



UNIVERSIDADE
LUSÓFONA

Segmentação de incêndios florestais

Trabalho Final de Curso

Relatório Intercalar 1º Semestre

Francisco Alves, a22207107, Ciencia de Dados

Orientador: João Carvalho

Orientador: Houda Harkat

Departamento de Engenharia Informática e Sistemas de Informação

Universidade Lusófona, Centro Universitário de Lisboa

01/12/2024

www.ulusofona.pt

Direitos de cópia

Segmentação de incêndios florestais, Copyright de Francisco Perpétuo Alves,
Universidade Lusófona.

A Escola de Comunicação, Arquitectura, Artes e Tecnologias da Informação (ECATI) e a Universidade Lusófona (UL) têm o direito, perpétuo e sem limites geográficos, de arquivar e publicar esta dissertação através de exemplares impressos reproduzidos em papel ou de forma digital, ou por qualquer outro meio conhecido ou que venha a ser inventado, e de a divulgar através de repositórios científicos e de admitir a sua cópia e distribuição com objectivos educacionais ou de investigação, não comerciais, desde que seja dado crédito ao autor e editor.

Resumo

Os incêndios florestais constituem um dos desastres naturais mais devastadores, causando impactos ambientais, sociais e económicos significativos. Estes eventos ameaçam ecossistemas inteiros, comprometem a biodiversidade, colocam vidas humanas e animais em risco e contribuem para o agravamento das alterações climáticas devido à emissão de gases com efeito de estufa. Assim, a deteção precoce e eficiente de incêndios é essencial para mitigar os seus danos e permitir intervenções rápidas e eficazes.

O desenvolvimento de incêndios florestais ocorre em duas fases principais: a emissão de fumo, que é o primeiro indicador visível do início de uma combustão, e, posteriormente, o aparecimento de chamas, que marcam um estado mais avançado e perigoso do fogo. A deteção de fumo é particularmente desafiante devido à sua natureza translúcida e à dispersão em condições ambientais adversas, como ventos ou luminosidade variável. Por outro lado, o fogo é mais fácil de identificar, mas quando visível, normalmente já representa uma situação crítica. Assim, um sistema de deteção eficaz deve ser capaz de identificar tanto o fumo quanto o fogo para facilitar uma intervenção atempada.

Esta trabalho propõe o desenvolvimento de um sistema avançado baseado em deep learning para a segmentação simultânea de fumo e fogo em imagens. Serão exploradas duas arquiteturas reconhecidas pela sua eficiência em tarefas de segmentação, a U-Net e a ResUNet-a, ambas capazes de realizar classificações pixel-a-pixel em cenários complexos. Os modelos serão treinados e avaliados utilizando dois conjuntos de dados públicos, Corsican e FLAME, que incluem imagens captadas por drones e sensores térmicos.

O estudo também investigará a melhoria dessas arquiteturas por meio de modificações estruturais para aumentar a sua precisão e robustez, especialmente em condições ambientais desafiadoras. A otimização das arquiteturas visa melhorar a deteção de padrões complexos, como os contornos irregulares de fumo e fogo, e garantir uma maior eficácia em diferentes cenários.

Espera-se que os resultados desta trabalho demonstrem a viabilidade e a eficácia de sistemas baseados em deep learning para a deteção precoce de incêndios florestais. Esta abordagem representa um avanço significativo ao combinar a identificação de fumo e fogo numa única solução, com potencial para transformar as estratégias de monitorização, resposta a emergências e preservação ambiental.

Palavras-chave: Incêndios florestais; Deep learning, Segmentação de imagens; U-Net, resUNet-a, Visão computacional.

Abstract

Wildfires are among the most devastating natural disasters, causing significant environmental, social, and economic impacts. These events threaten entire ecosystems, compromise biodiversity, put human and animal lives at risk, and contribute to worsening climate change due to greenhouse gas emissions. Therefore, early and efficient detection of fires is essential to mitigate their damage and enable quick and effective interventions.

Wildfires develop in two main phases: the emission of smoke, which is the first visible indicator of the start of combustion, and then the appearance of flames, which mark a more advanced and dangerous stage of the fire. Smoke detection is particularly challenging due to its translucent nature and dispersion in adverse environmental conditions, such as winds or variable lighting. On the other hand, fire is easier to identify, but when it is visible, it usually already represents a critical situation. Therefore, an effective detection system must be able to locate both smoke and fire to facilitate timely intervention.

This work proposes the development of an advanced system based on deep learning for the simultaneous segmentation of smoke and fire in images. Two architectures recognized for their efficiency in segmentation tasks, U-Net and ResUNet-a, will be explored, both capable of performing pixel-by-pixel classifications in complex scenarios. The models will be trained and evaluated using two public datasets, Corsican and FLAME, which include images captured by drones and thermal sensors.

The study will also investigate improving these architectures through structural modifications to increase their accuracy and robustness, especially in challenging environmental conditions. The optimization of the architectures aims to improve the detection of complex patterns, such as the irregular contours of smoke and fire, and ensure greater effectiveness in different scenarios.

The results of this work are expected to demonstrate the feasibility and effectiveness of deep learning-based systems for the early detection of wildfires. This approach represents a significant advance by combining smoke and fire identification in a single solution, with the potential to transform monitoring, emergency response, and environmental preservation strategies.

Keywords: Wildfires; Deep learning, Image segmentation; U-Net, ResUNet-a, Computer vision.

Índice

Resumo	2
Abstract	3
Índice	4
Lista de Figuras	5
Lista de Tabelas	6
1- Introdução	7
2- Estado de arte	10
3 - Metodologias	18
4 - Calendário	22
5 - Conclusão	24
Bibliografia	25

Listas de Figuras

1	Perda de cobertura global de árvores devido a incêndios, 2001-2022	10
2	Exemplo de uma imagem do dataset	11
3	Máscara criada manualmente para a imagem do dataset	11
4	Arquitetura MaskSU R-CNN	13
5	Arquitetura geral de segmentação semântica de incêndios	14
6	Arquitetura U-Net	15
7	Arquitetura ResUNet-a	20
8	Calendário de atividades do projeto	22

Lista de Tabelas

- 1 Comparação de algoritmos para segmentação de incêndios florestais . . . 16

1- Introdução

1.1 Motivação

Os incêndios florestais representam uma ameaça crescente em várias partes do mundo, afetando não só o meio ambiente, mas também a vida das pessoas e a economia das regiões afetadas. Estes eventos devastadores têm sido amplificados pelas alterações climáticas, resultando em incêndios mais frequentes e intensos. O impacto ambiental é particularmente grave, com a destruição de vastas áreas de floresta, a perda de biodiversidade e a contribuição significativa para as emissões de gases com efeito de estufa, exacerbando o problema global das mudanças climáticas. Além disso, as comunidades afetadas por incêndios florestais enfrentam grandes desafios económicos e sociais, desde a destruição de bens e infraestruturas até à perda de vidas humanas.

A deteção precoce de incêndios florestais é um dos fatores mais críticos para minimizar estes danos. Identificar um foco de incêndio logo nas suas fases iniciais pode ser determinante para um combate eficaz e para a proteção de vidas e bens. Tradicionalmente, a deteção de incêndios tem sido realizada por meio de vigilância humana ou sistemas baseados em sensores. No entanto, estes métodos apresentam limitações em termos de cobertura e precisão, especialmente em áreas remotas ou de difícil acesso, onde a intervenção rápida é crucial.

Com o avanço das tecnologias de monitorização por satélite, drones e sensores, surgem novas oportunidades para melhorar a deteção de incêndios em tempo real. Uma dessas abordagens emergentes é o uso de técnicas de visão computacional combinadas com algoritmos de deep learning. Estes algoritmos, ao processarem grandes volumes de dados provenientes de imagens aéreas, podem identificar padrões visuais complexos, como o fumo e o fogo, de forma muito mais eficiente do que os métodos tradicionais. A segmentação de imagens, por exemplo, permite a separação precisa de áreas de interesse, como os focos de incêndio, facilitando a tomada de decisões rápidas e informadas.

A grande importância deste trabalho reside na exploração de soluções tecnológicas que possam aumentar a eficácia dos sistemas de deteção de incêndios florestais, contribuindo para a sua antecipação e combate. A utilização de imagens aéreas e algoritmos avançados oferece uma oportunidade única para superar as limitações dos métodos tradicionais, proporcionando uma abordagem mais eficaz e rápida para detetar focos de incêndio logo nas suas fases iniciais. A relevância deste trabalho é ainda maior considerando que a deteção precoce é fundamental para uma resposta eficaz e para a minimização dos danos ambientais e sociais causados por esses eventos.

1.2 Objetivo

O objetivo deste estudo é desenvolver e avaliar um sistema baseado em deep learning para a deteção de incêndios florestais em tempo real, com ênfase na redução do tempo de resposta e na melhoria da eficácia das operações de combate. A proposta visa usar modelos de aprendizagem profunda para identificar e localizar focos de fogo e fumo, informação crucial para uma resposta precoce e eficiente. Através da utilização de técnicas avançadas de segmentação de imagens, este trabalho procura aprimorar a capacidade de distinguir entre áreas queimadas e áreas não afetadas, possibilitando uma análise detalhada e precisa.

Com isso, espera-se não apenas otimizar os processos de monitorização em tempo

real, mas também contribuir para a redução dos impactos ambientais, sociais e económicos provocados pelos incêndios florestais. A aplicação de soluções tecnológicas inovadoras permitirá um maior grau de automação e assertividade nas decisões de gestão de riscos, tornando o sistema mais eficaz na minimização dos danos causados por incêndios em áreas florestais e nas comunidades circundantes.

1.3 Pergunta de pesquisa

É possível aumentar significativamente a precisão e a eficácia na deteção precoce de incêndios florestais utilizando técnicas de deep learning aplicadas a imagens de incêndios?

1.4 Hipótese

A hipótese central deste estudo é que a aplicação de técnicas avançadas de deep learning, especificamente através de modelos de segmentação, pode melhorar substancialmente a deteção precoce e a precisão na identificação de incêndios florestais. A utilização de redes neurais profundas, quando aplicadas a imagens aéreas ou de satélite, permite não só a segmentação eficaz dos focos de fogo, mas também a capacidade de detetar o fumo de maneira antecipada, o que é crucial para uma resposta rápida.

Espera-se que, ao otimizar os modelos existentes e adaptá-los a diferentes conjuntos de dados, seja possível alcançar uma solução robusta e precisa que identifique focos de incêndio e fumo com uma taxa mínima de falsos positivos e falsos negativos. A ideia é que, com a melhoria contínua das arquiteturas utilizadas, a deteção de incêndios se torne mais eficaz, rápida e adaptável a diferentes cenários, reduzindo significativamente o tempo de resposta e minimizando os danos causados pelos incêndios florestais.

Este estudo também investiga como a adaptação desses modelos pode levar a uma maior eficiência na utilização de recursos em sistemas automáticos de monitorização, além de contribuir para o desenvolvimento de tecnologias mais eficazes na mitigação dos impactos ambientais e sociais dos incêndios. A hipótese é que, ao aplicar e refinar esses modelos, será possível criar uma abordagem mais precisa e eficiente para a deteção e monitorização em tempo real de incêndios florestais.

1.5 Etapas do relatório

O presente relatório está organizado em três capítulos principais, estruturados para apresentar, de forma clara e sistemática, os objetivos, os fundamentos teóricos e a abordagem prática adotada neste trabalho.

No Capítulo 1 - Introdução, são apresentados os aspectos motivadores deste estudo, evidenciando a importância e a relevância do tema abordado. Este capítulo também estabelece os objetivos gerais e específicos do trabalho, formula a pergunta de pesquisa e apresenta a hipótese proposta. Esses elementos fundamentam o problema em questão, oferecendo uma visão clara do propósito deste estudo e justificando sua realização no contexto científico e prático.

O Capítulo 2 - Estado da Arte, oferece uma revisão teórica e prática sobre os principais tópicos relacionados à pesquisa. Inicialmente, são descritos os principais datasets utilizados na área, com destaque para suas características, diversidade e representatividade. Em seguida, são exploradas as técnicas de anotação, detalhando os processos utilizados para criar um dataset de forma consistente e confiável. Posteriormente, são

analisados os algoritmos existentes, discutindo suas características, abordagens e desempenho. Por fim, é apresentada uma comparação entre os algoritmos, com base em métricas de avaliação e aplicações específicas, identificando as limitações e oportunidades para futuras melhorias.

O Capítulo 3 - Metodologia, descreve as abordagens utilizadas para implementar e avaliar os modelos propostos. Em particular, são exploradas as arquiteturas U-Net e ResUNet-a, com explicações detalhadas das arquiteturas e como aumentar a eficácia da segmentação. São também discutidas técnicas para otimizar o treino dos modelos, como ajustes de hiperparâmetros e funções de perda adequadas a problemas com classes desequilibradas.

O Capítulo 4 - Calendário, detalha as etapas organizacionais do trabalho, com base num cronograma estruturado que abrange desde a pesquisa científica até à redação do relatório. Este capítulo expõe o planeamento das atividades e discute as estratégias utilizadas para alcançar os objetivos de forma eficiente e no prazo estipulado. O calendário apresentado encontra-se atualizado até à data deste relatório, refletindo as etapas concluídas e os ajustes necessários para garantir a realização dos objetivos no período estabelecido.

O Capítulo 5 - Conclusão, apresenta as reflexões finais sobre o trabalho desenvolvido, destacando a relevância do tema e as contribuições do sistema proposto. Este capítulo discute os principais passos realizados durante o projeto, bem como as potencialidades da abordagem baseada em deep learning para a segmentação de fumo e fogo. Além disso, são abordados os benefícios esperados desta solução para a área de monitorização de incêndios florestais, concluindo com a importância do trabalho para avanços científicos e aplicações práticas.

2- Estado de arte

O impacto e desafios dos incêndios florestais

Os incêndios florestais são uma ameaça crítica para os ecossistemas naturais, causando perdas significativas para a biodiversidade, degradação ambiental e impactando diretamente o bem-estar humano. As florestas, frequentemente chamadas de "pulmões do planeta", desempenham um papel essencial na filtragem do ar, absorvendo dióxido de carbono (CO₂) e libertando oxigénio (O₂), além de abrigar uma ampla diversidade de vida selvagem e proteger as culturas agrícolas contra eventos climáticos extremos (Zanchi et al., 2021 [1]; San-Miguel-Ayanz et al., 2013 [2]). A gravidade desses impactos pode ser observada na perda global de cobertura arbórea entre os anos de 2001 e 2022, apresentada na Figura 1. Os pontos vermelhos destacam as áreas afetadas por incêndios florestais, algumas das quais estão sob esforços de restauração significativos.

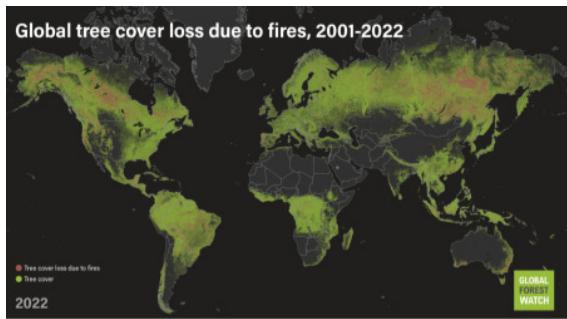


Figure 1: Perda de cobertura global de árvores devido a incêndios, 2001-2022

2.1 Datasets

O presente trabalho recorre a dois conjuntos de dados públicos utilizados na área de deteção de incêndios florestais: o Corsican[3] Dataset e o FLAME[4] Dataset. Estes datasets fornecem imagens de incêndios capturadas em diferentes condições, contribuindo significativamente para o treino e validação dos modelos de segmentação e deteção. Além disso, foi criado um dataset personalizado com imagens provenientes desses dois conjuntos públicos, mas com a adição manual de máscaras de anotação detalhadas, que incluem a segmentação de áreas de fogo, fumo e fundo (ground truths). Estas máscaras serão feitas manualmente, assegurando um elevado nível de precisão, o que permitirá ao modelo aprender de forma mais eficaz e com maior adaptação às características específicas do cenário em análise.

2.1.1 Datasets Description

Corsican Dataset

O Corsican Dataset[3] contém aproximadamente 2000 imagens de incêndios florestais capturadas sob diferentes configurações de câmaras, utilizando espectros visíveis e infravermelhos próximos. As imagens apresentam resolução de 1024 × 768 pixels e estão no formato PNG. Este conjunto de dados inclui imagens multimodais obtidas por câmaras

como a "JAI AD-080GE", capazes de captar simultaneamente espectros visíveis e infravermelhos através de óticas alinhadas. Cada imagem do dataset é acompanhada por uma máscara de segmentação correspondente, criada utilizando técnicas de homografia para registo preciso das imagens.

FLAME Dataset

O FLAME Dataset[4] consiste principalmente em vídeos capturados por drones utilizando câmaras como a Zenmuse X4S e Phantom 3, com uma resolução de 640×512 pixels e taxa de 30 quadros por segundo. Para fins de segmentação, frames foram extraídos dos vídeos e já incluem máscaras anotadas correspondentes, tornando o dataset adequado para experimentação em tarefas de deteção de incêndios. Este conjunto de dados fornece 2003 imagens anotadas, com variabilidade em condições de captura, como diferentes níveis de iluminação e presença de fumo.

2.1.2 Datasets annotation technique

As máscaras de anotação, essenciais para o treino supervisionado dos modelos, foram criadas manualmente utilizando a ferramenta MATLAB ImageLabeler [5]. O processo de anotação envolveu a classificação de cada pixel das imagens como fogo, fumo ou fundo. A Figura 2 apresenta uma imagem do dataset de incêndios florestais, e a Figura 3 mostra a sua máscara feita manualmente.



Figure 2: Exemplo de uma imagem do dataset



Figure 3: Máscara criada manualmente para a imagem do dataset

2.2 Algoritmos Existentes

A deteção de incêndios florestais utilizando imagens aéreas tem sido amplamente estudada com diferentes abordagens de algoritmos. Cada um desses algoritmos apresenta

vantagens e desvantagens dependendo do tipo de imagem utilizada, das condições ambientais e dos objetivos do estudo. A seguir, apresentamos uma análise comparativa de alguns estudos recentes que utilizam diferentes métodos para detetar incêndios florestais em imagens, com foco nas metodologias, resultados e limitações de cada um.

MaskSU R-CNN

O artigo "Forest Fire Segmentation from Aerial Imagery Data Using an Improved Instance Segmentation Model"[6] apresenta o modelo MaskSU R-CNN, ilustrado na Figura 4, uma evolução do Mask R-CNN, desenvolvida especificamente para melhorar a segmentação de áreas de fogo em cenários florestais. A deteção e segmentação de regiões de fogo em imagens são tarefas cruciais para a prevenção e controlo de incêndios, mas enfrentam desafios significativos devido ao baixo contraste entre o fogo e o fundo, à natureza dinâmica das chamas e às condições ambientais adversas. O MaskSU R-CNN propõe uma solução que aborda estas limitações, integrando melhorias ao Mask R-CNN para alcançar uma segmentação mais precisa e eficiente.

Inspirado pela arquitetura U-Net [7], o MaskSU R-CNN incorpora uma ramificação MaskIoU em formato de "U", o que otimiza a segmentação de incêndios florestais ao preservar características importantes durante o processamento. O modelo beneficia de aprendizagem por transferência, utilizando pesos pré-treinados no dataset COCO. Este método acelera o processo de treino, aproveitando as características extraídas de um grande conjunto de dados diversificado para melhorar o desempenho do modelo em bases de dados específicas e menores, como as usadas na deteção de fogo. Durante o treino, pixels da imagem são selecionadas e classificadas como positivas ou negativas com base no limite de IoU (Intersection over Union), o que impacta diretamente a função de perda e direciona o modelo para uma segmentação mais robusta.

O treino do MaskSU R-CNN 4 é orientado por uma função de perda multibranch, composta por múltiplos componentes: perda RPN (Region Proposal Network), perda de classificação, perda de regressão de caixas delimitadoras, perda de máscara e perda de MaskIoU. Estes componentes trabalham de forma integrada para detetar e segmentar áreas de fogo com alta precisão. A perda RPN é responsável pela geração de propostas de regiões relevantes, enquanto a regressão de MaskIoU mede a qualidade da segmentação ao comparar as máscaras previstas com as máscaras de referência (ground truth). Este mecanismo é fundamental para otimizar a sobreposição entre as máscaras previstas e as reais, contribuindo para a melhoria contínua do modelo durante o treino.

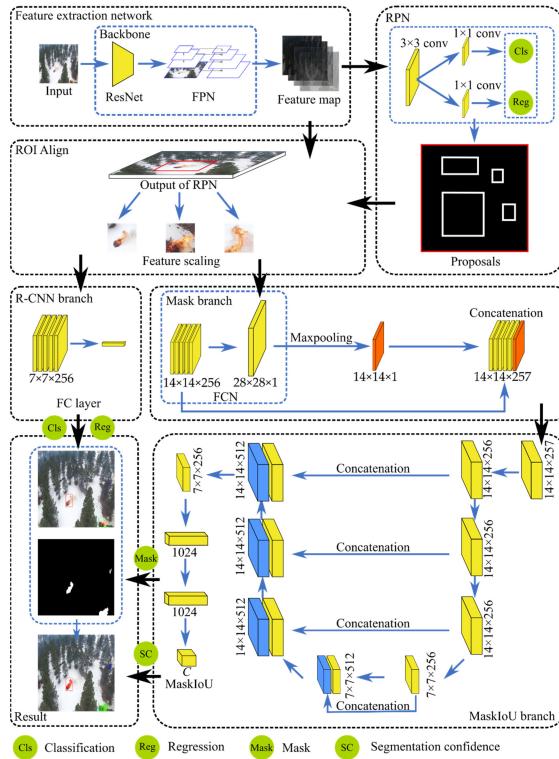


Figure 4: Arquitetura MaskSU R-CNN

A arquitetura equilibrada do MaskSU R-CNN permite ajustar os hiperparâmetros de forma a manter um equilíbrio eficaz entre classificação, regressão de caixas delimitadoras e segmentação de máscaras. Esta abordagem garante uma convergência estável durante o treino e resultados consistentes em diferentes cenários.

Os resultados apresentados no artigo demonstram a eficácia do MaskSU R-CNN, que alcança uma precisão de 91,85%, recall de 88,81%, F1-score de 90,30% e um mean intersection over union (mIoU) de 82,31%. Estes valores destacam a capacidade do modelo para realizar uma segmentação precisa e eficiente, mesmo em condições desafiadoras. Assim, o MaskSU R-CNN posiciona-se como uma ferramenta promissora para a deteção e monitorização de incêndios florestais, oferecendo suporte fundamental para a gestão e mitigação de desastres ambientais.

GPG e MEI

O artigo "Optimized Deep Learning Model for Fire Semantic Segmentation"[8] apresenta uma abordagem para segmentação semântica de incêndios, ilustrada na figura 5 , abordando desafios relacionados à identificação precisa de áreas afetadas pelo fogo. O trabalho propõe melhorias específicas em modelos de deep learning, com foco na correção de erros nas bordas das áreas em chamas e na captação de informações contextuais importantes para o cenário de incêndios.

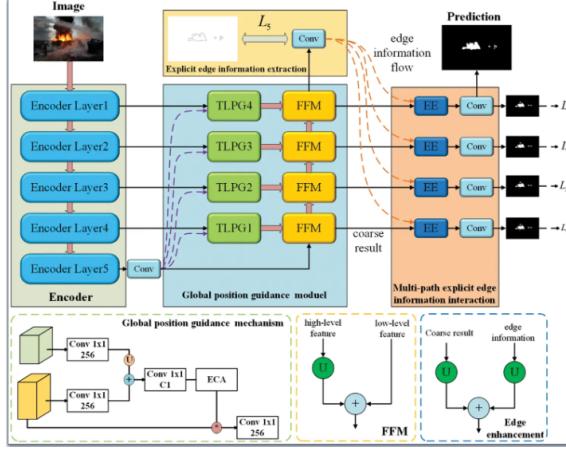


Figure 5: Arquitetura geral de segmentação semântica de incêndios

Os autores desenvolveram dois módulos principais para otimizar a segmentação. O Guia de Posição Global (Global Position Guidance - GPG) concentra-se na correção de falhas na identificação de características espaciais de baixo nível, oferecendo um mapeamento mais preciso da posição dos incêndios. O Módulo de Interação Explícita de Informações de Bordas (Multi-path Explicit Edge Information - MEI) tem como objetivo melhorar a definição e precisão das bordas das áreas de fogo, utilizando uma arquitetura multi-vias que integra informações de diferentes resoluções hierárquicas. Este design permite uma análise detalhada e consistente, mesmo em condições desafiadoras, como a presença de fumo ou reflexos.

Os modelos foram avaliados em bases de dados específicas de incêndios utilizando métricas amplamente reconhecidas no campo. Para a métrica Intersection over Union (IoU), o modelo alcançou um desempenho de 82,3%, superando métodos anteriores como DeepLabv3+ e U-Net. Além disso, obteve uma precisão global de 92,1%, demonstrando uma significativa melhoria na segmentação de áreas afetadas pelo fogo e na identificação de bordas com maior detalhe.

U-net

O artigo "Aerial Imagery Pile Burn Detection Using Deep Learning: The FLAME Dataset" [9] apresenta uma metodologia para segmentação de imagens, ilustrada na figura 6, focada na identificação de áreas com fogo em imagens aéreas. A segmentação é tratada como um problema de classificação binária por pixel, visando gerar máscaras que identificam regiões ativamente em chamas. Essa abordagem é particularmente útil em cenários como a deteção de pequenos focos de fogo e o monitorização de áreas de queima.

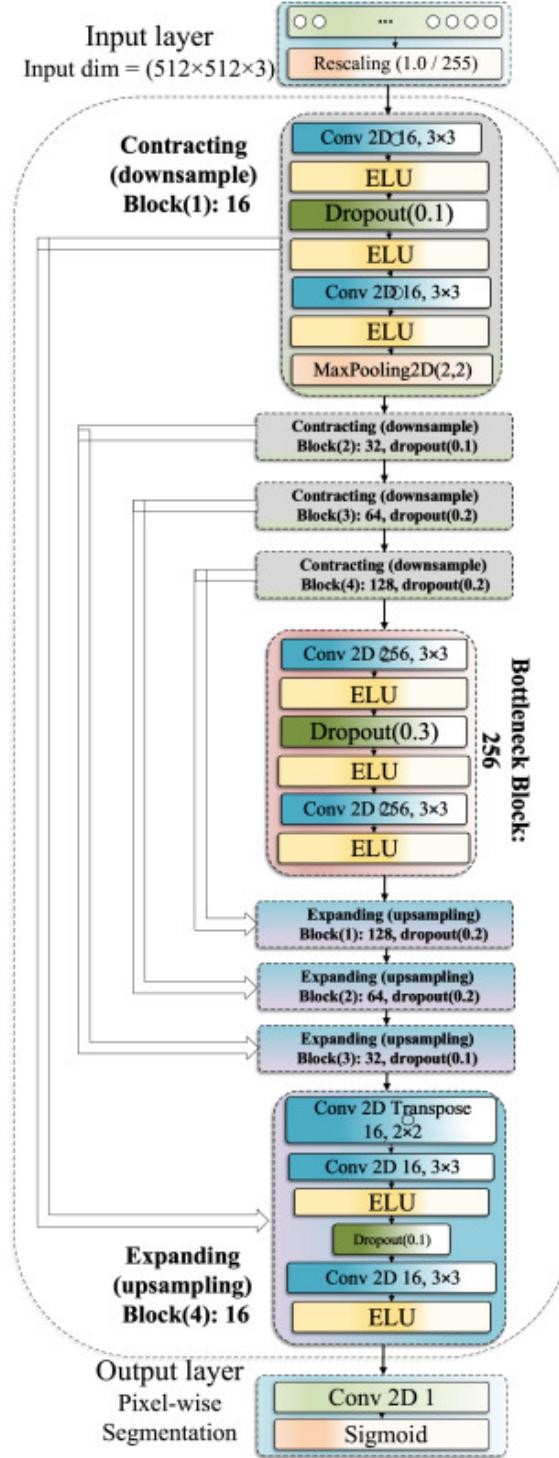


Figure 6: Arquitetura U-Net

Para alcançar esse objetivo, os autores utilizaram uma versão modificada da rede U-Net, conhecida por sua aplicação em segmentação de imagens biomédicas. No modelo adaptado, a função de ativação ReLU foi substituída pela Exponential Linear Unit (ELU), visando suavizar os resultados e melhorar a precisão. A arquitetura do modelo manteve a estrutura clássica da U-Net, composta por caminhos simétricos de contração e expansão conectados por operações de concatenação, o que preserva informações de alta resolução durante o processamento das imagens. Além disso, para minimizar problemas de overfitting devido à quantidade limitada de dados, foi aplicada a técnica de

dropout durante o treino.

O conjunto de dados utilizado foi o FLAME dataset, que contém imagens aéreas anotadas manualmente para a criação de máscaras que representam as regiões de fogo. A ferramenta MATLAB Image Labeler foi usada para gerar as anotações, e o dataset foi composto por 2003 frames. As imagens RGB foram normalizadas, com cada pixel classificado como "fogo" ou "não-fogo" (fundo). O modelo foi treinado com a função de perda binary cross entropy, otimizada pelo algoritmo Adam.

Os resultados obtidos demonstram a eficácia do modelo U-Net modificado. O desempenho foi avaliado com base em várias métricas: uma precisão de 91,99%, recall de 83,88%, área sob a curva (AUC) de 99,85%, F1-Score de 87,75%, sensibilidade de 83,12%, especificidade de 99,96% e Intersection over Union (IoU) de 78,17%. Esses resultados indicam uma capacidade robusta do modelo em segmentar regiões de fogo, oferecendo uma ferramenta valiosa para o monitorização e a gestão de incêndios em ambientes aéreos.

2.3 Comparação de Algoritmos Existentes

A Tabela 1 apresenta uma comparação entre diferentes algoritmos de deteção e segmentação de incêndios florestais, com base nos artigos analisados.

Artigo	Algoritmo(s)	Métricas (%)
<i>Forest Fire Segmentation from Aerial Imagery Data Using an Improved Instance Segmentation Model [6]</i>	MaskSU R-CNN	<ul style="list-style-type: none"> Precisão: 91,85% Recall: 88,81% F1-Score: 90,30% mIoU: 82,31%
<i>Optimized Deep Learning Model for Fire Semantic Segmentation[8]</i>	PGP e MEI	<ul style="list-style-type: none"> IoU: 82,3% Precisão global: 92,1%
<i>Aerial Imagery Pile Burn Detection Using Deep Learning: The FLAME Dataset[9]</i>	U-Net	<ul style="list-style-type: none"> Precisão: 91,99% Recall: 83,88% AUC: 99,85% F1-Score: 87,75% Sensibilidade: 83,12% Especificidade: 99,96% IoU: 78,17%

Table 1: Comparação de algoritmos para segmentação de incêndios florestais

2.4 Benchmarking

A aplicação de tecnologias avançadas de deep learning para a deteção de incêndios florestais tem vindo a crescer em todo o mundo, incluindo na Europa, onde iniciativas relevantes mostram como estas abordagens podem ser adaptadas a diferentes contextos ambientais. Este subcapítulo explora implementações relevantes de sistemas baseados em inteligência artificial em países europeus.

Sistema EFFIS (União Europeia)

O *European Forest Fire Information System (EFFIS)* é uma iniciativa da Comissão Europeia, criada para monitorizar e prever incêndios florestais em toda a Europa. O sistema utiliza principalmente dados de satélites para fornecer informações em tempo real sobre incêndios florestais e os riscos associados. Recentemente, esforços têm sido feitos para integrar algoritmos de inteligência artificial, visando melhorar a análise preditiva do risco de incêndios e sua propagação, ajudando a planejar estratégias de resposta mais eficazes. Além disso, o EFFIS é uma plataforma colaborativa que permite a troca de dados entre países da União Europeia, melhorando a capacidade de resposta coletiva no combate a incêndios.

Sistema FireSense (Grécia)

O FireSense é um sistema de monitorização de incêndios florestais desenvolvido na Grécia, com foco na deteção precoce de focos de incêndio em áreas de risco elevado, como parques naturais e regiões turísticas. Este sistema utiliza câmaras de infravermelhos e tecnologias de deep learning para a segmentação de imagens térmicas, permitindo a identificação de incêndios em fases iniciais. A tecnologia do FireSense tem sido usada em áreas de difícil acesso, como as ilhas de Creta e as ilhas Egeias, para melhorar a resposta a incêndios e reduzir os danos ambientais. A abordagem de deep learning aplicada no sistema permite uma análise mais precisa e eficaz das imagens térmicas, contribuindo para uma gestão mais rápida e eficaz dos incêndios.

Relevância destas Implementações para Portugal

As soluções implementadas em países da União Europeia oferecem ensinamentos valiosos para Portugal, um país que enfrenta desafios semelhantes devido ao seu clima mediterrâneo e às vastas áreas florestais. A adoção de tecnologias semelhantes, adaptadas às particularidades do território nacional, poderia fortalecer a eficácia das estratégias de prevenção e combate aos incêndios florestais, contribuindo para uma maior resiliência ambiental e garantindo mais segurança para as comunidades afetadas.

3 - Metodologias

3.1 Metodologia

A classificação e segmentação de focos de incêndio em imagens aéreas representam um desafio significativo devido à complexidade dos padrões visuais presentes, como a interferência do fumo, a translucidez das chamas e a variabilidade dos cenários ambientais. Métodos tradicionais de processamento de imagens, que utilizam abordagens baseadas em modelagem matemática ou redes neurais simples, frequentemente enfrentam limitações na identificação precisa de incêndios, resultando em falsos positivos ou negativos. Em contraste, as redes neurais profundas, especialmente as Redes Neuronais Convolucionais (CNNs), têm se destacado pela sua capacidade de aprender representações robustas e características discriminativas diretamente dos dados, superando essas limitações.

Neste estudo, foi utilizado arquiteturas U-Net e ResUNet para a segmentação semântica de fogo e fumo. Essas redes são reconhecidas por sua eficiência em preservar informações relevantes durante o processamento de imagens e por sua capacidade de lidar com detalhes em múltiplas escalas. A U-Net, uma arquitetura amplamente empregada em segmentação de imagens, é composta por caminhos simétricos de contração e expansão conectados por operações de concatenação, permitindo a extração e a combinação eficaz de características de diferentes resoluções. Por outro lado, a ResUNet incorpora blocos residuais, que aprimoram o fluxo de gradientes, facilitando o treino de redes profundas ao reduzir problemas como o vanishing gradient. Isso torna o modelo mais estável e eficaz na segmentação de padrões visuais complexos presentes nas imagens aéreas de incêndios.

A abordagem proposta também considera a necessidade de capturar características multi-escala, fundamentais para a identificação de chamas e fumo em diferentes níveis de detalhe. Para isso, mecanismos de aprimoramento de representação, como o uso de conexões residuais na ResUNet, são explorados para aumentar a capacidade de generalização da rede, preservando as informações críticas das imagens. Esses aspectos são particularmente importantes no contexto de incêndios florestais, onde os padrões variam significativamente com base na intensidade do fogo, nas condições climáticas e no tipo de vegetação.

O pipeline de implementação segue as etapas fundamentais de treino e validação dos modelos, utilizando dados anotados manualmente para a criação de máscaras que representam as áreas de interesse. O objetivo é avaliar o desempenho das arquiteturas U-Net e ResUNet em termos de métricas como accuracy, jaccard score, f1-score e intersection over union (IoU), garantindo a robustez e a eficácia dos modelos em cenários reais.

U-Net

O U-Net[7] é uma arquitetura de rede neural convolucional projetada para segmentação de imagens, inicialmente desenvolvida para aplicações biomédicas. O seu design exclusivo em forma de "U" combina um caminho de contração, que extrai características de alto nível, e um caminho de expansão, que reconstrói as dimensões originais da imagem. Este layout permite que o modelo capture tanto os detalhes contextuais mais amplos quanto os elementos espaciais mais finos, sendo altamente eficiente para tarefas que exigem segmentação pixel a pixel.

No caminho de contração, também conhecido como encoder, a imagem é processada por uma série de camadas convolucionais seguidas de operações de max pooling. Isso reduz progressivamente a resolução da imagem, ao mesmo tempo que aumenta a abstração das características extraídas. Já no caminho de expansão, ou decoder, a resolução é recuperada usando convoluções transpostas e operações de up-sampling, recriando a estrutura da imagem original com a segmentação desejada.

Um dos elementos-chave do U-Net é o uso de conexões de salto (skip connections), que transferem informações diretamente das camadas do encoder para as camadas correspondentes do decoder. Essas conexões garantem que detalhes críticos de alta resolução sejam preservados, o que é essencial para tarefas onde a precisão nas bordas ou contornos das regiões segmentadas é crucial, como na identificação de focos de incêndio em imagens.

A arquitetura U-Net é amplamente utilizada em problemas de segmentação semântica devido à sua capacidade de processar imagens inteiras e realizar previsões precisas de classificação para cada pixel. Para aumentar sua eficiência, podem ser aplicadas funções de perda como binary cross-entropy ou dice loss, e a rede é frequentemente otimizada com métodos como o Adam. Além disso, sua flexibilidade permite adaptações, como a inclusão de camadas adicionais ou o ajuste de funções de ativação, para lidar com desafios específicos, como desequilíbrio de classes nos dados.

ResUNet-A

O ResUNet-A [10] é uma evolução significativa da arquitetura ResUNet, concebido para lidar com tarefas de segmentação de alta complexidade, como a identificação de áreas de incêndios florestais em imagens de alta resolução. Este modelo combina os blocos residuais característicos da ResNet com elementos adicionais que ampliam sua capacidade de extração de características relevantes para a segmentação. A arquitetura do ResUNet-A 7 integra funcionalidades avançadas, como convoluções dilatadas (atrous convolutions), mecanismos de atenção espacial e camadas de agregação multi-escala, o que permite uma captura mais rica e detalhada de padrões espaciais e contextuais.

Uma das principais inovações do ResUNet-A é a inclusão de mecanismos de atenção que ajudam o modelo a focar seletivamente nas regiões mais relevantes da imagem, ignorando elementos de fundo que possam introduzir ruído no processo de segmentação. Esse enfoque é especialmente útil em cenários como a deteção de incêndios florestais, onde os sinais de fogo podem ser pequenos ou parcialmente obscurecidos por fumo ou vegetação.

Além disso, o uso de convoluções dilatadas expande o campo recetivo do modelo sem aumentar o número de parâmetros, permitindo a captura de informações em diferentes escalas de maneira eficiente. Isso é complementado pelas conexões residuais, que ajudam a evitar o problema de desaparecimento de gradientes, melhorando a propagação de informações durante o treino de redes profundas.

No ResUNet-A, a estrutura em forma de "U" do U-Net original é preservada, com caminhos de contração e expansão simétricos conectados por operações de concatenação. Entretanto, ele incorpora camadas de normalização avançadas e funções de ativação otimizadas, que aumentam sua robustez e precisão. A função de perda utilizada frequentemente inclui Focal Loss ou IoU Loss, que são adequadas para lidar com datasets desequilibrados, como aqueles onde as regiões de fogo são pequenas em comparação ao fundo.

Essa arquitetura é particularmente eficiente em tarefas de segmentação semântica devido à sua capacidade de processar imagens de forma end-to-end, gerando máscaras de segmentação altamente detalhadas. O ResUNet-A tem-se destacado em benchmarks que envolvem deteção de padrões complexos, mostrando-se uma solução robusta e

adaptável a diferentes desafios no campo de segmentação.

Após a implementação e avaliação inicial dos modelos U-Net e ResUNet-a, será realizado um segundo estágio de experimentação com o objetivo de explorar melhorias na estrutura das arquiteturas e nas funções de perda utilizadas. Essa etapa procura identificar ajustes que possam aumentar a eficácia dos modelos em termos de precisão e generalização para o problema de segmentação de incêndios

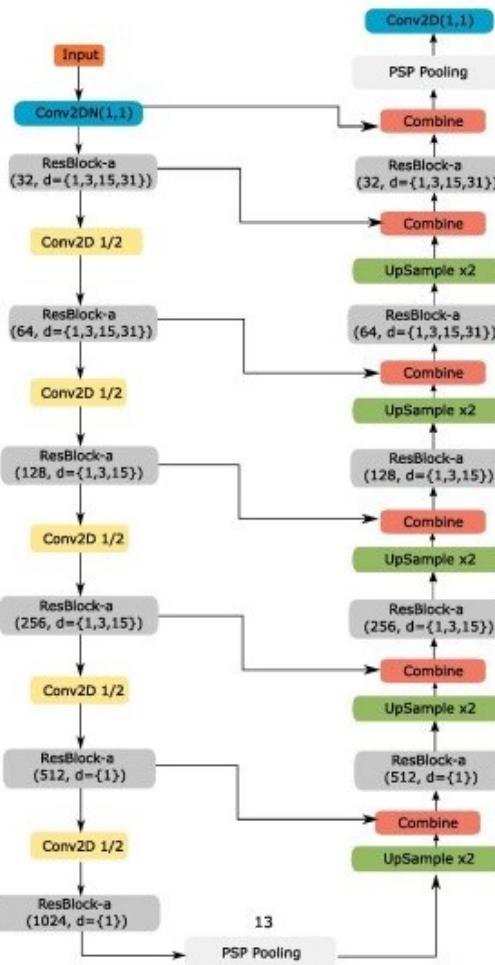


Figure 7: Arquitetura ResUNet-a

3.2 Métricas de avaliação

As métricas utilizadas para avaliar o desempenho dos modelos de segmentação incluem: *global accuracy*, *mean accuracy*, *mean IoU*, *weighted IoU*, e *mean BF score*.

- **Global Accuracy:** Esta métrica fornece uma visão geral sobre a percentagem de pixels corretamente classificados, independentemente da classe a que pertencem. Embora seja amplamente utilizada pela sua simplicidade, torna-se menos fiável em conjuntos de dados altamente desbalanceados, onde a classe dominante pode influenciar desproporcionalmente o resultado final.
- **Mean Accuracy:** Indica a proporção de pixels corretamente classificados con-

siderando todas as classes. É matematicamente definida como:

$$\text{Accuracy} = \frac{\text{TPs}}{\text{TPs} + \text{FNs}} \quad (1)$$

onde TPs representa o número de verdadeiros positivos, isto é, o número de píxeis positivos corretamente classificados, e FN refer-se aos falsos negativos.

- **Mean IoU (Intersection over Union):** Também conhecida como medida de similaridade de Jaccard, é calculada como a média da IoU de todas as classes de dados, considerando todas as imagens. Esta métrica estatística penaliza os falsos positivos (*FPs*) e é definida como:

$$\text{IoU} = \frac{\text{TPs}}{\text{TPs} + \text{FPs} + \text{FNs}} \quad (2)$$

- **Weighted IoU:** Métrica que leva em consideração as classes minoritárias em conjuntos de dados desbalanceados. O valor final é mais realista, sendo calculado como a média ponderada da IoU de cada classe pelo número de píxeis dessa classe.
- **Mean BF Score:** Refere-se ao *Boundary F1 (BF) score*, utilizado para medir o grau de correspondência entre os contornos dos objetos previstos e o *ground truth*. Matematicamente, é definido como:

$$\text{BF} = \frac{\text{TPs}}{\text{TPs} + 0.5 \times (\text{FPs} + \text{FNs})} \quad (3)$$

O *Mean BF Score* é calculado como a média do *BF score* de todas as imagens para uma classe correspondente.

4 - Calendário

O cronograma apresentado na figura 8 detalha as etapas principais do projeto, distribuídas entre os meses de outubro de 2024 e junho de 2025. Este plano de trabalho organiza-se em cinco atividades principais: pesquisa científica, construção do dataset, desenvolvimento do algoritmo U-Net, desenvolvimento do algoritmo ResUNet-a e redação do relatório. A seguir, é apresentada uma descrição detalhada de cada uma das fases.

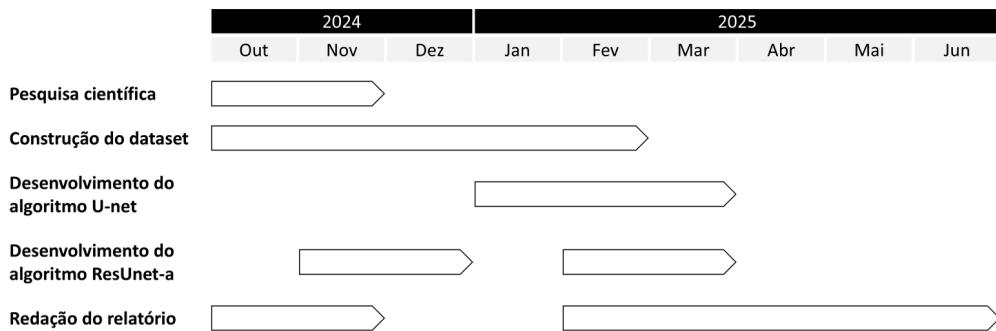


Figure 8: Calendário de atividades do projeto

Pesquisa Científica

Esta etapa inicial foca na revisão de literatura e exploração de metodologias já existentes para a segmentação de incêndios florestais. Serão analisados artigos científicos, algoritmos de aprendizagem profunda (como U-Net e ResUNet-a) e estudos que visam identificar as métricas mais relevantes para avaliar a eficácia de modelos de segmentação. O objetivo principal desta fase é construir uma base teórica sólida para guiar o desenvolvimento técnico do projeto.

Construção do Dataset

Durante este período, será realizada a construção de um dataset personalizado para a segmentação de incêndios florestais, utilizando imagens dos conjuntos de dados Corsican e FLAME. Serão criadas manualmente máscaras detalhadas para identificar os pixels de fogo, fumo e fundo nas imagens, com o objetivo de garantir a precisão e a qualidade dos dados para o treino dos modelos.

Desenvolvimento do Algoritmo U-Net

Uma vez construído o dataset, inicia-se o desenvolvimento do modelo U-Net. Nesta fase, será implementado um modelo de segmentação que utiliza a arquitetura U-Net para identificar e segmentar áreas afetadas por incêndios. O processo inclui o treino do modelo, ajustes nos seus hiperparâmetros e a avaliação. O objetivo desta fase é criar um modelo robusto que consiga segmentar áreas de fogo e fumo com alta fiabilidade.

Desenvolvimento do Algoritmo ResUNet-a

Durante este período, será desenvolvido o algoritmo ResUNet-a, uma versão avançada do modelo de segmentação, projetada para lidar com tarefas complexas, como a seg-

mentação de áreas afetadas por incêndios florestais. O ResUNet-a incorpora técnicas avançadas, como convoluções dilatadas (atrous convolutions) e mecanismos de atenção espacial (spatial attention mechanisms), que permitem ao modelo capturar padrões espaciais e contextuais mais ricos em imagens de alta resolução. O foco será melhorar a precisão da segmentação em condições desafiadoras. Durante esta fase, o modelo será treinado, ajustado e avaliado, com especial atenção aos hiperparâmetros, visando otimizar seu desempenho e garantir sua capacidade de generalização para novos dados.

Redação do Relatório

A redação do relatório ocorrerá ao longo de todo o projeto, com maior intensidade nos meses finais. Esta atividade inclui a documentação detalhada das etapas do projeto, desde a metodologia e construção do dataset até os resultados obtidos nos testes com os modelos. A entrega de um relatório bem estruturado e completo é essencial para consolidar os resultados do projeto e garantir que as conclusões sejam devidamente comunicadas.

Este cronograma foi elaborado para permitir uma progressão lógica das atividades, assegurando que cada etapa seja concluída com a profundidade necessária para alcançar os objetivos estabelecidos no projeto.

5 - Conclusão

A presente dissertação aborda um tema de elevada relevância e impacto global: a deteção precoce e eficiente de incêndios florestais. Este trabalho propõe o desenvolvimento de um sistema avançado baseado em deep learning para a segmentação simultânea de fumo e fogo, explorando arquiteturas de referência como a U-Net e a ResUNet-a. A estrutura do estudo está organizada para assegurar uma abordagem abrangente, incluindo uma revisão do estado da arte, a definição de metodologias robustas e um plano de execução detalhado.

O cronograma do projeto estabelece cinco fases principais, que guiarão os próximos passos do trabalho. Na etapa inicial de pesquisa científica, foi realizada uma análise aprofundada de estudos e algoritmos existentes, com o objetivo de construir uma base teórica sólida. De seguida, o foco é na construção de um dataset personalizado, através da criação manual de máscaras detalhadas para identificar fogo, fumo e fundo, garantindo dados de alta qualidade para treinar os modelos.

As fases de desenvolvimento dos algoritmos serão conduzidas de forma sequencial. O modelo U-Net será implementado e treinado primeiro, utilizando o dataset construído para identificar padrões específicos de fumo e fogo. Posteriormente, o modelo ResUNet-a será desenvolvido, explorando técnicas avançadas para melhorar a precisão da segmentação em condições desafiadoras. Por fim, a redação do relatório consolidará todo o trabalho realizado, garantindo uma documentação completa e clara das metodologias, resultados e conclusões.

Espera-se que este trabalho contribua significativamente para o avanço científico e tecnológico na área de monitorização de incêndios florestais, oferecendo um sistema integrado e eficiente capaz de melhorar a capacidade de resposta a esses desastres.

Bibliografia

- [1] Giuliana Zanchi et al. "Simulation of water and chemical transport of chloride from the forest ecosystem to the stream". In: *Environmental Modelling & Software* (2021). URL: <https://www.sciencedirect.com/science/article/pii/S136481522100027X>.
- [2] Jesús San-Miguel-Ayanz, Jose Manuel Moreno, and Andrea Camia. "Analysis of large fires in European Mediterranean landscapes: Lessons learned and perspectives". In: *Forest Ecology and Management* (2013). URL: <https://www.sciencedirect.com/science/article/pii/S0378112712006561>.
- [3] Tom Toulouse et al. *Computer vision for wildfire research: An evolving image dataset for processing and analysis*. July 07, 2017. URL: <https://www.sciencedirect.com/science/article/abs/pii/S0379711217302114?via%3Dihub> (visited on 11/2024).
- [4] Alireza Shamsoshoara et al. *Aerial imagery pile burn detection using deep learning: The FLAME dataset*. March 23, 2021. URL: <https://www.sciencedirect.com/science/article/pii/S1389128621001201?via%3Dihub> (visited on 11/2024).
- [5] *Image Labeler*. URL: <https://www.mathworks.com/help/vision/ug/get-started-with-the-image-labeler.html>.
- [6] Guan et al. "Forest Fire Segmentation from Aerial Imagery Data Using an Improved Instance Segmentation Model". In: *Remote Sensing* 14.13 (2022). ISSN: 2072-4292. DOI: 10.3390/rs14133159. URL: <https://www.mdpi.com/2072-4292/14/13/3159>.
- [7] *Papers with Code - U-Net Explained*. URL: <https://paperswithcode.com/method/u-net>.
- [8] Songbin Li et al. *Optimized Deep Learning Model for Fire Semantic Segmentation*. April 21, 2022. URL: <https://www.techscience.com/cmc/v72n3/47492/html>.
- [9] Songbin Li et al. *Aerial imagery pile burn detection using deep learning: The FLAME dataset*. March 23, 2021. URL: <https://www.sciencedirect.com/science/article/pii/S1389128621001201>.
- [10] Foivos I. Diakogiannis et al. "ResUNet-a: A deep learning framework for semantic segmentation of remotely sensed data". In: *ISPRS Journal of Photogrammetry and Remote Sensing* (2020). DOI: <https://doi.org/10.1016/j.isprsjprs.2020.01.013>. URL: <https://www.sciencedirect.com/science/article/pii/S0924271620300149>.