticas do relevo da região já disponíveis no site da USGS [116]. De posse desta região vetorizada, classificada e processada com o MDE, suas informações são exportadas em arquivo “.shp” (shapefile) para finalização no SJD.

Para esta CT, cada pixel da imagem será um tipo de terreno, que pode ser uma floresta, uma área urbana, um rio, entre outros. Este pixel deve

ser identificado de maneira fidedigna com a realidade, atribuindo a ele uma mobilidade aceita pelo simulador (deslocamento a pé, em viatura sobre rodas, viatura sobre lagarta, como é o caso do Carro-Lagarto Anfíbio (CLAnf)).

No caso dos inputs para a criação dos terrenos do SJD, um código desenvolvido na linguagem de programação python recebe estas camadas vetorizadas, e efetua uma série de cálculos de forma a converter as cores das camadas em códigos (letras) a serem lidos pelo simulador. Observamos na fig. 2.4 e fig. 2.6 exemplos de trafegabilidades utilizadas no SJD.

Consideramos agora o uso das máscaras como dados de entrada para criar o terreno do SJD. Para isso, estas obtidas após a predição, identificadas como “binários” na figura 4.4, devem ser verificadas dentro do ArcGIS. Esta verificação visa verificar se os pixels possuem informações relativas às cores.

Após este processo, as imagens são convertidas em shapefiles no próprio ArcGIS, sendo inseridas no mesmo código em Python citado nos parágrafos anteriores. A partir deste ponto, o processo de criação do terreno virtual do SJD segue o mesmo trâmite. Como resultado final, observamos a Figura 6.1, contendo a imagem da trafegabilidade a ser utilizada no Simulador.

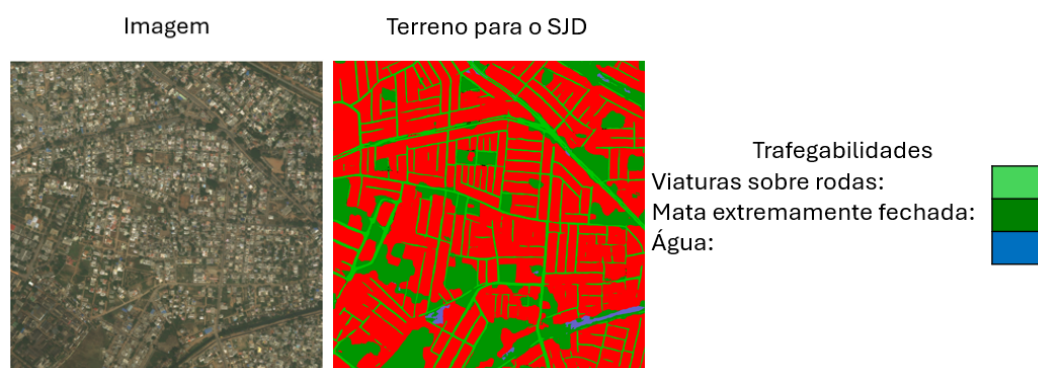


Figura 6.1: A esquerda observamos a imagem satélite original e, a direita, a imagem criada para uso como trafegabilidade no SJD.

Considerando as informações acima citadas, podemos afirmar nossa hipótese básica “O pipeline desenvolvido, que abrange desde a segmentação semântica até a geração de terrenos virtuais para simuladores e jogos de guerra, é capaz de fornecer dados suficientes e condizentes com situações reais, possibilitando a realização de testes efetivos de planejamento militar em ambientes simulados”.

6.2

Uso das imagens concatenadas como Carta de Trafegabilidade

Esta seção visa averiguar se as imagens obtidas fruto dos 3 diferentes tipos de concatenações na seção 4.7 poderiam ser utilizadas como CT. Estas

combinações foram obtidas com a predição realizada de imagens de satélite. Para isso, confeccionamos um questionário contendo as imagens e o enviamos a planejadores militares.

O questionário foi criado no Google Forms [36] e é composto por 3 diferentes seções. Todos os participantes consentiram em responder clicando em uma opção obrigatória antes de entrarem no questionário propriamente dito.

Na primeira seção, foram realizadas perguntas visando obter informações sobre o planejador, tais como sua experiência no planejamento de operações militares. Já na segunda seção, buscamos averiguar como as imagens poderiam contribuir no planejamento de operações através de questionários dissertativos. Finalizando, a terceira seção busca detectar qual seria uma provável concatenação ideal para utilização. Cabe destacar aqui que um questionário piloto foi utilizado e depois descartado.

Os questionários foram enviados a planejadores que servem ou serviram em seções que normalmente lidam com estes tipos de imagens, a saber, as seções de operações e seções de inteligência. As respostas são exibidas nos parágrafos a seguir.

Como respostas relativas à primeira seção de perguntas, nosso público-alvo é composto, em sua maioria, por Capitães de Fragata (Figura 6.2), o que significa que os respondedores possuem experiência nas Forças Armadas.

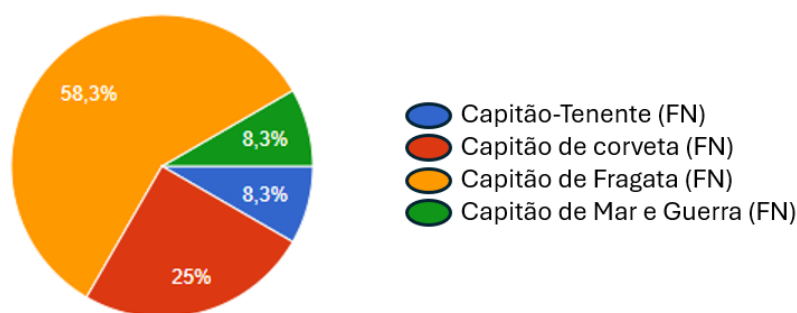


Figura 6.2: Respostas sobre o posto dos planejadores, onde obtemos o tempo de experiência dentro das Forças Armadas.

Além disso, conforme a Figura 6.3, observamos que a maioria dos militares (83.3%) possui acima de 400 dias de manobra e exercício. Estes valores significam que o militar, ao longo de sua carreira, passou a quantidade de dias efetivamente em treinamentos ou em operações militares e longe de casa. Neste sentido, os militares relacionados no planejamento possuem, em sua maioria, experiência em operações e treinamentos militares.

Cerca de quantos dias de manobra e exercício possui?

12 respostas

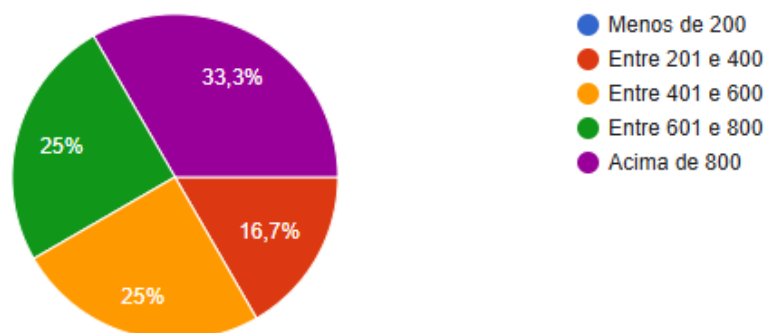


Figura 6.3: Respostas sobre a quantidades de dias de manobra e exercício, que denota experiência em treinamentos e operações militares.

Perguntado ainda quantos planejamentos os militares já participaram e se poderiam descrevê-los, 75% dos participantes já planejaram mais de 5 operações militares (Figura 6.4), denotando experiência, por parte dos planejadores, na realização do PPM. As participações envolvem Operações de Paz, Operações de Garantia da lei e da ordem (GLO), Operações em áreas de desastres naturais e de ajuda humanitária, operações Ágata, Jogos Mundiais da Juventude, operações em áreas ribeirinhas entre diversas outras. Neste contexto, possuímos um vasto leque de operações realizadas.

Já realizou o planejamento de quantas manobras ou exercícios?

12 respostas

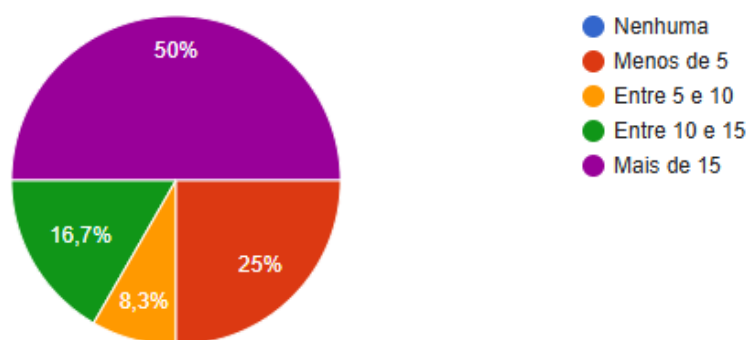


Figura 6.4: Respostas sobre a quantidades planejamentos, que significa que os militares possuem experiência em realizar o PPM.

De maneira geral, podemos observar um público-alvo que contém grande vivência e experiência em operações militares, em que pese a grande quantidade de tipos de operações realizadas e áreas de operações visitadas. Estas

informações atribuem grande credibilidade às respostas dos militares dado ao seu constante trato com planejamentos e documentos militares.

Sobre a segunda seção, que obtém informações sobre a contribuição das imagens no planejamento de operações, enunciaremos suas perguntas e as principais respostas, analisando as respostas obtidas de maneira geral. Destacamos aqui que as perguntas e respostas constam no anexo B.

Sobre a pergunta “Como o Sr utilizaria as informações da Máscara para o planejamento de uma operação militar nesta região”, destacamos as seguintes respostas:

- Usaria as duas para marcar Objetivos, traçar eixos de progressão, limites entre SU e linhas de coordenação.
- Iria sobrepor a imagem satélite a fim de estabelecer itinerários e caso fosse necessário no curso da operações da RAT/RET itinerários e a até para estabelecimento da defensiva.
- Ao indicar as medidas de coordenação e controle do calco de operações e ao mostrar o avanço e desenvolvimento das tropas como um todo. Por essa imagem, me parece adequado para uma valor de batalhão ou superior.
- As informações de máscara são importantes para identificar pontos de importância, trafegabilidade do terreno, as vezes até informação de trânsito pode auxiliar no planejamento.
- Subsídio para trafegabilidade e deslocamento dos meios e tropas; cobertas e abrigos; locais de posicionamento dos apoios; dependendo da situação, setores a serem limpos. Muito próximo a uma Carta de Trafegabilidade.

Referente à pergunta “O Sr poderia citar informações da Máscara que poderiam ser extraídas para planejar operações militares”, destacamos as principais respostas:

- Possíveis obstáculos para o avanço das universidades/SU, dependendo do meio de transporte empregado, vias de acessos mais favoráveis para progressões;
- Obstáculos, que poderiam gerar efeito de canalizar ou dividir as tropas. Uso dos arruamentos para um deslocamento mais rápido, porém aberto a campos de tiro. Terreno com observação extremamente limitada, a partir do solo, face à densidade de construções;
- Informações práticas para o movimento de tropas e meios), ela serve justamente para transformar o mapa “puro” em algo operacional. Direções

de provável desenvolvimento; Obstáculos naturais e artificiais; Trechos impeditivos/restritivos para Vtr sr/sl; possíveis VA; Cobertas e Abrigos; Campos de Tiro; e Rotas Logísticas e de Apoio;

- Principais faixas do terreno a serem utilizadas pela natureza de tropa empregada. Seleção de local adequado para emprego do esforço principal. Seleção de áreas passivas ou com posturas defensivas mais permissivas;
- Máscara com os sentidos das vias urbanas; com o fluxo dos rios e as lagoas existentes no terreno;
- Restrições/impeditivos/facilitadores ao deslocamento dos meios e tropa, tais como casarios (área urbana), locais alagadiços, corte de estradas, mata densa;

Sobre a pergunta “comparando com uma carta de trafegabilidade convencional, o Sr poderia citar vantagens para o uso da Máscara como uma carta de trafegabilidade”, destacamos agora as respostas abaixo:

- Essas cartas possuem um detalhamento maior, com dimensões com proporcionalidade mais próximas ao real;
- A carta convencional é rica em detalhes brutos, mas exige interpretação técnica; já a máscara entrega uma síntese “pronta” para a decisão militar, acelerando o planejamento e reduzindo falhas.
- Possibilidade de seleção de filtros e camadas de forma mais fácil para focar em uma atividade ou natureza de tropa específica;
- Facilita a sobreposição de imagens e consegue-se tomar decisões mais rápidas
- Creio que a atualização dos dados e possibilidade de escolha das informações a serem avaliadas são vantagens; e
- Ainda não percebi as possíveis vantagens.
- Facilidade de identificar itinerários.

Finalizando, as principais respostas da pergunta “comparando com uma carta de trafegabilidade convencional, o Sr poderia citar desvantagens para o uso da Máscara como uma carta de trafegabilidade” são exibidas abaixo:

- A carta da esquerda, talvez, possa ser um pouco mais difícil de se orientar em certos casos.
- Os obstáculos nem sempre aparecem de modo claro.

- A máscara é excelente para síntese rápida e visão operacional, mas perde em profundidade, precisão e flexibilidade de interpretação. Na prática, o ideal é usar as duas em conjunto: a máscara para a decisão rápida e briefing, e a carta convencional para o planejamento detalhado e navegação.
- A ausência da imagem satélite ou outros elementos que são retirados para clarear a imagem reforçam a dependência da análise do software ou da Inteligência Artificial aplicada, impedindo uma avaliação própria do planejador, fato que passa a ser relevante quando o planejador é um oficial mais antigo. A Máscara retira o fator experiência do processo de planejamento.
- Ainda não percebi desvantagem.
- A princípio, não vejo uma desvantagem.
- Não apresenta curvas de nível, para avaliação da altitude dos pontos notáveis.

De maneira geral, obtivemos respostas bastante positivas, onde podemos destacar a identificação possível das trafegabilidades e a sua possível utilização em conjunto com a carta militar, em complemento a esta. Cabe destacar aqui que a carta de trafegabilidade utilizada no PPM é utilizada em conjunto com as cartas militares de forma a complementá-las.

Além disso, destacamos ainda a possibilidade de “entregar uma síntese pronta para a decisão militar, acelerando o planejamento e reduzindo falhas” destacado anteriormente. Combinando esta resposta com outras como a possibilidade de “transformar o mapa “puro” em algo operacional”, podemos afirmar a nossa hipótese básica, que versa que “Os produtos gerados pela segmentação semântica – particularmente as classes identificadas – podem ser utilizados para criar cartas de trafegabilidade que atendem aos requisitos operacionais das Operações militares, acelerando o processo de análise e planejamento do terreno”.

7

Conclusões

O presente trabalho tem como objetivo melhorar a análise detalhada do terreno, a realização de treinamentos e testes de planos estratégicos em contextos reais e simulados. Para isso, desenvolvemos e validamos um pipeline baseado em técnicas modernas de deep learning para segmentação semântica de imagens de sensoriamento remoto. O propósito foi automatizar e aprimorar o processo de geração de cartas de trafegabilidade, terrenos virtuais e ambientes simulados para apoio ao planejamento de operações militares. Para atingir esse objetivo, abordamos o problema proposto em 6 fases.

Para isso apresentamos a seguinte questão: “Como desenvolver uma abordagem que forneça melhores condições de planejamento, com recursos atualizados e de maneira rápida, capaz de auxiliar no entendimento do estudo dos aspectos topológicos e táticos do terreno, bem como auxiliar na realização dos confrontos deste plano?”

Durante as fases, foram levantados os requisitos necessários que orientassem a como confeccionar uma CT. Além disso, foi realizada pesquisa visando à obtenção das melhores arquiteturas, modelos de redes neurais já em utilização no mercado e datasets.

Posteriormente, foram realizados treinamentos e predições com arquiteturas de redes neurais encontradas e com os diversos datasets obtidos, de forma a obtermos a melhor combinação de parâmetros. Foram realizadas predições em imagens do dataset para ser possível a criação da CT, do terreno para o SJD, do terreno virtual 3D e um visualizador para observá-lo, chamado de Simulador Virtual para Estudo Topotático do Terreno (SVETT). Finalizando, foi conduzida uma pesquisa sobre o uso do SVETT como ferramenta de análise do terreno.

Dentre as diferentes arquiteturas e modelos estudados e testados para o treinamento, a melhor solução encontrada foi através da abordagem de Comitê de modelos. Nele, cada classe foi treinada individualmente, e as decisões para a segmentação são tomadas em decisões coletivas [73]. Já os datasets Land Cover Classification Challenge e Road Extraction Challenge, ambos do DeepGlobe 2018 Satellite Image Understanding Challenge [43] foram utilizados como dados para nossa solução proposta.

De posse das máscaras obtidas fruto da predição do modelo, o processo de criação do terreno virtual pôde ser realizado. Para a inserção das camadas, utilizamos o asset Vegetation studio Pro [109], em que pese o fato deste último

não estar mais disponível.

Além disso, desenvolvemos o SVETT e o avaliamos com 54 alunos. A intenção foi obter informações de forma a comprovar que, usando o SVETT entre a teoria e a ida ao ET, os alunos o acharão melhor quando participarem do Exercício devido ao seu primeiro contato com a região ainda em sala de aula.

Sobre a criação do terreno para o SJD, utilizamos as máscaras como dados de entrada para o ArcGIS. Estas imagens, chamadas de “binários” neste trabalho, foram verificadas de forma a averiguar se os pixels possuíam informações relativas às cores. Após este processo, as imagens são convertidas em shapefiles no próprio ArcGIS, sendo inseridas no mesmo código em Python citado nos parágrafos anteriores. A partir deste ponto, o processo de criação do terreno virtual do SJD pôde seguir com a criação do terreno.

Finalizando, criamos cartas de trafegabilidades a partir de 3 diferentes tipos de concatenações, elencados na seção 4.7. Estas concatenações foram obtidas com a predição realizada de imagens de satélite. Além disso, buscamos averiguar como as 3 imagens produzidas poderiam contribuir no planejamento de operações militares, por meio de um questionário enviado aos planejadores militares por meio do Google Forms.

Referente aos resultados obtidos, observamos na tabela 4.3 que os mIoU obtidos pelo modelo s4GAN + Entropy foram superados em sua maioria pelas predições das classes do experimento 4 (Tabela 4.2). Cabe destacar que, durante as procuras por trabalhos que utilizassem o mesmo dataset, o trabalho de Desai e Ghose foi o mais atual encontrado. Sendo assim, fomos capazes de afirmar nossa hipótese básica que diz que “O método de Comitê de modelos (do inglês, “Ensemble methodology”), para segmentação semântica binária, apresenta desempenho superior à abordagem multiclases tradicional, permitindo identificar com maior precisão as classes relacionadas à trafegabilidade do terreno em imagens de sensoriamento remoto”.

Através de questionário aplicado a planejadores militares, obtivemos resultados que afirmam que a máscara obtida “são importantes para identificar pontos de importância, trafegabilidade do terreno”. Desta forma, afirmamos a hipótese básica que diz que “Os produtos gerados pela segmentação semântica – particularmente as classes identificadas – podem ser utilizados para criar cartas de trafegabilidade que atendem aos requisitos operacionais das Operações militares, acelerando o processo de análise e planejamento do terreno”.

Referente ao SVETT, os usuários se sentiram mais seguros quando receberam a instrução em campo depois de terem recebido a instrução por meio do simulador (nota 4,69 de 5,00). Desta forma, o conceito de instrução

que consistiu no uso do simulador seguido de exercício em campo, na mesma região do simulador, facilitou o aprendizado (nota 4,66 de 5,00, com o segundo menor desvio padrão, 0,47).

O fato acima nos ajudou a afirmar as duas hipóteses básicas:

- “A integração de ambientes virtuais para estudo topotático do terreno, combinando recursos didáticos tradicionais com simulação virtual, potencializa o aprendizado e a confiança dos usuários (alunos e planejadores), sendo eficiente no treinamento antes do deslocamento ao campo real para análise de missão”; e
- “O pipeline desenvolvido, que abrange desde a segmentação semântica até a geração de terrenos virtuais para simuladores e jogos de guerra, é capaz de fornecer dados suficientes e condizentes com situações reais, possibilitando a realização de testes efetivos de planejamento militar em ambientes simulados”.

Fato importante observado durante o trabalho foi a escassez, ou até inexistência, de um dataset com as diferentes regiões do Brasil, contendo terrenos anotados, não sendo possível treinar o modelo nessas condições.

Outro aspecto detectado está na carência ainda de um recurso para inserir as residências em suas posições exatas, haja vista a identificação ser apenas da área contendo construções. Estudos futuros visando a identificação positiva das construções podem ser realizados de forma a complementar este trabalho.

Tomando como base os estudos realizados durante todo o período, foram publicados, ao todo, 3 artigos. O primeiro foi publicado na Interservice / Industry Training Simulation and Education Conference (I/ITSEC), realizada em Orlando, Florida, com o título de “From Classroom to Field: Topological and Tactical Terrain Analysis Inside a Learning Environment”.

Já o segundo e terceiro artigos foram publicados no XXVI Symposium on Virtual and Augmented Reality (SVR 2024), realizado em Manaus, AM. O primeiro artigo deste Simpósio, “Using learning environments for military study”, foi publicado na trilha principal, e o segundo, intitulado “A simulator for military terrain study”, foi publicado na categoria “XR Experience”, onde ganhou o prêmio “Best demo”.

Como trabalhos futuros, sugerimos a criação de um dataset com o foco nas diferentes regiões do Brasil, de forma a permitir a elaboração dos materiais propostos neste trabalho em melhores condições. Além disso, fruto do constante emprego das Forças Armadas em resposta às calamidades englobando desastres naturais, este trabalho sugere que o tema seja incluído na elaboração dos datasets.

Após a criação do terreno, utilizamos o recurso do Vegetation Studio Pro para inserir a vegetação no terreno. Este processo carece ainda de um recurso para inserir as residências em suas posições exatas, haja vista identificarmos apenas a área contendo construções. Estudos futuros visando à identificação positiva das construções podem ser realizados de forma a complementar este trabalho.

Referências bibliográficas

- [1] NOWACKA, A.; RZEMIENIAK, M.. The impact of the vuca environment on the digital competences of managers in the power industry. *Energies*, 15(1):185, 2021. 1
- [2] SOEMITRA, A.; LUBIS, A. S.; DEWI, R. S. ; OVAMI, D. C.. Essential hard skill for student in vuca era: Literature study. *Journal of Trends Economics and Accounting Research*, 3(4):605–610, 2023. 1
- [3] ABIDI, S.; JOSHI, M.. *The Vuca Learner: Future Proof Your Relevance*. Sage Publications, 2018. 1
- [4] LICHTMAN, S.. Office of naval research: Preparing the marine corps for battlefields of the future. *LEATHERNECK*, p. 7, 2023. 1
- [5] NAVY, D. O. T.; MARINES CORPS, U. S.. *Force design 2030 annual update*. . 1
- [6] SANDERFIELD, C. D.. The arms room concept: How multi-disciplinary marines provide maximum flexibility to a distributed force by capt devon sanderfield. *Marine Corps Gazette*, p. 4, 2023. 1
- [7] ARMADA, E.-M. D.. *Doutrina militar naval*. . 1, 2.1
- [8] FEDERAL, G.. *Livro branco de defesa nacional*. (document), 1, 1, 2.1, 2.2
- [9] BRASIL, B.. *Política nacional de defesa e estratégia nacional de defesa*. 1, 2.1.1
- [10] ARMADA, E.-M. D.. *Manual de planejamento operativo da marinha*. . 1, 2.2
- [11] NAVAIS, C.-G. C. F.. *Manual de operações terrestres de fuzileiros navais*. . 1, 2.1.1
- [12] ALMEIDA, R. D. S. E.; GONÇALVES, T. ; RAPOSO, A.. Using learning environments for military study. In: *SYMPOSIUM ON VIRTUAL AND AUGMENTED REALITY*, p. 185–192. ACM, 2024. 1, 2.3, 4.7

- [13] THIJSSSEN, D.; BOSMA, R.. Recommendation system in an integrated digital training environment for the 5th generation air force. In: THE INTERSERVICE INDUSTRY, TRAINING, SIMULATION, AND EDUCATION CONFERENCE SEARCH ARCHIVES, p. 11. NTSA, 2022. 2.3
- [14] NAVAIS, C.-G. C. F.. Manual de planejamento dos grupamentos operativos de fuzileiros navais. . 1, 1, 2.1.1, 2.2, 2.2, 5
- [15] L. C. CASTRO, A.; B. CALHEIROS, L.; I. R. CUNHA, M. ; L. N. C. BRINGEL, M.. Manual de desastres volume i, desastres naturais. 1
- [16] NAVAL, E. D. G.. Jogos de guerra. . 1
- [17] FANG, L.; YANG, Z.; MA, T.; YUE, J.; XIE, W.; GHAMISI, P. ; LI, J.. Open-world recognition in remote sensing: Concepts, challenges, and opportunities. IEEE Geosci. Remote Sens. Mag., 12(2):8–31, 2024. 1
- [18] CASAGLI, N.; INTRIERI, E.; TOFANI, V.; GIGLI, G. ; RASPINI, F.. Landslide detection, monitoring and prediction with remote-sensing techniques. Nature Reviews Earth & Environment, 4(1):51–64, 2023. 1, 2.5
- [19] One world terrain - university of southern california institute of creative technologies, 2025. 1
- [20] ARBIA, B.; OUERSIGHNI, R.. Simulation systems for armed forces training simulation system nato’s standards and proposal of environment model. In: ICCRK, p. 2. Conal, 2012. Publisher: Unpublished. 1, 1.3, 5
- [21] BRUMADINHO, P.. Histórico do rompimento das barragens da vale na mina córrego do feijão, May 2024. 1.1
- [22] MOREIRA, R. E. D. A.. Deslocamentos induzidos por desastres relacionados à chuva no brasil entre 2013 e 2022. REMHU: Revista Interdisciplinar da Mobilidade Humana, 32:18, 2024. 1.2
- [23] APUFSC, I.. Entidades alertam para possível uso indevido de dados pessoais pelo aplicativo sou gov.br, July 2021. 1.2
- [24] , K.. Kaggle, June 2025. 1.4

- [25] NAVAIS, C.-G. C. F.. Cgcfn 1-1 - manual de operações da força de desembarque. . 2
- [26] RIO, G.. Forças armadas assumem ocupação de 15 comunidades da maré, rio. 2.1
- [27] RIO, U.. Marinha faz treinamento para Copa das Confederações e Jornada Mundial da Juventude no Rio, May 2013. 2.1
- [28] DEFESA, M. D.. Histórico da participação brasileira em missões da ONU, June 2025. 2.1
- [29] BRASIL, N. U.. Brasileiros das forças de paz concluem missão no Líbano após uma década de participação, Dec. 2020. 2.1
- [30] MENDONÇA, V.. Marinha mobiliza grupamento de Fuzileiros Navais para apoiar vítimas de temporal no RJ, Feb. 2024. 2.1
- [31] NAVAIS, C.-G. C. F.. Manual do combatente anfíbio. . 2.1.1, 3
- [32] NAVAIS, C.-G. C. F.. Manual de inteligência de fuzileiros navais. . 2.2
- [33] PEREIRA, R.. O que é a amazônia azul e por que o brasil quer se tornar potência militar no atlântico, Nov. 2019. 2.2.1
- [34] DE SOUZA E ALMEIDA, R.. Topological and tactical study modeling into a virtual environment. In: ITEC EXTENDED ABSTRACT, p. 8, 2019. 2.2.1
- [35] Passagens de ônibus de rio de janeiro, rj aracaju, se. 4
- [36] WORKSPACE, G.. Formulários on-line para receber insights rapidamente, May 2024. 2.3, 6.2
- [37] AHMED, A.; SUTTON, M. J.. Gamification, serious games, simulations, and immersive learning environments in knowledge management initiatives. World Journal of Science, Technology and Sustainable Development, 14(2):78–83, 2017. 2.4
- [38] Ören, T.; Zeigler, B. P. ; Tolk, A., editors. Body of Knowledge for Modeling and Simulation: A Handbook by the Society for Modeling and Simulation International. Simulation Foundations, Methods and Applications. Springer International Publishing, Cham, 2023. 2.4

- [39] LIANG, Y.; WANG, X.; YUE, T.; XU, X.; YAO, C.; LIU, D. ; LIU, Q.. **Review of live-virtual-constructive simulation technology**. In: JOURNAL OF Physics: Conference Series, 2022. 2.4
- [40] JASPERS, F.. **VR Team Training for military special forces**. In: PAPER Sessions I/ITSEC 2024 Conference, p. 9, Orlando, Florida, 2024. NTSA. 2.4
- [41] MIRANDA, J.; CUMMINS, C.. **Adapting**. Marine Corps Gazette, 1(April 2023):5, Apr. 2023. 2.4
- [42] PERLA, P.. **Wargaming and the cycle of research and learning**. Scandinavian Journal of Military Studies, 5(1):197–208, 2022. 2.4
- [43] DEMIR, I.; KOPERSKI, K.; LINDENBAUM, D.; PANG, G.; HUANG, J.; BASU, S.; HUGHES, F.; TUIA, D. ; RASKAR, R.. **Deepglobe 2018: A challenge to parse the earth through satellite images**. In: 2018 IEEE/CVF CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION WORKSHOPS (CVPRW), p. 172–17209. IEEE, 2018. 2.5, 3, 4.1, 4.3, 4.4, 7
- [44] SENSING, C. C. F. R.. **Fundamentals of remote sensing**. . 2.5
- [45] DI MAIO, A.; F. T. RUDORFF, B.; C. MORAES, E.; PEREIRA, G.; A. MOREIRA, M.; M. SAUSEN, T. ; G. FLORENZANO, T.. **Curso astronáutica e ciências do espaço, Sensoriamento remoto**. AEB Escola, 2008. 2.5
- [46] SKANSI, S.. **Introduction to Deep Learning: From Logical Calculus to Artificial Intelligence**. Undergraduate Topics in Computer Science. Springer International Publishing, 2018. 2.6.1
- [47] TURING, A. M.. **Computing machinery and intelligence**. 2.6.1
- [48] MCCARTHY, J.; MINSKY, M. L.; ROCHESTER, N. ; SHANNON, C. E.. **A proposal for the dartmouth summer research project on artificial intelligence**. AI Magazine, 27:3, 2006. 2.6.1
- [49] SAS, S.. **História das redes neurais**, 2023. 2.6.1
- [50] FAVARON, G.. **Os invernos da ia: Ciclos de ascensão e queda na história da inteligência artificial**, July 2024. 2.6.1
- [51] HAYKIN, S. S.. **Neural networks and learning machines**. Prentice-Hall, 3. ed edition, 2009. (document), 2.6.2, 2.6.2, 2.10, 2.6.2

- [52] SANTOS, V. S. D.. **Neurônios**, 2025. 2.6.2
- [54] CHOLLET, F.. **Francois Chollet - Deep Learning with Python**. Manning Publications Co., 2018. (document), 2.6.2, 2.6.3, 2.11
- [55] KRIZHEVSKY, A.; SUTSKEVER, I. ; HINTON, G. E.. **Imagenet classification with deep convolutional neural networks**. Communications of the ACM, 60(6):84–90, 2017. Publisher: Association for Computing Machinery (ACM). 2.6.3
- [56] LUNA, J. C.. **Um guia para iniciantes sobre a unidade linear retificada (ReLU)**, Jan. 2025. 2.6.3
- [57] NIMMA, D.; UDDAGIRI, A.. **Advancements in deep learning architectures for image recognition and semantic segmentation**. International Journal of Advanced Computer Science and Applications, 15(8), 2024. Publisher: The Science and Information Organization. 2.6.3
- [58] LI, Z.; LIU, F.; YANG, W.; PENG, S. ; ZHOU, J.. **A survey of convolutional neural networks: Analysis, applications, and prospects**. IEEE Transactions on Neural Networks and Learning Systems, 33(12):6999–7019, Dec. 2022. Publisher: Institute of Electrical and Electronics Engineers (IEEE). 2.6.3, 2.6.3
- [59] ZHU, E.; CHEN, Z.; WANG, D.; SHI, H.; LIU, X. ; WANG, L.. **Unet-mamba: An efficient unet-like mamba for semantic segmentation of high-resolution remote sensing images**. IEEE Geosci. Remote Sensing Lett., 22:1–5, 2025. 2.6.3
- [60] RONNEBERGER, O.; FISCHER, P. ; BROX, T.. **U-net: Convolutional networks for biomedical image segmentation**. (document), 2.12, 2.6.3, 2.7.1
- [61] HE, K.; ZHANG, X.; REN, S. ; SUN, J.. **Deep residual learning for image recognition**. 2.6.3
- [62] ZHANG, L.; ZHANG, L. ; DU, B.. **Deep learning for remote sensing data: A technical tutorial on the state of the art**. IEEE Geoscience and Remote Sensing Magazine, 4(2):22–40, June 2016. 2.7
- [63] WILEY, J.. **Scale Issues in Remote Sensing**. John Wiley & Sons, 2014. 2.7

- [64] ZHOU, N.; HONG, J.; CUI, W.; WU, S. ; ZHANG, Z.. **A multiscale attention segment network-based semantic segmentation model for landslide remote sensing images.** *Remote Sensing*, 16(10):1712, 2024. 2.7
- [65] ZHENG, X.; CHEN, T.. **High spatial resolution remote sensing image segmentation based on the multiclassification model and the binary classification model.** *Neural Computing and Applications*, 35(5):3597–3604, 2023. 3
- [66] TONG, X.-Y.; XIA, G.-S.; LU, Q.; SHEN, H.; LI, S.; YOU, S. ; ZHANG, L.. **Land-cover classification with high-resolution remote sensing images using transferable deep models.** *Remote Sensing of Environment*, 237:111322, 2020. 3
- [67] ABURAED, N.; AL-SAAD, M.; ALKHATIB, M. Q.; ZITOUNI, M. S.; ALMANSOORI, S. ; AL-AHMAD, H.. **Semantic segmentation of remote sensing imagery using an enhanced encoder-decoder architecture.** *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, X-1/W1-2023:1015–1020, 2023. 3
- [68] CASTILLO-NAVARRO, J.; AUDEBERT, N.; BOULCH, A.; LE SAUX, B. ; LEFEVRE, S.. **What data are needed for semantic segmentation in earth observation?** In: 2019 JOINT URBAN REMOTE SENSING EVENT (JURSE), p. 1–4. IEEE, 2019. 3
- [69] AUDEBERT, N.; SAUX, B. L. ; LEFÈVRE, S.. **Beyond RGB: Very high resolution urban remote sensing with multimodal deep networks.** *ISPRS Journal of Photogrammetry and Remote Sensing*, 140:20–32, 2018. 3
- [70] SIMONYAN, K.; ZISSERMAN, A.. **Very deep convolutional networks for large-scale image recognition.** 3
- [71] ROTTENSTEINER, F.; SOHN, G.; JUNG, J.; GERKE, M.; BAILLARD, C.; BENITEZ, S. ; BREITKOPF, U.. **The isprs benchmark on urban object classification and 3d building reconstruction.** *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, I-3:293–298, 2012. Publisher: Copernicus GmbH. 3
- [72] HE, K.; ZHANG, X.; REN, S. ; SUN, J.. **Delving deep into rectifiers: Surpassing human-level performance on imagenet classification.**

- In: 2015 IEEE INTERNATIONAL CONFERENCE ON COMPUTER VISION (ICCV). IEEE, 2015. 3
- [73] KUMAR, S.; KAUR, P. ; GOSAIN, A.. **A comprehensive survey on ensemble methods**. In: 2022 IEEE 7TH INTERNATIONAL CONFERENCE FOR CONVERGENCE IN TECHNOLOGY (I2CT), p. 1–7. IEEE, 2022. (document), 3, 3.1, 4.5, 7
- [74] ALIZADEH, H.; YOUSEFNEZHAD, M. ; BIDGOLI, B. M.. **Wisdom of crowds cluster ensemble**. IDA, 19(3):485–503, 2015. 3
- [75] MISHRA. **Ensemble learning: A beginner’s guide**. 3
- [76] TAYLOR, N. E.; CLAYTON, A. S.. **Form from function: Applying flow driven experiential learning to the integration of immersive technology in formal military aviation training programs**. In: THE INTERSERVICE INDUSTRY, TRAINING, SIMULATION, AND EDUCATION CONFERENCE SEARCH ARCHIVES, p. 13. NTSA, 2021. 3, 3
- [77] EHRLICH P.D., D. B.. **Establishing connections: Interactivity factors for a distance education course**. Journal of Educational Technology & Society, 5(1):48–54, 2002. 3
- [78] KOZHEVNIKOV, M.; KOZHEVNIKOV, M.. **The effect of environment immersivity on spatial perspective-taking task performance**. In: I/ITSEC EXTENDED ABSTRACT, p. 12. NTSA, 2022. 3
- [79] ANACONDA. **Anaconda.org**. 4.5
- [80] MICROSOFT. **Visual studio code**. 4.5
- [81] GUO, Z.; SHI, X.; ZHANG, H.; HUANG, D.; SONG, X.; YAN, J. ; SHIBASAKI, R.. **Enhancing building semantic segmentation accuracy with super resolution and deep learning: Investigating the impact of spatial resolution on various datasets**. 4.1
- [82] BUSLAEV, A.; SEFERBEKOV, S.; IGLOVIKOV, V. ; SHVETS, A.. **Fully convolutional network for automatic road extraction from satellite imagery**. In: 2018 IEEE/CVF CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION WORKSHOPS (CVPRW), p. 197–1973. IEEE, 2018. 4.1
- [83] GOOGLE. **Google colab**. 4.2, 4.3
- [84] AWARI. **Tensorflow vs. pytorch: Comparando frameworks**. 4.2

- [85] SHARMA, P.. **A step-by-step guide to image segmentation techniques (part 1)**. 2.7.1
- [86] DE OCAMPO, A. L. P.. **Normalized difference vegetation index (ndvi) estimation based on filter augmented imaging**. In: 2023 INTERNATIONAL ELECTRICAL ENGINEERING CONGRESS (IEEECON), p. 84–88. IEEE, 2023. 4.2
- [87] PEREIRA, L. E.; AMORIM, G.; GRIGIO, A. M. ; PARANHOS, A. C.. **Comparative analysis of normalized difference water index (ndwi) methods in continental wetland**. Anuário do Instituto de Geociências - UFRJ, 41(2):654–662, 2018. 4.2
- [88] SEFERBEKOV, S.; IGLOVIKOV, V.; BUSLAEV, A. ; SHVETS, A.. **Feature pyramid network for multi-class land segmentation**. In: 2018 IEEE/CVF CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION WORKSHOPS (CVPRW), p. 272–2723. IEEE, 2018. 4.2
- [89] DESAI, S.; GHOSE, D.. **Active learning for improved semi-supervised semantic segmentation in satellite images**. In: 2022 IEEE/CVF WINTER CONFERENCE ON APPLICATIONS OF COMPUTER VISION (WACV), p. 1485–1495. IEEE, 2022. 3, 4.7
- [90] SHAO, Z.; ZHOU, W.; DENG, X.; ZHANG, M. ; CHENG, Q.. **Multilabel remote sensing image retrieval based on fully convolutional network**. IEEE J. Sel. Top. Appl. Earth Observations Remote Sensing, 13:318–328, 2020. 3
- [91] MACKOWIAK, R.; LENZ, P.; GHORI, O.; DIEGO, F.; LANGE, O. ; ROTHER, C.. **Cereals - cost-effective region-based active learning for semantic segmentation**, 2018. 3
- [92] MITTAL, S.; TATARCHENKO, M. ; BROX, T.. **Semi-supervised semantic segmentation with high- and low-level consistency**. IEEE Trans. Pattern Anal. Mach. Intell., 43(4):1369–1379, 2021. 3.4
- [93] TIAN, C.; LI, C. ; SHI, J.. **Dense fusion classmate network for land cover classification**. In: 2018 IEEE/CVF CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION WORKSHOPS (CVPRW), p. 262–2624. IEEE, 2018. 3
- [94] JIA, Y.; SHELHAMER, E.; DONAHUE, J.; KARAYEV, S.; LONG, J.; GIRSHICK, R.; GUADARRAMA, S. ; DARRELL, T.. **Caffe: Convolutional architecture for fast feature embedding**, 2014. 3

- [95] RAHIMI, F.; SADEGHI-NIARAKI, A. ; CHOI, S.-M.. **Generative ai meets virtual reality: A comprehensive survey on applications, challenges, and future direction.** IEEE Access, 13:94893–94909, 2025. 3
- [96] BULKAS, D.; CHIAMPI, J.; CHAPMAN, J. ; PAVILL, V. F.. **Creating a virtual reality environment with a fusion of suas and tls point-clouds.** International Journal of Image and Data Fusion, 11(2):136–161, 2020. 3
- [97] **inceptionresnetv2**, 2025. 4.3
- [98] 365 TEAM. **What is cross-entropy loss function?** 4.3
- [99] DEMIR, A.; YILMAZ, F.. **Inception-ResNet-v2 with leakyrelu and averagepooling for more reliable and accurate classification of chest x-ray images.** In: 2020 MEDICAL TECHNOLOGIES CONGRESS (TIPTEKNO), p. 1–4. IEEE, 2020. 4.3
- [100] KINGMA, D. P.; BA, J.. **Adam: A method for stochastic optimization.** In: INTERNATIONAL CONFERENCE FOR LEARNING REPRESENTATIONS, p. 15, 2015. 4.3
- [101] PACIOCCO, J.; HASAN, A.; CHAN, J.; TANIGUCHI, D.; AMON, C.; KAROUBI, G. ; BAZYLAK, A.. **Computer vision to predict cell seeding coverage in re-endothelialized mouse lungs.** Scientific Reports, 15(1):26236, 2025. 4.4
- [102] REZATOFIGHI, H.; TSOI, N.; GWAK, J.; SADEGHIAN, A.; REID, I. ; SAVARESE, S.. **Generalized intersection over union: A metric and a loss for bounding box regression.** In: 2019 IEEE/CVF CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION (CVPR), p. 658–666. IEEE, 2019. 4.4
- [103] SIGOULAKIS, D.. **Ajp-01_edf_v1_e__(1)_2437**, 2022. 5
- [104] UNITY TECHNOLOGIES. **Unity**. 5
- [105] BLENDER. **The freedom to create**, 2025. . 5
- [106] MIRROR. **Mirror**, 2025. 5
- [107] POSTGRESQL GLOBAL DEVELOPMENT GROUP. **Postgresql: The world’s most advanced open source relational database**, 2025. 5

- [108] CODE, I.. **Infinity code**, 2025. 5
- [109] **Vegetation studio pro extensions**, 2025. 5, 7
- [110] **Survey monkey**, 2025. 5.2
- [111] ALMEIDA, R. G. D. S.; MAZZO, A.; MARTINS, J. C. A.; BAPTISTA, R. C. N.; GIRÃO, F. B. ; MENDES, I. A. C.. **Validation to portuguese of the scale of student satisfaction and self-confidence in learning**. *Revista Latino-Americana de Enfermagem*, 23(6):1007–1013, 2015. 5.2
- [112] **Where can i find 3d elevation program (3dep) product updates and status maps?**, July 2024. 6.1
- [113] GOOGLE. **Google earth**, May 2024. 6.1
- [114] OPEN STREET MAPS. **Open Street Maps**, 2025. 6.1
- [115] ESRI, E.. **Arcgis online**, Feb. 2025. 6.1
- [116] , U.. **United states geological survey**, 2025. 4.2, 6.1

A

Respostas do questionário qualitativo e quantitativo aplicado aos instrutores antigos e atuais do C-ApA-CFN

Neste anexo, cada título de seção é composto pela pergunta realizada pelos instrutores. Cada item relacionado na seção é a resposta referente à pergunta realizada.

Em sua opinião, quais os pontos positivos em se realizar a análise do terreno no Exercício no Terreno (ET)?

- Padroniza os termos militares;
- Entender as diversas formas de tatear o terreno e utiliza-las para o melhor entendimento do seu público alvo (pode ser um Pelotão, ou um Recon de Líderes, ou uma Equipe ReconVig, entre outros);
- Uniformizar o conhecimento do terreno;
- Comparar o que vemos e estudamos nas cartas com o terreno real e que, esse mesmo terreno, pode ter evoluído desde a época de confecção da carta;
- Após realizar o planejamento como oficial de Estado-Maior em sala de aula, restam dúvidas que são sanadas no terreno;
- Adaptação do planejamento em sala de aula ao terreno;
- Identificação de dissonâncias entre as cartas, imagens satélites e demais meios de apoio com o terreno;
- Desenvolvimento da capacidade analítica dos OA;
- Os ET Provém ao aluno o conhecimento e real noção do planejamento. Podendo verificar se determinados ações, distâncias envolvidas, natureza da tropa são exequíveis no terreno;
- Oportunidade dos alunos verificarem a aplicabilidade prática dos planejamentos realizados em sala de aula, confrontando com a realidade do terreno;
- Oportunidade dos alunos aprimorarem os conhecimentos teóricos a luz da realidade que o terreno impõe;
- Verificação concreta de aspectos do planejamento, dificilmente “mensuráveis” sem a ida ao terreno (Ex. Fatores fixos, como granulometria, consistência da praia, regime de cheias e secas das hidrovias, etc);
- Somente o terreno fornece a real dimensão ao aluno acerca da necessidade de adaptabilidade do que é planejado em sala de aula com a realidade e, portanto, o entendimento de que o papel, de fato, não deve aceitar tudo;

- Traz a realidade do terreno para os alunos;
- O ponto mais positivo é levar o aluno a confirmar / retificar os aspectos táticos do terreno, após os levantar em sala de aula;
- O estudo do terreno no ET é uma ferramenta muito importante para mostrar aos OA se o que está sendo planejado pode ser levado em consideração ou necessitará de modificações. Isso ocorre porque durante o planejamento tudo é aceito no papel, mas quando se vai ao terreno pode-se verificar se é possível, pois pode-se achar um terreno desfavorável para algum tipo de progressão de viatura, por exemplo;
- Permite que o OA aplique seus conhecimentos aprendidos em sala no terreno, de forma a saber efetivamente interpretar a carta com o terreno, onde efetivamente será realizada a operação planejada;
- Poder dar ao Oficial Aluno uma perspectiva real do que é planejado na carta ou em representações;
- Permite uma melhor análise do terreno, auxiliando na tomada de decisão;
- Estimula o OA a aplicar os conceitos doutrinários de forma prática;
- Permite o OA a adaptar o seu planejamento às condições reais de emprego;
- Permitir que os OA entendam as influências da AOp em sua manobra;
- Permite ao Aluno aplicar os conceitos teóricos na visualização do terreno, incluindo as possíveis alterações das condições topográficas em comparação com o que é representado tanto na carta quanto em imagens de Google Earth bidimensionais;
- Outro fator importante que influencia neste processo são as condições atmosféricas/climatológicas que podem afetar a luminosidade / alcance de visada / trafegabilidade dos diferentes tipos de terreno;
- Aplicar os conhecimentos de sala de aula com a realidade do terreno;
- Aplicação de planos;
- Dimensionamento das frentes ofensivas e defensivas;
- Conhecimento aprofundado de obstáculos, objetivos e ZReu;
- Reconhecimento de bons/maus PO; e
- Permite ao oficial visualizar os pontos importantes que foram utilizados na manobra, e assim RET/RAT posteriormente.

Em sua opinião, durante o exercício, quais as dificuldades em se realizar as instruções nos ET?

- Fazer com que os alunos observem o que está na carta com o que está no terreno. Principalmente quando o terreno está diferente da carta por diversos motivos evolutivos, como construção de casarios e outras obras de arte que acabam confundindo curvas de nível ou diferenças entre elevações;
- Normalmente, as grandes distâncias envolvidas;
- Questões logísticas e coordenação com os responsáveis pelos locais planejados como PO;
- Administrativo: Hoje um dos principais problemas para qualquer tipo de exercício é o CLG. O C-ApA-CFN não apresentou esse tipo de problema mas a inserção de qualquer outra necessidade já não seria atendida;
- Operativo: a subjetividade do entendimento. Algumas ações são de caráter subjetivo o que é natural nas atividades militares. A dificuldade é transmitir o pensamento ideal de modo que esteja alinhado com a doutrina;
- Dificuldades inerentes a aspectos de ordem logística (escassez de pontos de apoio com estrutura adequada a execução dos exercícios, limitações de meios, etc);
- Poucas oportunidades de ida ao terreno pelos oficiais que não sejam da infantaria;
- A parte mais difícil é analisar um terreno heterogêneo, com diversas características;
- A dificuldade é mais administrativa. Arrumar viatura é conciliar datas com o calendário do curso. E até mesmo verba para que o exercício ocorra;
- Logísticas, principalmente transporte;
- A carta estar desatualizada por alterações recentes ocorridas no terreno;
- Problemas meteorológicos, como chuva;
- Dificuldade em encontrar o acesso ao PO planejado;
- Condições climáticas e logística;
- Restrições orçamentárias para aquisição de passagens;
- Logística;
- Normalmente são de caráter logístico (numero de vtr), condições de trafegabilidade até os locais de PO, e em alguns casos acesso negado aos mesmos PO ou outros locais que poderiam ser usados como bons locais de PO;
- Tornar o planejamento realizado em sala de aula aceitável, prático e adequado;
- Desapego da carta e planos; turmas muito grandes; dificuldade com o linguajar e estimativa de distâncias; e

- Em se tratando “que já estamos no ET”, há dependência das condições meteorologias em determinadas situações. As maiores dificuldades são antes do ET, logística e preparo dos instrutores.

Em sua opinião, estando em um local de instrução (ocupando o PO com os alunos e instrutores), quais as dificuldades para se instruir os alunos sobre o estudo topotático do terreno?

- Boa parte dos alunos não tinham tido essa instrução ou essa experiência durante a carreira, isso faz com que eles fiquem muito amarrados a carta e tenham dificuldade de entender o terreno. Uma consequência dessa dificuldade é fazer o aluno compreender que a manobra planejada deverá ser ajustada por conta das variáveis observadas no terreno;
- No terreno, é necessário assumir diversas perspectivas diferentes. Por exemplo, o melhor posto de observação para mostrar os campos de tiro pode não ser suficiente para observarmos os obstáculos. Dessa forma, a quantidade de deslocamento necessário pode ser considerada dificuldade logística;
- Manter todos os OA atentos ao conteúdo apresentado mesmo com as dificuldades naturais de estar no PO;
- O aluno assimilar o que foi feito na carta no terreno. Muitos alunos após uma pergunta ficam olhando para a carta;
- Nem todo local de instrução reúne as condições adequadas para a condução dos exercícios, sobretudo em relação aos aspectos doutrinários essenciais que devem trabalhados (muitas vezes por imposição curricular) com os alunos de acordo com o tipo de operação estudada. Mesmo em se tratando somente da análise topotática, a depender do tipo de operação, cuja característica seja melhor adaptável a um determinado ambiente operacional, o perfil topográfico e demais idiossincrasias do terreno afetarão a análise sobremaneira. Ou seja, por mais que seja utilizado um mesmo padrão mnemônico (OCOAV), o estudo deve ser modulado de acordo com as peculiaridades de cada ambiente operacional. Isso inevitavelmente impõe a necessidade da ida ao terreno específico, a fim de que se consiga trabalhar tais particularidades;
- Após as primeiras idas ao terreno, nota-se uma grande evolução dos alunos;
- Alguns tem dificuldade em expressar em palavras o que está vendo no terreno;
- A grande dificuldade é conseguir mostrar para o OA como o terreno se comporta de acordo com a carta, de forma a propiciar o estudo válido para o planejamento. O estudo do terreno é uma arte;

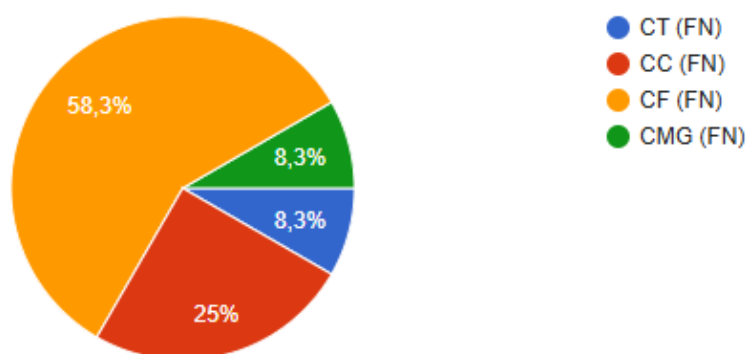
- Falta de certeza se os alunos realmente identificaram o que se pretende mostrar. Dificuldade com o volume da voz, quando se têm muitos alunos;
- Propagação do som e distrações;
- Condições meteorológicas desfavoráveis, como pouco visibilidade;
- Às vezes o tempo limitava o aprofundamento da análise;
- As características que são “visualizadas” no terreno exigem do instrutor e alunos, quando estão descrevendo o terreno, a capacidade de bem detalhá-lo, o que nem sempre é claro;
- O aluno observar o terreno propriamente dito pela primeira vez;
- Por vezes, o local do PO não suporta a quantidade de Alunos ao mesmo tempo, implicando em atrasos e dificuldade de passar as mesmas informações para todos os Alunos. Nestes casos o local do PO pode não oferecer boas condições de visibilidade para todos os Alunos que estejam no mesmo PO ou ainda terem visão parcial de pontos de interesse no terreno que permitiriam explorar bem a manobra por parte dos Instrutores;
- Fazer com que os Alunos deixem de pensar na dimensão 2D da carta e focar na visão 3D do terreno;
- Clareza sobre dominância (de fogos por exemplo) das elevações sobre determinada posição;
- Condições de solo devido as distância de observação; e
- Transmitir o entendimento que as vezes não é fácil para quem não tem vivência.

B

Respostas do questionário qualitativo e quantitativo aplicado a Oficiais

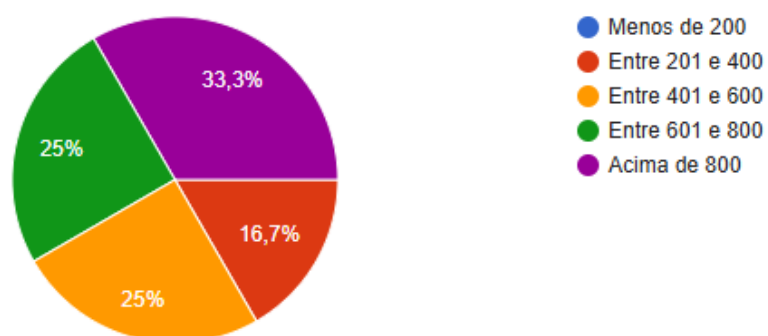
Posto:

12 respostas



Cerca de quantos dias de manobra e exercício possui?

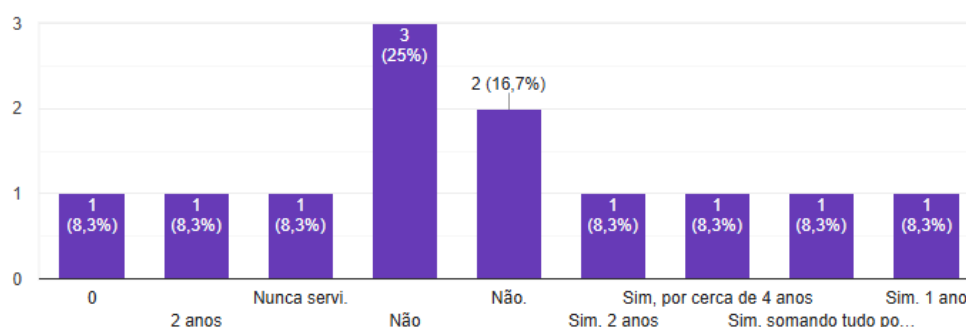
12 respostas



Serve ou serviu em seção de inteligência de Unidade operativa? Caso sim, por quanto tempo?

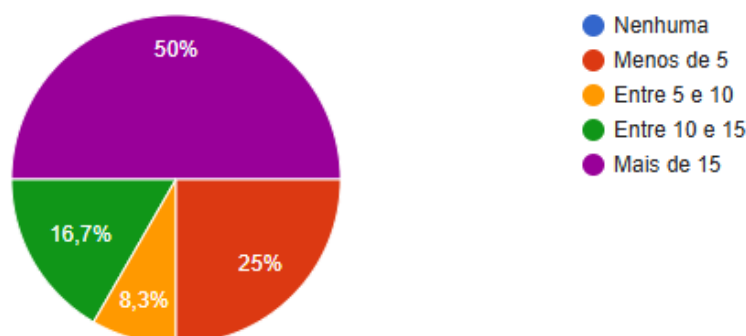
[Copiar gráfico](#)

12 respostas



Já realizou o planejamento de quantas manobras ou exercícios?

12 respostas



Já realizou planejamento de alguma Operação Militar? Caso sim, poderia citar os tipos de operações que já planejou e as quantidades?

- Nunca realizei um planejamento de uma Operação.
- Formosa (5 vezes, talvez), Dragão/UAnfEx (4 vezes, talvez), Ribeirex (3 vezes), OpPaz (1 vez), GLO (5 vezes talvez), Ajuda Humanitária (1 vez - TAQUARI), etc
- Ciclo de adestramento da FFE (6 Adst-Equipe, 6 Adst-Sub / Subex, 02 Três Corações, 5 Formosa, 4 Dragão); ET para EsGANf (E-QTe-GANf, EAO e C-Ap-EGANf - CADIM, Avelar, Resende, Rio Bonito, São Pedro da Aldeia, Macaé, Pirai e Furnas).
- Operações Anfíbias, terrestres e GLO - mais de 10.
- Nenhum.
- Ajuda humanitária - 7; Desembarque Anfíbia - 2 e GLO - 10.
- Operação de GLO na Maré e Op de paz no Haiti

- Sim, Operações Rio (muitas), Operações Formosa (muitas), Ágata Oeste (uma), ACRUX (uma), Operação Jogos Mundiais da Juventude (uma), Olimpíadas, entre outras.
- Sim. Operações Agata na Amazônia (4), Operações de patrulhamento (mais de 10) e Operações de Cerca e Vasculhamento (mais de 5). Todas reais.
- GLO. Duas
- Não
- Não.

Como o Sr utilizaria as informações da Máscara para o planejamento de uma operação militar *
nesta região?



- Usaria as duas para marcar Objetivos, traçar eixos de progressão, limites entre SU e linhas de coordenação.
- Identificação de obstáculos e acessos.
- Iria sobrepor a imagem satélite a fim de estabelecer itinerários e caso fosse necessário no curso da operações da RAT/RET itinerários e a até para estabelecimento da defensiva.
- Ao indicar as medidas de coordenação e controle do calco de operações e ao mostrar o avanço e desenvolvimento das tropas como um todo. Por essa imagem, me parece adequado para uma valor de batalhão ou superior.
- Para orientar o deslocamento da tropa.
- Máscara
- As informações de máscara são importantes para identificar pontos de importância, trafegabilidade do terreno, as vezes até informação de trânsito pode auxiliar no planejamento.
- Subsídio para trafegabilidade e deslocamento dos meios e tropas; cobertas e abrigos; locais de posicionamento dos apoios; dependendo da situação, setores a serem limpos. Muito próximo a uma Carta de Trafegabilidade.
- Colocaria sobreposto na imagem satelite.
- Trafegabilidade
- Trafegabilidade e traçar medidas de coordenação e controle
- Utilizaria para definir os itinerários para deslocamento de tropa em direção aos objetivos ou para retraimento.

O Sr poderia citar informações da Máscara que poderiam ser extraídas para planejar operações militares?

*



- Possíveis obstáculos para o avanço das universidades/SU, dependendo do meio de transporte empregado, vias de acessos mais favoráveis para progressões.
- Obstáculos, que poderiam gerar efeito de canalizar ou dividir as tropas. Uso dos arruamentos para um deslocamento mais rápido, porém aberto a campos de tiro. Terreno com observação extremamente limitada, a partir do solo, face à densidade de construções.
- Informações práticas para o movimento de tropas e meios), ela serve justamente para transformar o mapa “puro” em algo operacional. Direções de provável desenvolvimento; Obstáculos naturais e artificiais; Trechos impeditivos/restritivos para Vtr sr/sl; possíveis VÁ; Cobertas e Abrigos; Campos de Tiro; e Rotas Logísticas e de Apoio.
- Principais faixas do terreno a serem utilizadas pela natureza de tropa empregada. Seleção de local adequado para emprego do esforço principal. Seleção de áreas passivas ou com posturas defensivas mais permissivas.
- Regiões que permitem ou não o deslocamento da tropa.
- Terreno, vias de acesso, arruamento, pontos críticos.
- Máscara com os sentidos das vias urbanas; com o fluxo dos rios e as lagoas existentes no terreno
- Restrições/impeditivos/facilitadores ao deslocamento dos meios e tropa, tais como casarios (área urbana), locais alagadiços, corte de estradas, mata densa, ...
- Água, mata densa, estrada e arruamentos.
- Ruas largas duas vias / Elevações e dobras no terreno.

- Todas as constantes na legenda são importantes e úteis.
- Locais de homizio, zonas de reunião, obstáculos, itinerários, pontos críticos etc.

Comparando com uma carta de trafegabilidade convencional, o Sr poderia citar vantagens para o uso da Máscara como uma carta de trafegabilidade?

*



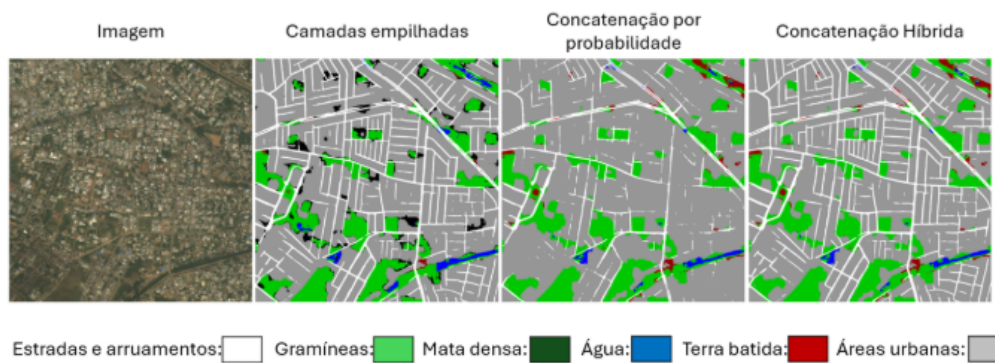
- Essas cartas possuem um detalhamento maior, com dimensões com proporcionalidade mais próximas ao real.
- Identificação das construções (possível até ver a altura de determinadas construções).
- A carta convencional é rica em detalhes brutos, mas exige interpretação técnica; já a máscara entrega uma síntese “pronta” para a decisão militar, acelerando o planejamento e reduzindo falhas.
- Possibilidade de seleção de filtros e camadas de forma mais fácil para focar em uma atividade ou natureza de tropa específica.
- Ainda não percebi as possíveis vantagens.
- Facilidade para identificar os itens citados anteriormente
- Excluir informações que não são relevantes para a trafegabilidade.
- Uma definição maior.
- Facilita a sobreposição de imagens e consegue-se tomar decisões mais rápidas
- Facilidade de identificar itinerários.
- Creio que a atualização dos dados e possibilidade de escolha das informações a serem avaliadas são vantagens.
- Os acidentes do terreno ficam melhor delimitados.

Comparando com uma carta de trafegabilidade convencional, o Sr poderia citar desvantagens *
para o uso da Máscara como uma carta de trafegabilidade?



- A carta da esquerda, talvez, possa ser um pouco mais difícil de se orientar em certos casos.
- Os obstáculos nem sempre aparecem de modo claro.
- A máscara é excelente para síntese rápida e visão operacional, mas perde em profundidade, precisão e flexibilidade de interpretação. Na prática, o ideal é usar as duas em conjunto: a máscara para a decisão rápida e briefing, e a carta convencional para o planejamento detalhado e navegação.
- A ausência da imagem satélite ou outros elementos que são retirados para clarear a imagem reforçam a dependência da análise do software ou da Inteligência Artificial aplicada, impedindo uma avaliação própria do planejador, fato que passa a ser relevante quando o planejador é um oficial mais antigo. A Máscara retira o fator experiência do processo de planejamento.
- Ainda não percebi desvantagem.
- Possível identificação do tipo de construção e da unidade de seu terraço para facilitar o deslocamento da tropa.
- Algum detalhe importante pode não ser observado
- A princípio, não vejo uma desvantagem.
- Possibilidade de dificuldade de leitura da carta sobreposta
- Identificação de facilidades logísticas. Tipos de construções.
- Não há
- Não apresenta curvas de nível, para avaliação da altitude dos pontos notáveis.

Escolha a melhor representação: *



- Concatenação híbrida
- Camadas empilhadas
- Concatenação híbrida
- Camadas empilhadas
- Concatenação híbrida
- Camadas empilhadas
- Camadas empilhadas
- Camadas empilhadas
- Camadas empilhadas
- Concatenação híbrida
- Concatenação híbrida
- Concatenação híbrida