



XỬ LÝ ẢNH VÀ ỨNG DỤNG

BÁO CÁO ĐỒ ÁN: HỆ THỐNG PHÁT HIỆN VÀ CẢNH BÁO CHÁY RỪNG SỬ DỤNG THIẾT BỊ BAY KHÔNG NGƯỜI LÁI (DRONE)



Thành viên:
Nguyễn Công Phát - 23521143 ○
Nguyễn Xuân An - 23520023 ○
Trương Hoàng Thành An - 23520032 ○
Vũ Việt Cương - 23520213 ○ 1

NỘI DUNG

I/ TỔNG QUAN ĐỀ TÀI

II/ PHƯƠNG PHÁP THỰC HIỆN

III/ THỰC NGHIỆM

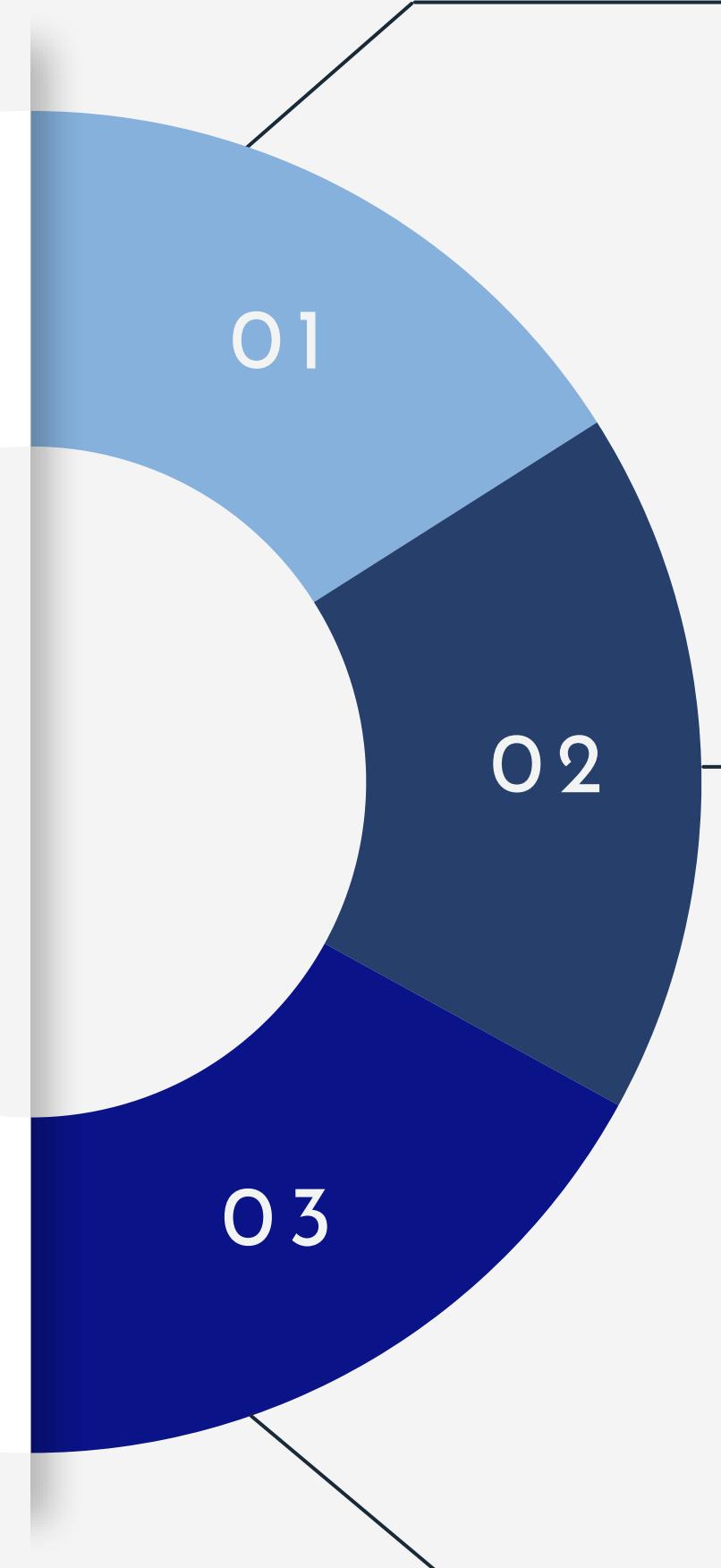
IV/ DEMO

V/ KẾT LUẬN



0000

I/ TỔNG QUAN ĐỀ TÀI



LÝ DO CHỌN ĐỀ TÀI

CÁC NGHIÊN CỨU LIÊN QUAN

PHÁT BIỂU BÀI TOÁN

I.I/LÍ DO CHỌN ĐỀ TÀI

- **Thực trạng:**

- Cháy rừng gia tăng toàn cầu do biến đổi khí hậu, gây thiệt hại lớn về môi trường và con người.
- Phương pháp truyền thống còn nhiều hạn chế: Quan sát mặt đất với tầm nhìn hẹp, khó tiếp cận địa hình hiểm trở; Vệ tinh: độ phân giải thấp, cập nhật chậm → không kịp thời.

- **Mục tiêu:**

- Xây dựng mô hình Deep Learning có khả năng phát hiện chính xác vùng cháy từ luồng video RGB và hồng ngoại thời gian thực thu thập bởi drone, đồng thời giảm thiểu nhầm lẫn với các yếu tố gây nhiễu từ góc nhìn trên cao.
- Thực hiện phân vùng chi tiết đám cháy ở mức pixel, cho phép xác định ranh giới, diện tích và tâm vùng cháy, làm cơ sở cho việc ước lượng vị trí địa lý và mức độ nguy hiểm.
- Phát triển hệ thống xử lý video thời gian thực, tích hợp thông tin không gian (GPS, độ cao, hướng camera) để tự động kích hoạt cảnh báo cháy, dựa trên độ tin cậy dự đoán và tính liên tục theo thời gian.

I.2/CÁC NGHIÊN CỨU LIÊN QUAN

- **Phương pháp truyền thống:** Dựa trên màu sắc, chuyển động, kết cấu
 - Dễ báo động giả, kém ổn định trong môi trường phức tạp
 - Tiêu biểu: Qi et al. (2009) - Video-based fire detection
- **Deep Learning:**
 - Classification: Chỉ xác định có/không có cháy
 - Segmentation (U-Net): Xác định ranh giới cháy ở mức pixel, chính xác hơn
 - Tiêu biểu: Haeri Boroujeni et al. (2024) - Survey AI for UAV wildfire management, Ronneberger et al. (2015) - U-Net segmentation,
- **FLAME Dataset:** Dữ liệu UAV chuẩn với RGB + ground-truth mask, dùng rộng rãi cho phân vùng cháy
 - Tiêu biểu: Miller et al. (2022) - FLAME Dataset, Bouguettaya et al. (2022) - Review UAV wildfire detection
- **State-of-the-art:**
 - YOLO / Faster R-CNN: Nhanh nhưng chỉ bounding box: Liu et al. (2022) - Image-adaptive YOLO or object detection in adverse weather conditions, Redmon & Farhadi (2018) - YOLOv3,
 - Xu hướng mới: RGB + nhiệt & temporal consistency để giảm báo động giả: Zhao et al. (2020) - RGB + saliency for UAV fire detection, Martinez-de Dios et al. (2015) - Multi-UAV wildfire monitoring.

I.3/ PHÁT BIỂU BÀI TOÁN

Input:

- Tập dữ liệu huấn luyện N phần tử: $D = \{(x_i, y_i)\}$ với i chạy từ 1 đến N . Trong đó:
 - $x_i \in \mathbb{R}^H \times W \times C$: Ảnh đầu vào từ UAV (RGB hoặc kết hợp RGB-Thermal).
 - $y_i \in \{0: \text{Non-fire}, 1: \text{Fire}\}^H \times W$: Mặt nạ phân vùng (segmentation mask) ở mức pixel.
- Ngoài ra, với mỗi khung hình video tại thời điểm t , hệ thống nhận thêm metadata đồng bộ:
 - **GPS**: ($\text{lat}_t, \text{lon}_t, \text{alt}_t$)
 - **Hướng camera**: ($\text{yaw}_t, \text{pitch}_t, \text{roll}_t$)

Output:

- $\hat{y}_i \in [0, 1]^H \times W$: Mặt nạ phân vùng cháy được dự đoán bởi mô hình Deep Learning (UNet).
- $\hat{c}_i = (\text{lat}_i, \text{lon}_i)$: Tọa độ địa lý ước lượng của vùng cháy, được xác định từ **tâm (centroid)** của vùng phân đoạn.
- $a_i \in \{0, 1\}$: Tín hiệu cảnh báo cháy thời gian thực, được kích hoạt khi vùng cháy thỏa mãn các điều kiện về **diện tích và độ tin cậy theo thời gian**.

I.3/ PHÁT BIỂU BÀI TOÁN (tt.)

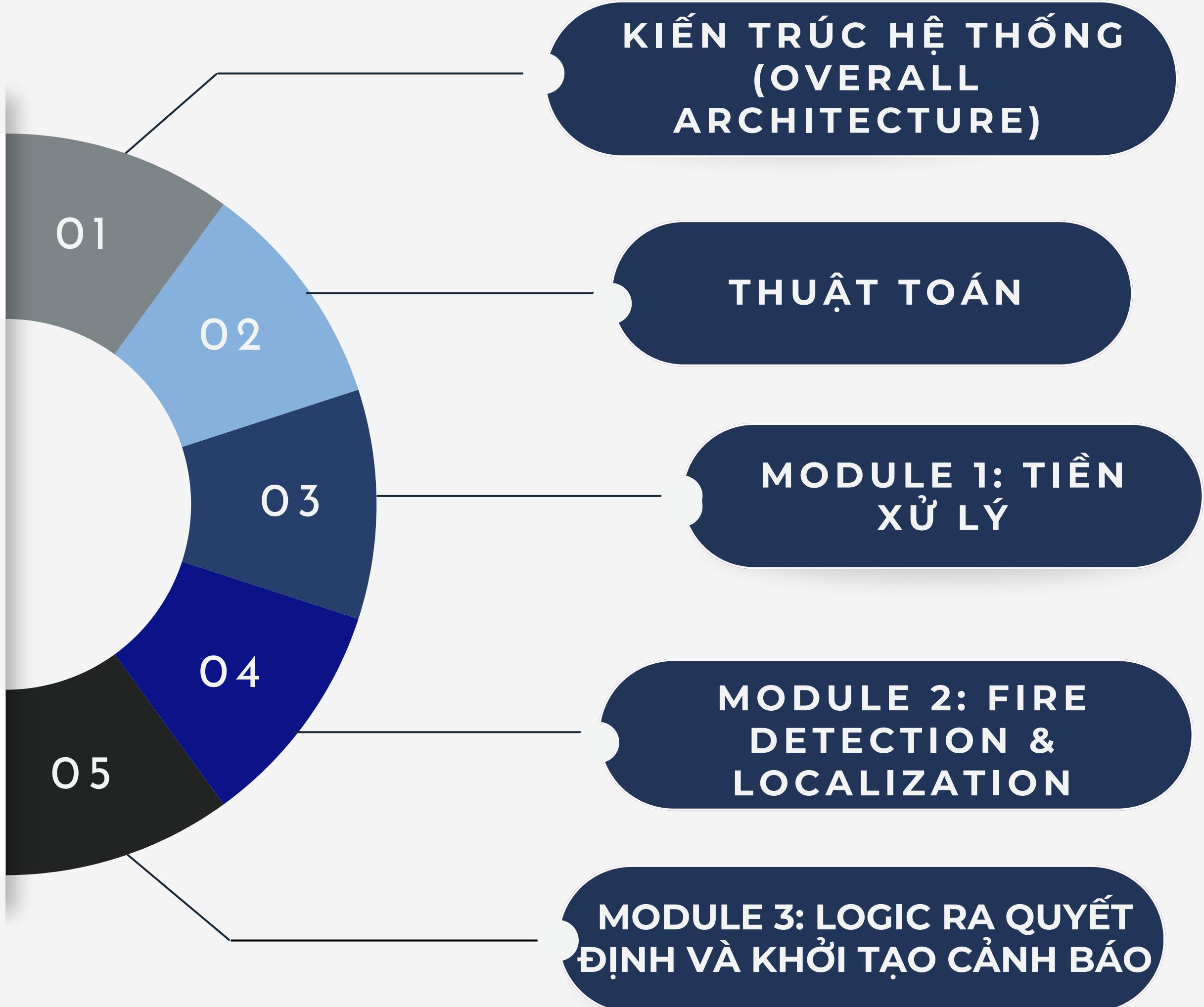
Ràng buộc hệ thống (Constraints):

- **Xử lý thời gian thực:** Luồng video đầu vào phải được xử lý theo thời gian thực, sử dụng định dạng mã hóa **H.264** (.MP4 hoặc .MOV), với tốc độ **24-30 FPS** và độ phân giải tối thiểu **512 * 512 pixel**.
- **Giới hạn camera UAV:** Góc nhìn camera (Field of View - FOV) nằm trong khoảng **60°-90°**.
- **Điều kiện bay:**
 - Độ cao bay trong khoảng **50-250 m**.
 - Tốc độ bay duy trì trong khoảng **15-25 km/h**.
- **Giới hạn không gian:** Hoạt động định vị GPS bị giới hạn trong các **khu vực rừng được xác định trước**.

Yêu cầu hệ thống (Requirements):

- **Hiển thị trực quan:** Vùng cháy phải được làm nổi bật rõ ràng trên luồng video để hỗ trợ quan sát và đánh giá nhanh.
- **Độ chính xác định vị:** Vị trí cháy ước lượng phải đủ chính xác để hỗ trợ phản ứng khẩn cấp kịp thời.
- **Cảnh báo kịp thời:** Tín hiệu cảnh báo cháy phải được **phát ra ngay sau khi phát hiện**, nhằm giảm thiểu độ trễ phản ứng.
- **Ưu tiên phát hiện đầy đủ:** Hệ thống cần **ưu tiên recall cao** để tránh bỏ sót các sự kiện cháy thực, đồng thời **chấp nhận một tỷ lệ báo động giả giới hạn**.

II/ PHƯƠNG PHÁP THỰC HIỆN



II.1/ OVERALL ARCHITECTURE

The proposed system is architected as a multi-stage image processing pipeline optimized for real-time UAV-based wildfire surveillance, comprising three primary functional modules: Data Preprocessing, Fire Detection & Localization, and Alert Generation.

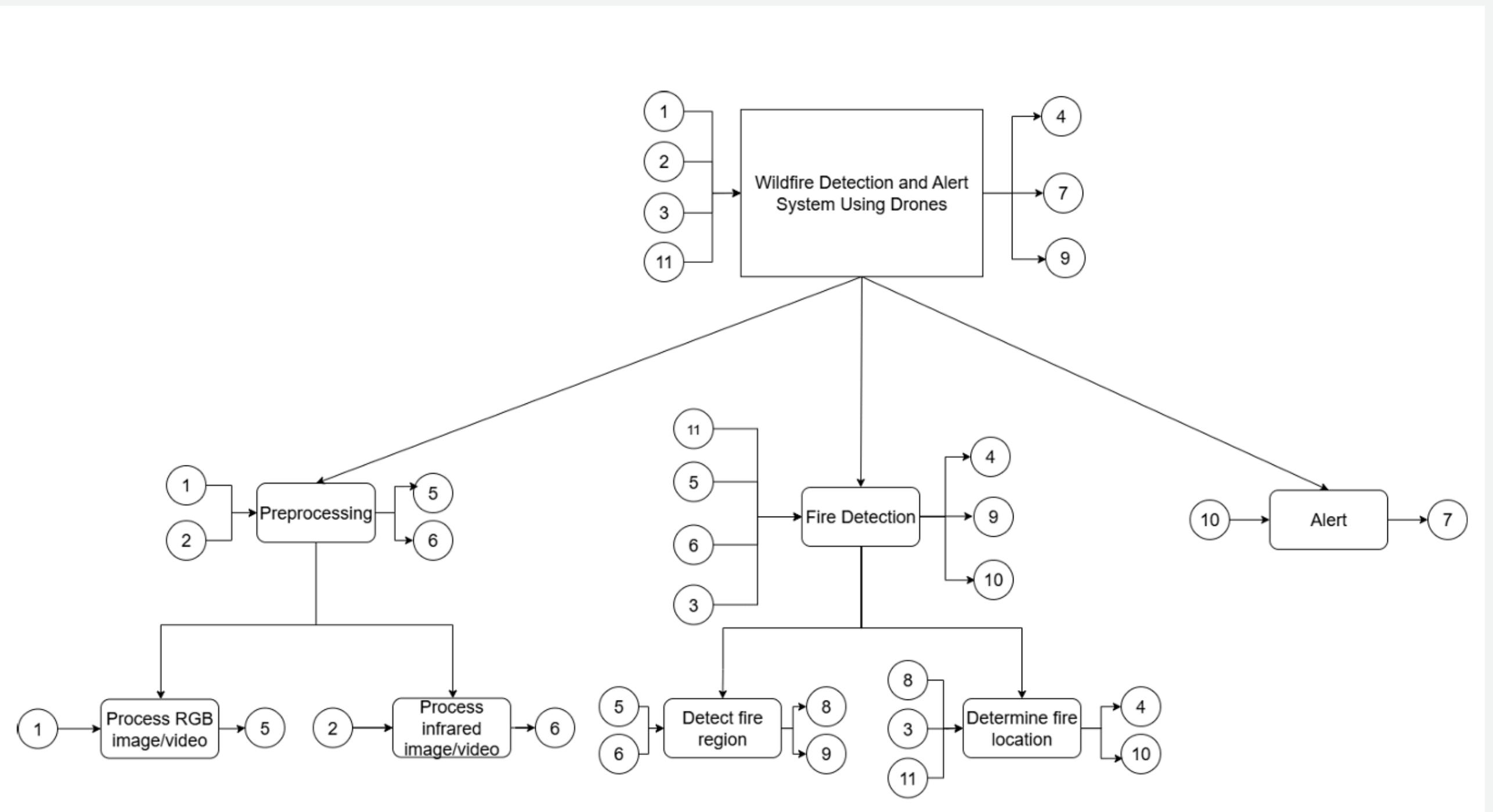
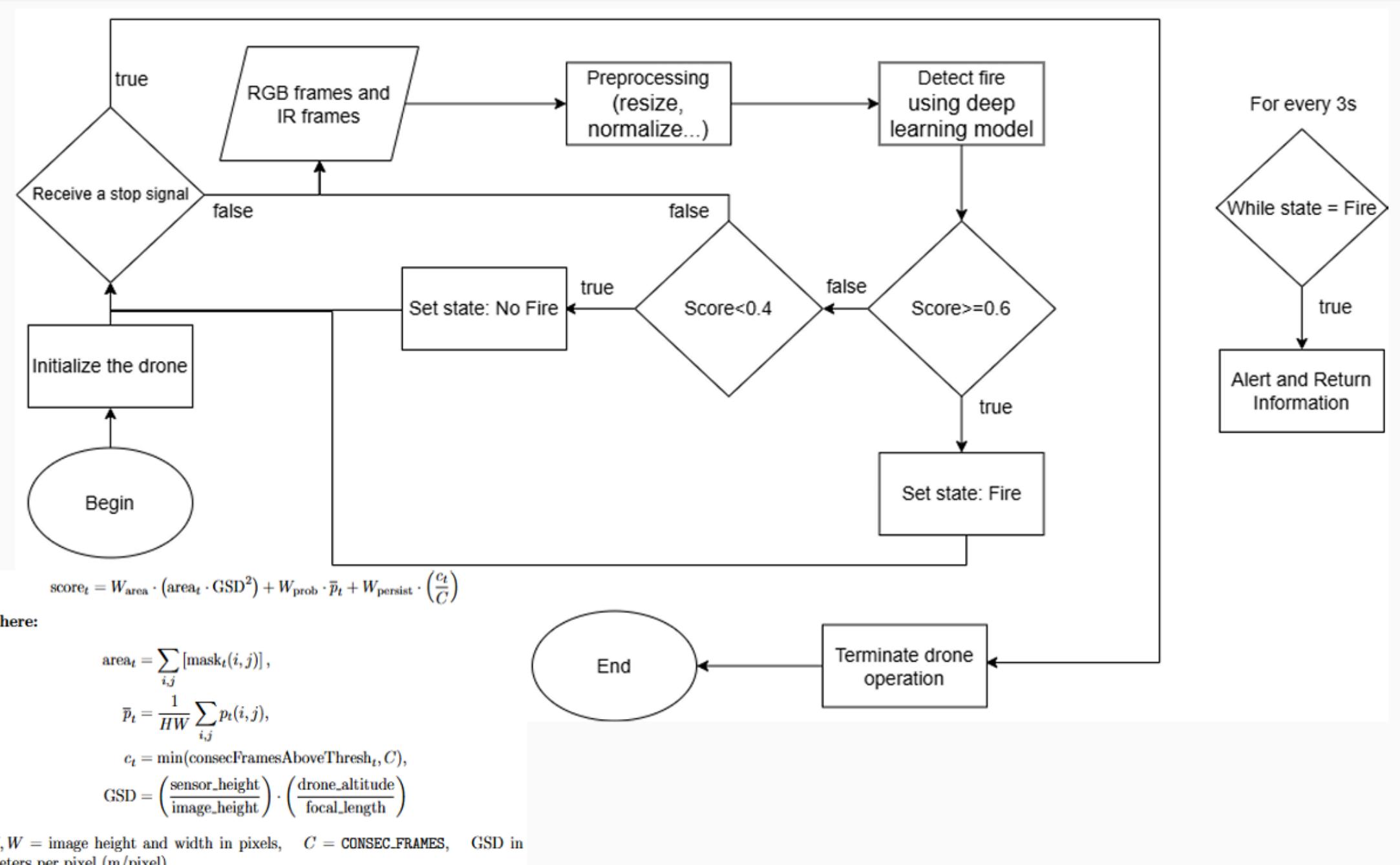


Figure 1. Overall Architecture

II.2/ THUẬT TOÁN HỆ THỐNG (SYSTEM ALGORITHM)

Luồng xử lý và Cơ chế ngưỡng kép (Hysteresis)



Hình 2. Thuật toán hệ thống

Kiến trúc luồng đôi (Dual-Loop Architecture):

- **Detection Loop:** Chạy tần suất cao, phân tích hình ảnh thời gian thực.
- **Alert Loop:** Chạy song song, độ trễ thấp, chuyên trách việc ra quyết định và gửi cảnh báo.

Cơ chế ổn định Hysteresis (Ngưỡng kép):

- Thay vì dùng một ngưỡng cố định (0.5) dễ gây nhiễu/nhấp nháy(flickering).
- **Ngưỡng bật (FIRE):** Score ≥ 0.6
- **Ngưỡng tắt (NO_FIRE):** Score < 0.4
- Giữ nguyên trạng thái: Khi Score nằm trong khoảng [0.4, 0.6]

Lợi ích: Loại bỏ các báo động giả nhất thời, đảm bảo hệ thống chỉ báo động khi có sự thay đổi rõ rệt và ổn định.

III.2/ THUẬT TOÁN HỆ THỐNG (tt.)

```
1 # System Initialization
2 global_state = NO_FIRE
3 # Execute the alert loop in a concurrent thread to ensure non-blocking operation
4 start_thread(alert_loop)
5
6 # 1. Main Detection Loop (High-frequency inference)
7 while True:
8     # Acquire synchronized RGB and Infrared frames from UAV sensors
9     rgb, ir = capture_frames()
10    # Generate wildfire probability score using the U-Net model
11    score = model.predict(rgb, ir)
12
13    # State Management using Hysteresis Thresholding to filter signal noise
14    if score >= 0.6:
15        global_state = FIRE
16    elif score < 0.4:
17        global_state = NO_FIRE
18    # Implicit: maintain current state if score is in range [0.4, 0.6]
19
20 # 2. Alert Generation Loop (Runs concurrently every 3 seconds)
21 def alert_loop():
22     while True:
23         if global_state == FIRE:
24             # Calculate absolute GPS coordinates using pinhole camera geometry
25             location = estimate_geospatial_coords()
26             # Dispatch telemetry package (GPS, Overlay, Timestamp) to command center
27             send_alert(location, visual_overlay)
28             # Sleep for 3 seconds to prevent alert fatigue and network congestion
29             sleep(3.0)
30         else:
31             # Low-latency polling when no fire is detected to save system resources
32             sleep(0.1)
```

Hình 3. Mã giả thuật toán (System Pseudocode)

II.3/ MODULE 1: TIỀN XỬ LÝ DỮ LIỆU

1. Mục tiêu cốt lõi (Core Objectives)

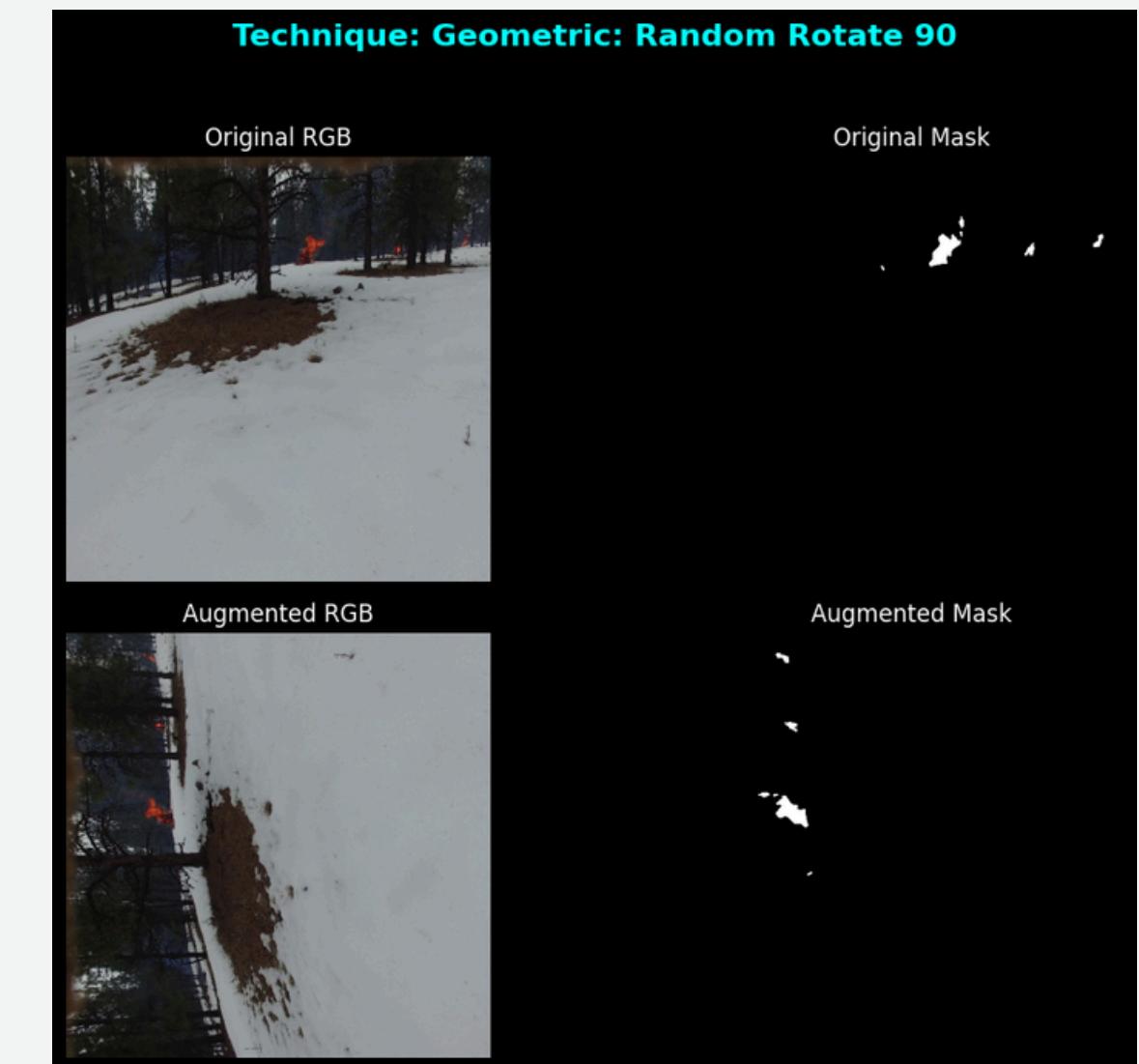
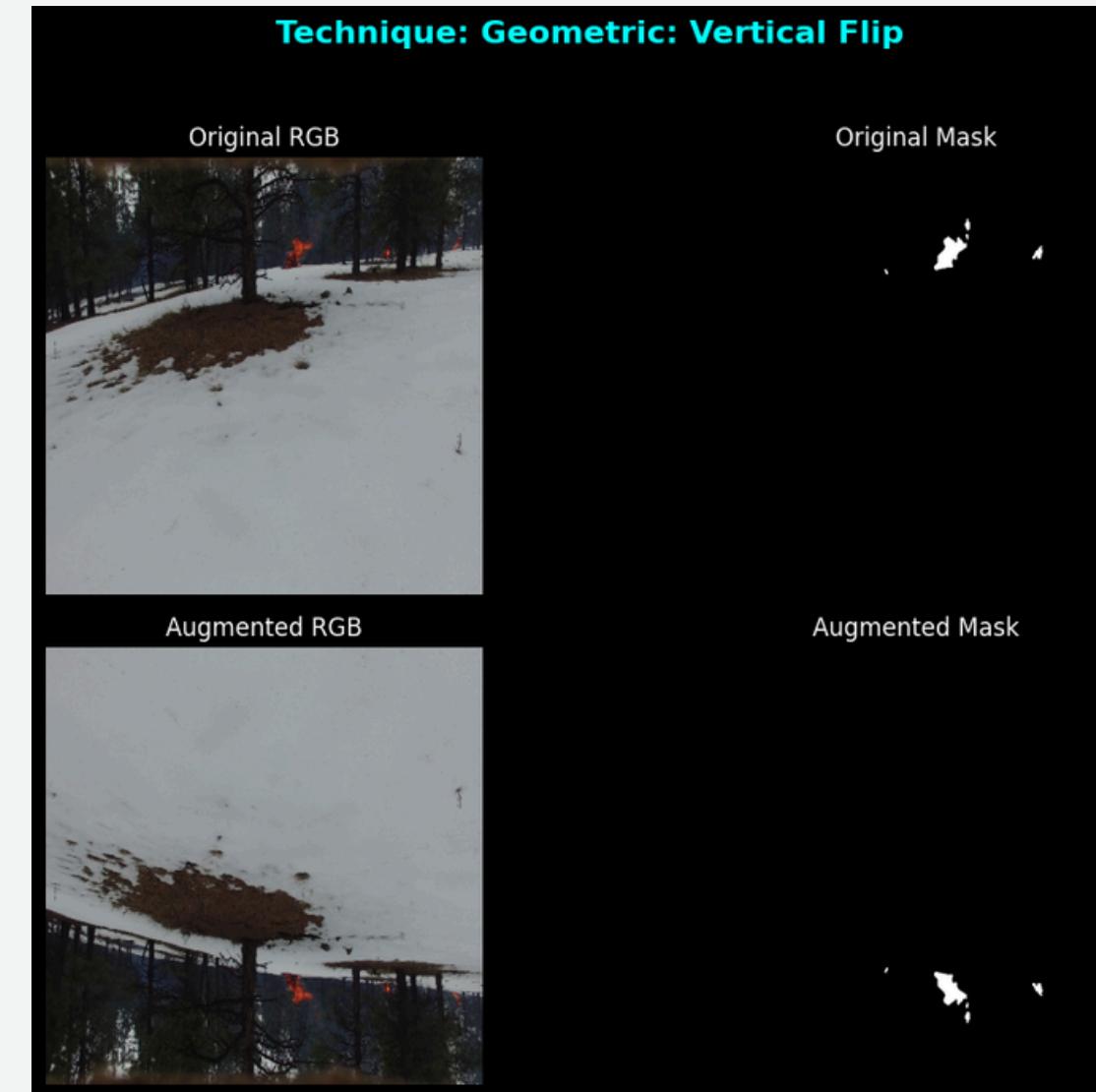
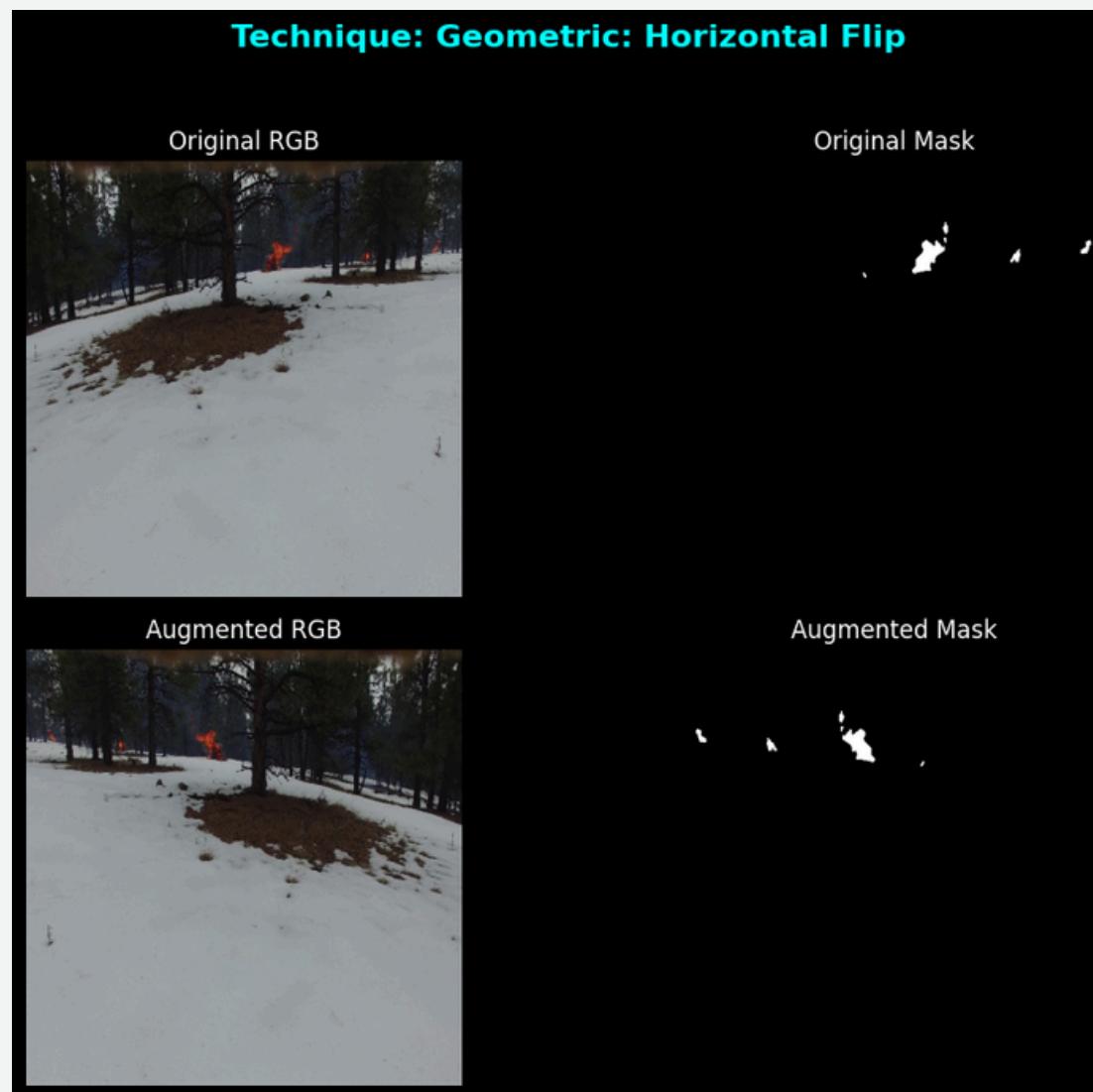
- Chuẩn hóa (Standardization):** Thay đổi kích thước về 512×512 và chuẩn hóa Pixel $[0, 1]$ để ổn định quá trình huấn luyện.
- Đồng bộ hóa (Synchronization):** Căn chỉnh khớp các điểm ảnh giữa cảm biến màu (RGB) và cảm biến nhiệt (Infrared/LWIR).
- Tính bền bỉ (Robustness):** Sử dụng thư viện Albuementations để chuẩn bị cho mô hình đối phó với các điều kiện thực tế không lường trước.

2. Đóng góp chính (Key Contributions)

- Giảm thiểu nhiễu:** Loại bỏ các lỗi kỹ thuật từ cảm biến và hiện tượng mù khí quyển.
- Chuẩn hóa đặc trưng:** Chuyển đổi các tín hiệu thô thành các Tensor (mảng dữ liệu) chất lượng cao và đồng bộ.
- Tăng cường khả năng tổng quát hóa:** Đảm bảo chỉ số về độ đo mô hình (IoU và Recall) duy trì ở mức cao bất kể địa hình hay điều kiện ánh sáng.

II.3/ MODULE 1: TIỀN XỬ LÝ DỮ LIỆU (tt.)

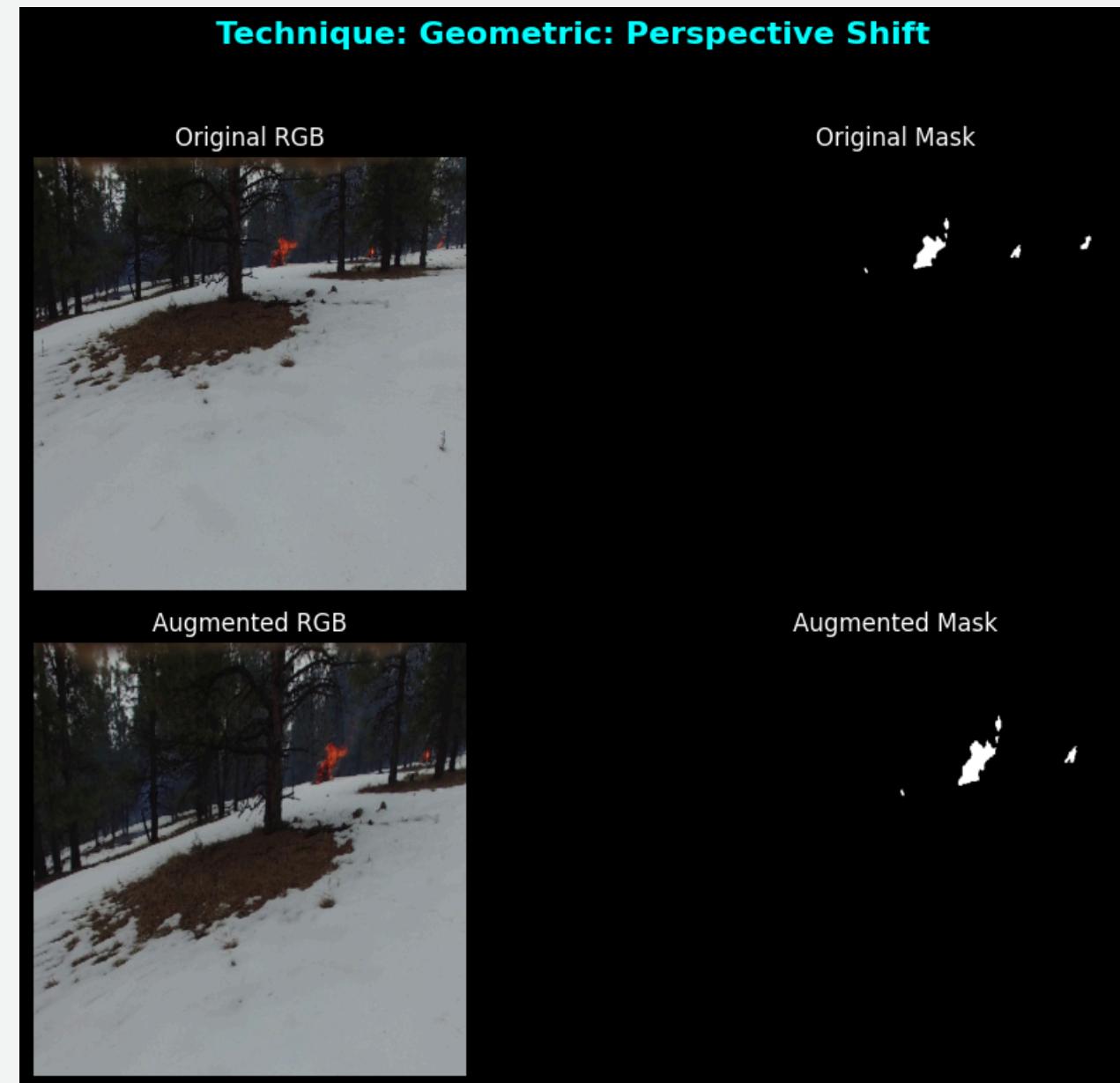
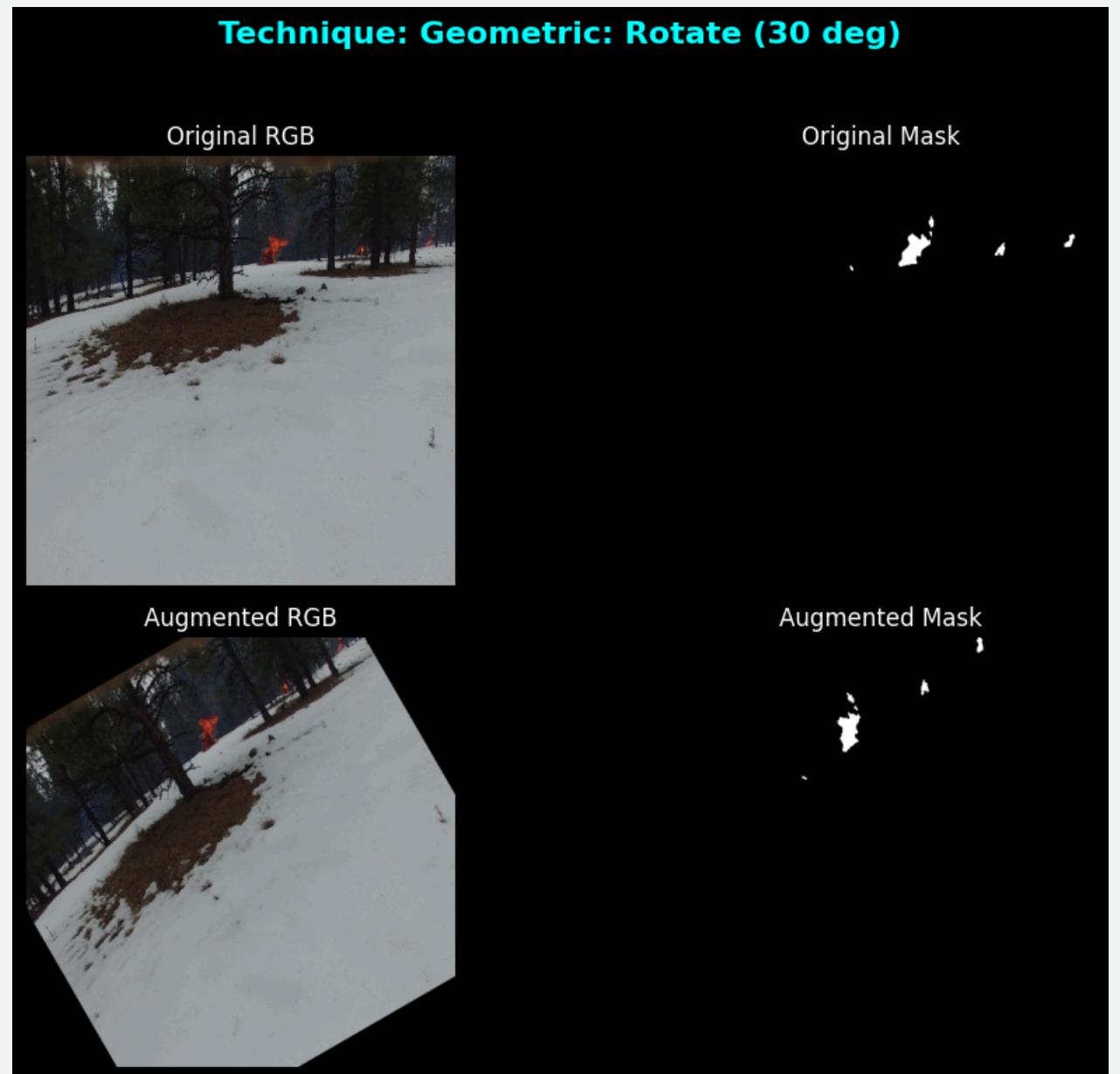
3. CÁC CHIẾN LƯỢC TĂNG CƯỜNG DỮ LIỆU



Horizontal & Vertical Flip / RandomRotate90: Giả lập việc Drone thay đổi hướng bay (Bắc-Nam, Đông-Tây) mà vẫn nhận diện được cấu trúc lửa.

II.3/ MODULE 1: TIỀN XỬ LÝ DỮ LIỆU (tt.)

3. CÁC CHIẾN LƯỢC TĂNG CƯỜNG DỮ LIỆU (tt.)

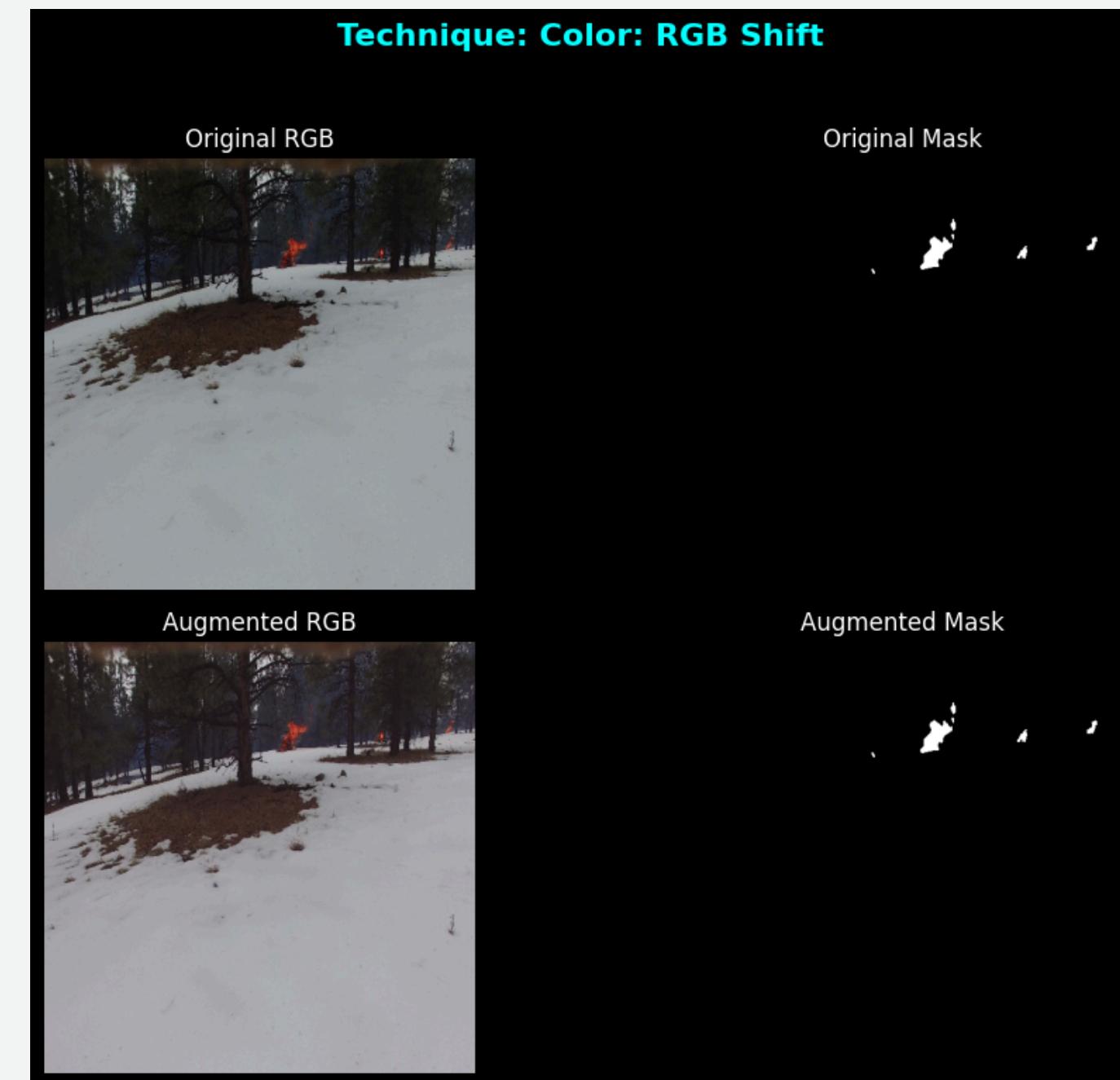
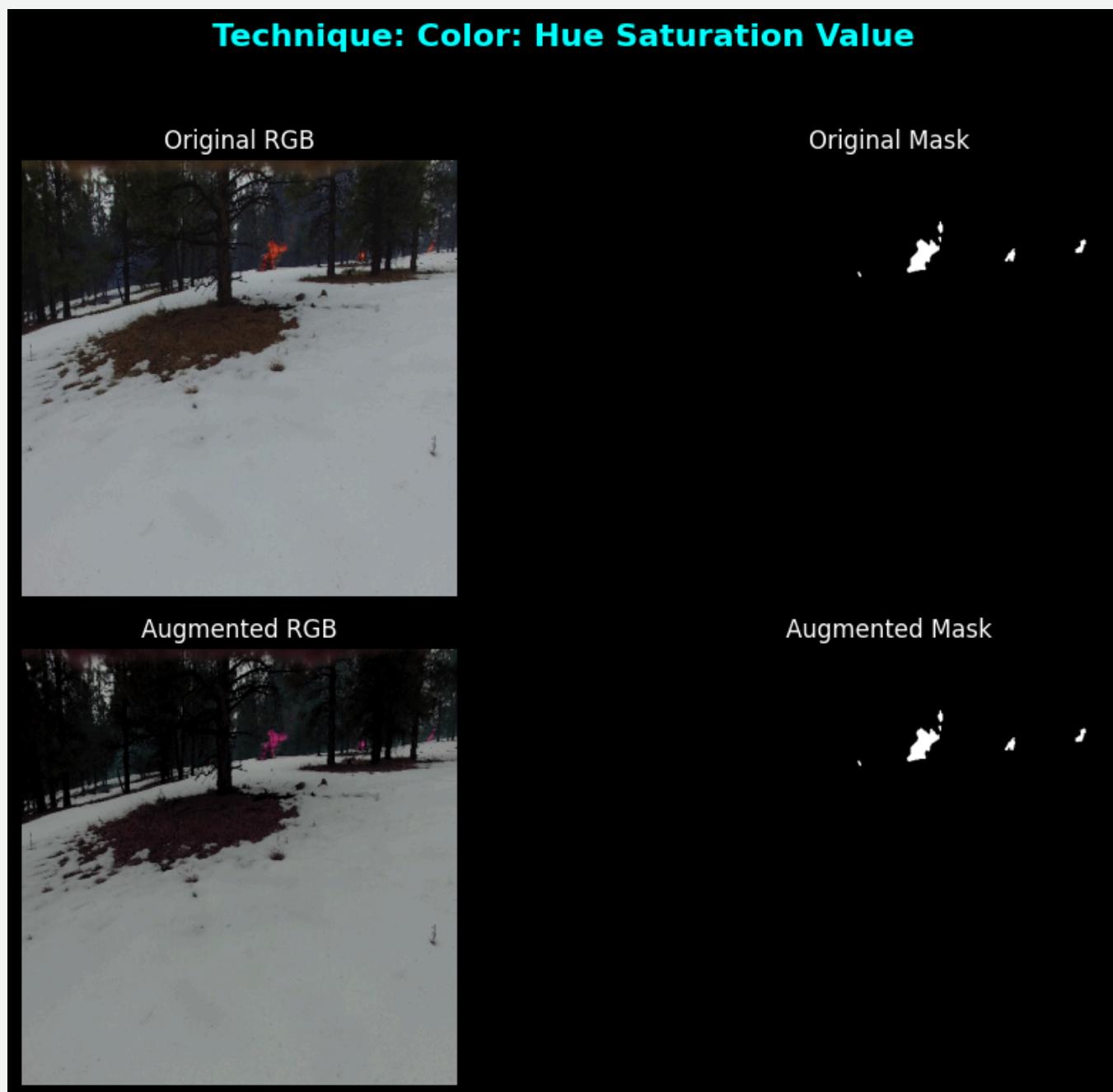


Rotate (30°): Mô phỏng các pha nghiêng mình (Roll) của Drone khi chịu tác động của gió hoặc khi chuyển hướng đột ngột.

Perspective: Thay đổi góc nhìn (Pitch/Yaw) của camera, giúp mô hình học được các biến dạng hình học của đám cháy từ các góc nhìn chéo.

II.3/ MODULE 1: TIỀN XỬ LÝ DỮ LIỆU (tt.)

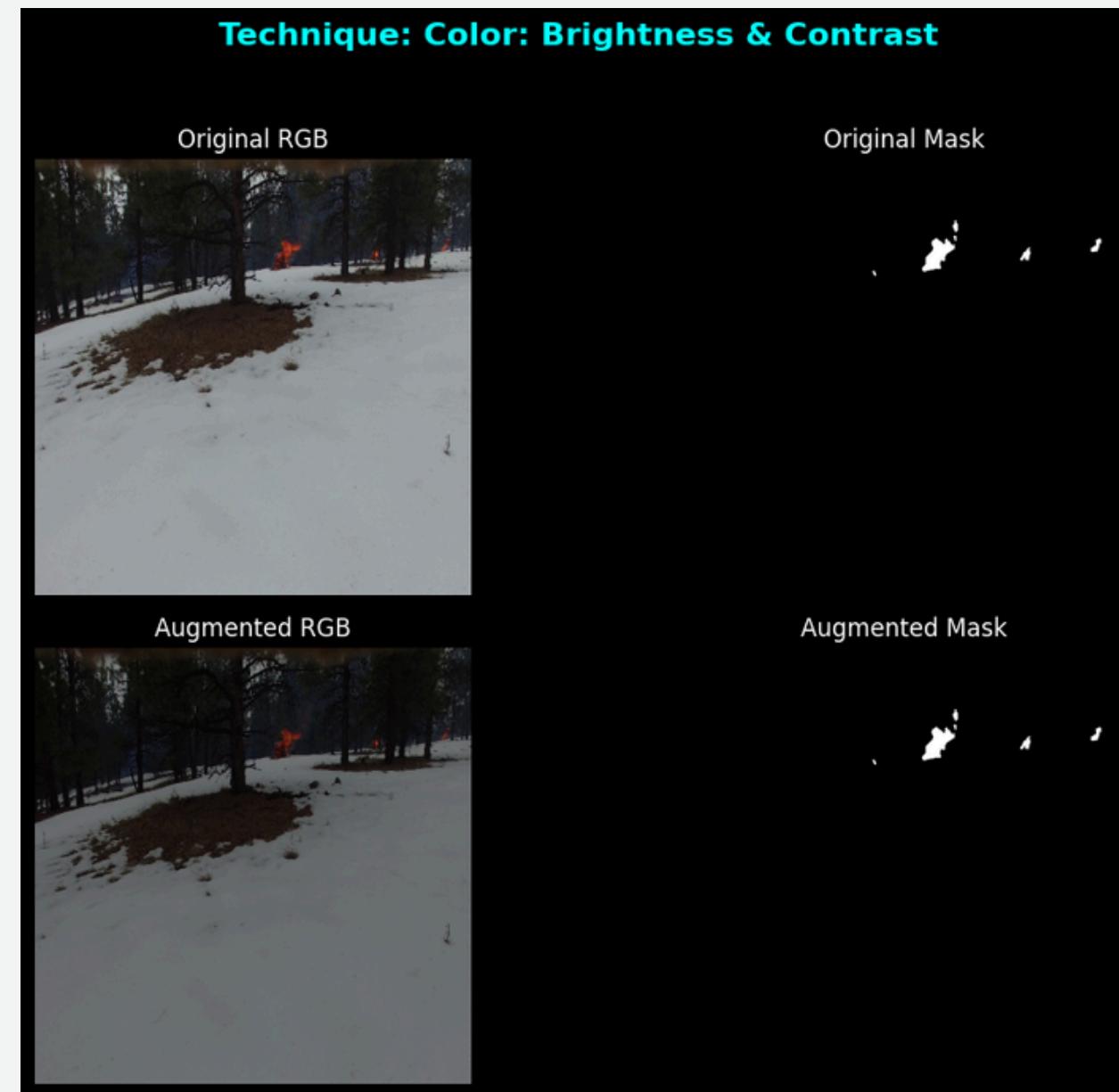
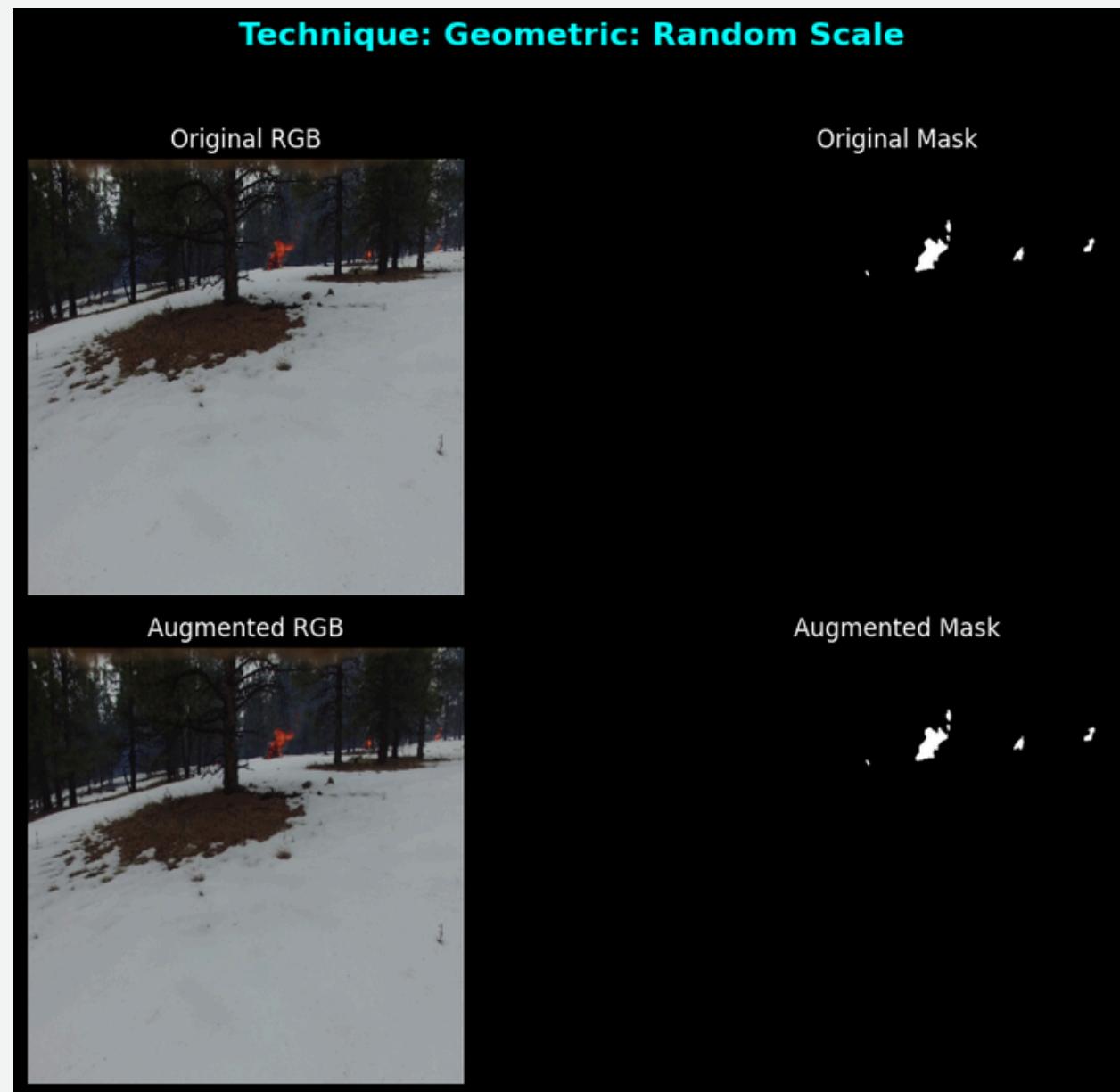
3. CÁC CHIẾN LƯỢC TĂNG CƯỜNG DỮ LIỆU (tt.)



HueSaturationValue & RGBShift: Xử lý sự sai lệch màu sắc giữa các loại camera Drone khác nhau hoặc do hiện tượng cân bằng trắng tự động bị lỗi trong môi trường rừng.

II.3/ MODULE 1: TIỀN XỬ LÝ DỮ LIỆU (tt.)

3. CÁC CHIẾN LƯỢC TĂNG CƯỜNG DỮ LIỆU (tt.)

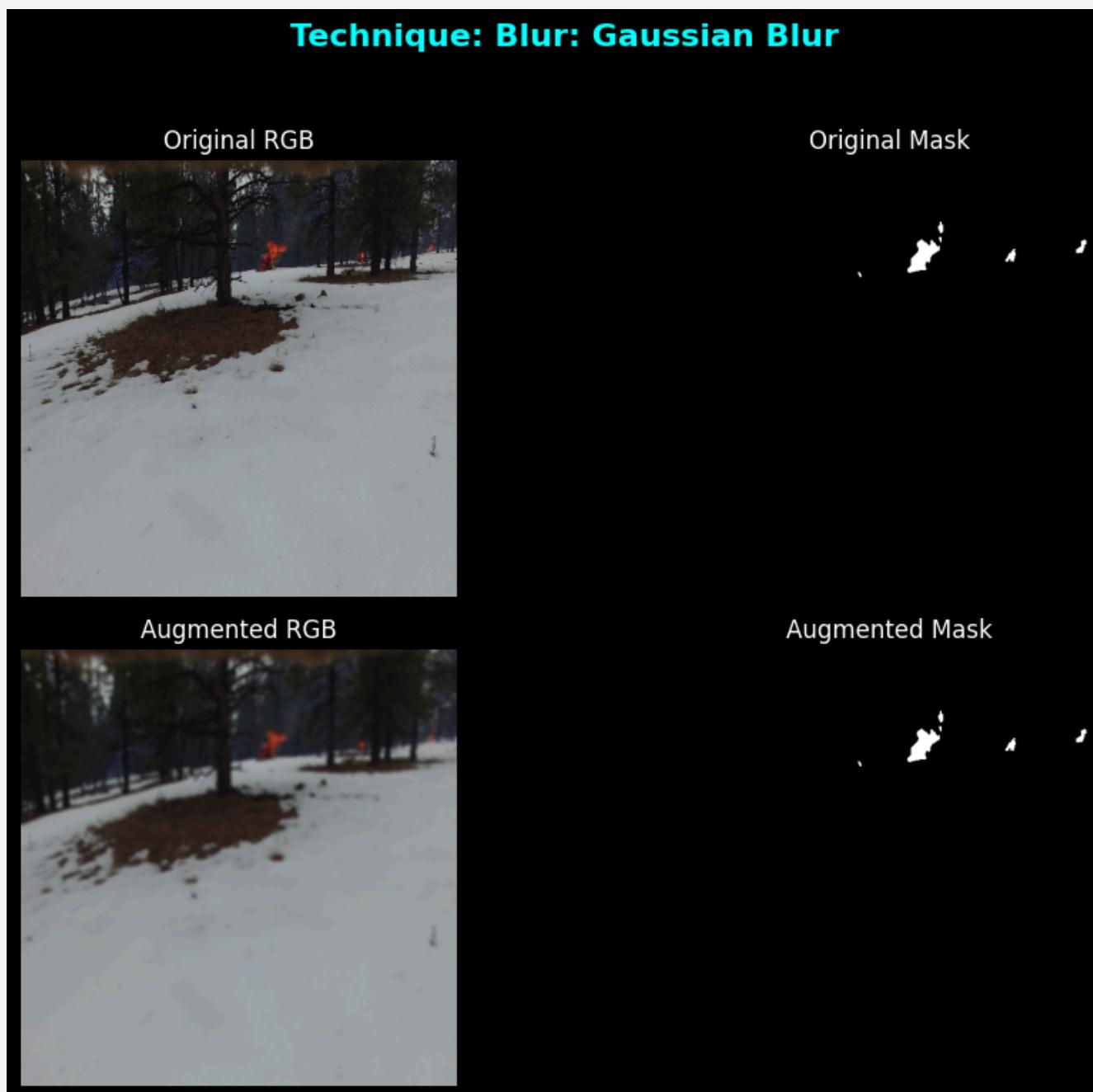


RandomScale: Giả lập sự thay đổi độ cao của Drone (từ 50m đến 250m). Đám cháy sẽ trông to hoặc nhỏ hơn tùy vào cao độ.

RandomBrightnessContrast: Đối phó với sự thay đổi cường độ ánh sáng mặt trời tùy theo thời điểm trong ngày (sáng sớm, trưa nắng gắt).

II.3/ MODULE 1: TIỀN XỬ LÝ DỮ LIỆU (tt.)

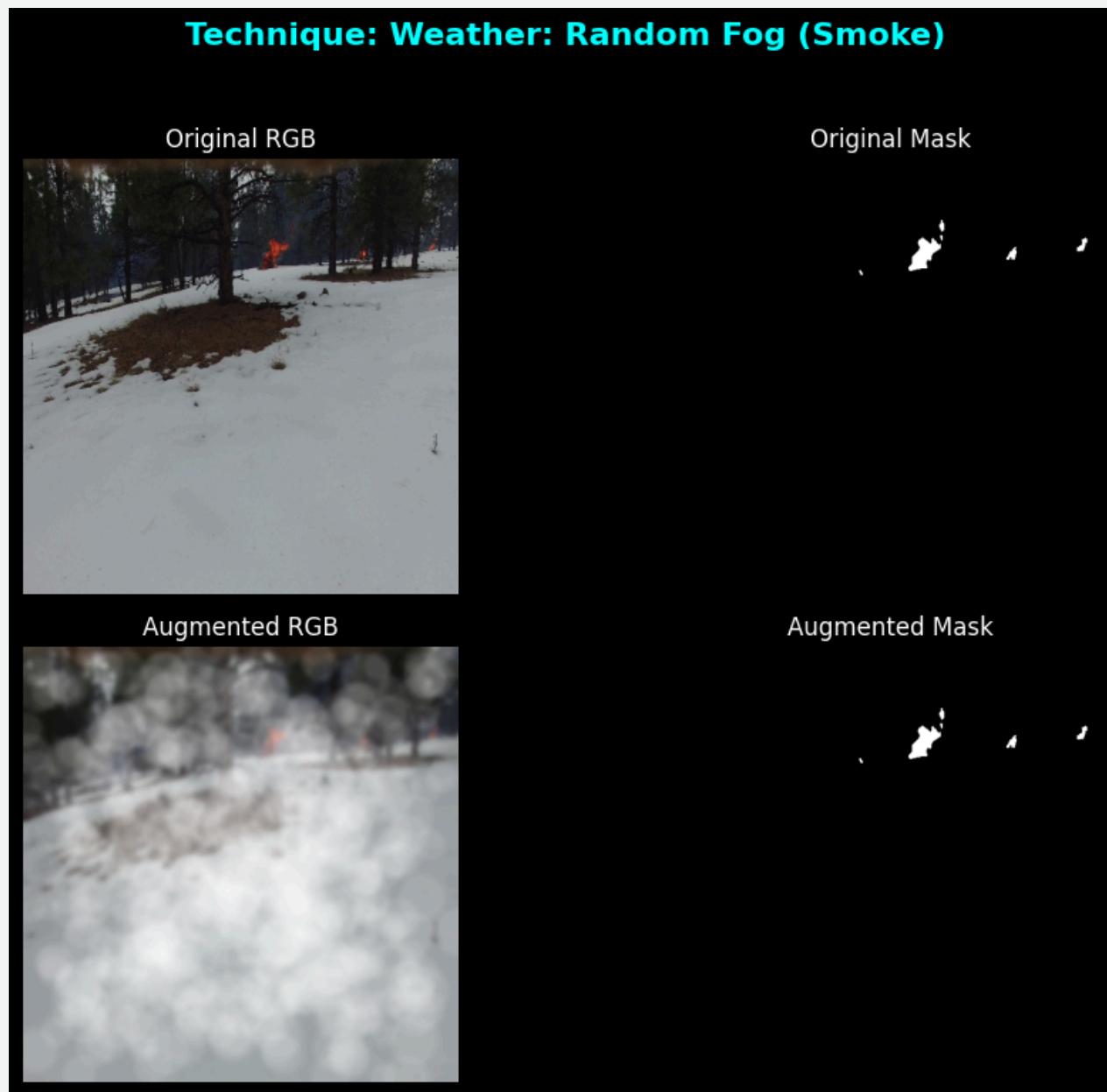
3. CÁC CHIẾN LƯỢC TĂNG CƯỜNG DỮ LIỆU (tt.)



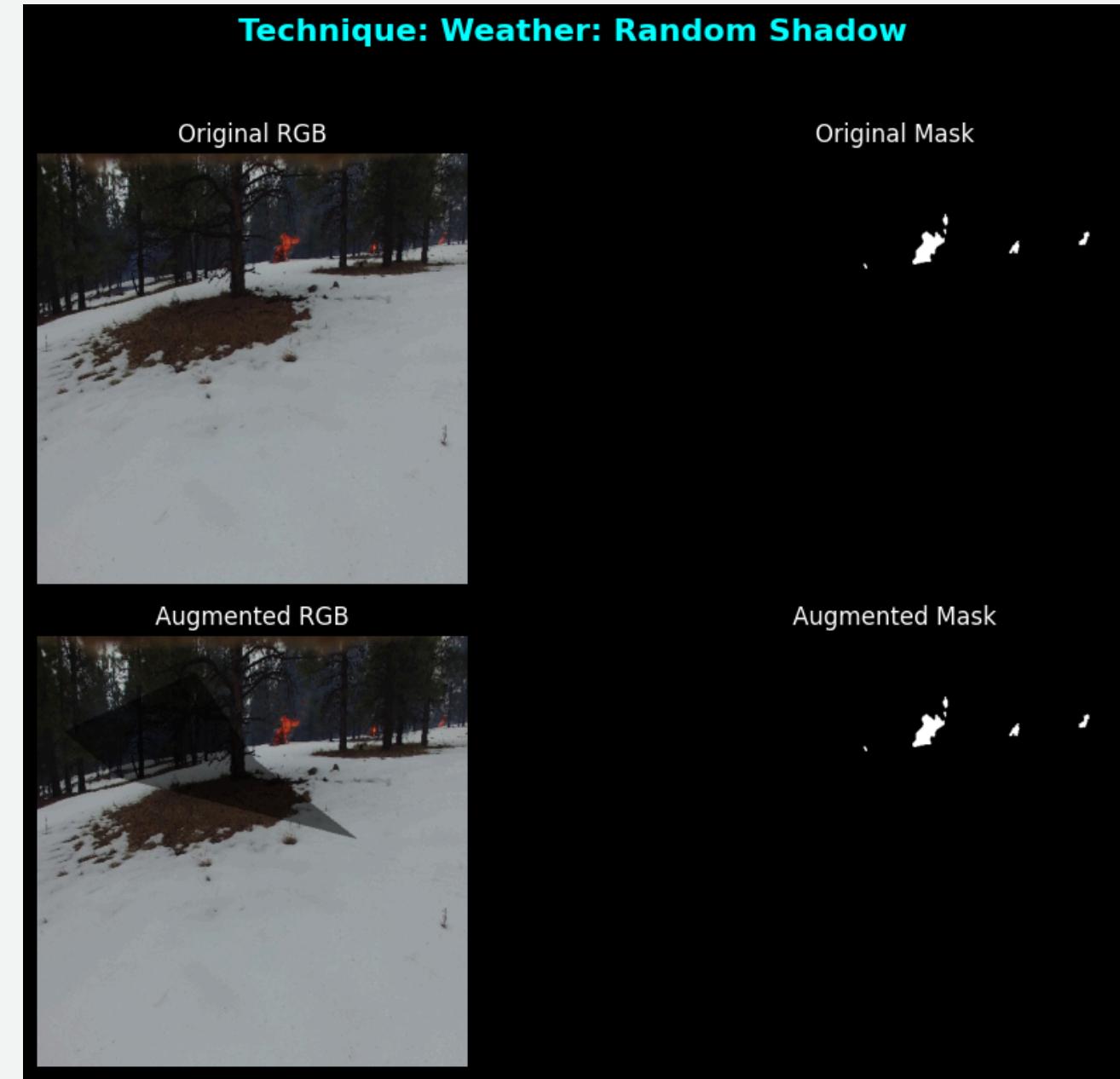
GaussianBlur: Giả lập hiện tượng nhòe hình do Drone bị rung lắc (vibration) hoặc camera bị mất nét tạm thời khi bay nhanh.

II.3/ MODULE 1: TIỀN XỬ LÝ DỮ LIỆU (tt.)

3. CÁC CHIẾN LƯỢC TĂNG CƯỜNG DỮ LIỆU (tt.)

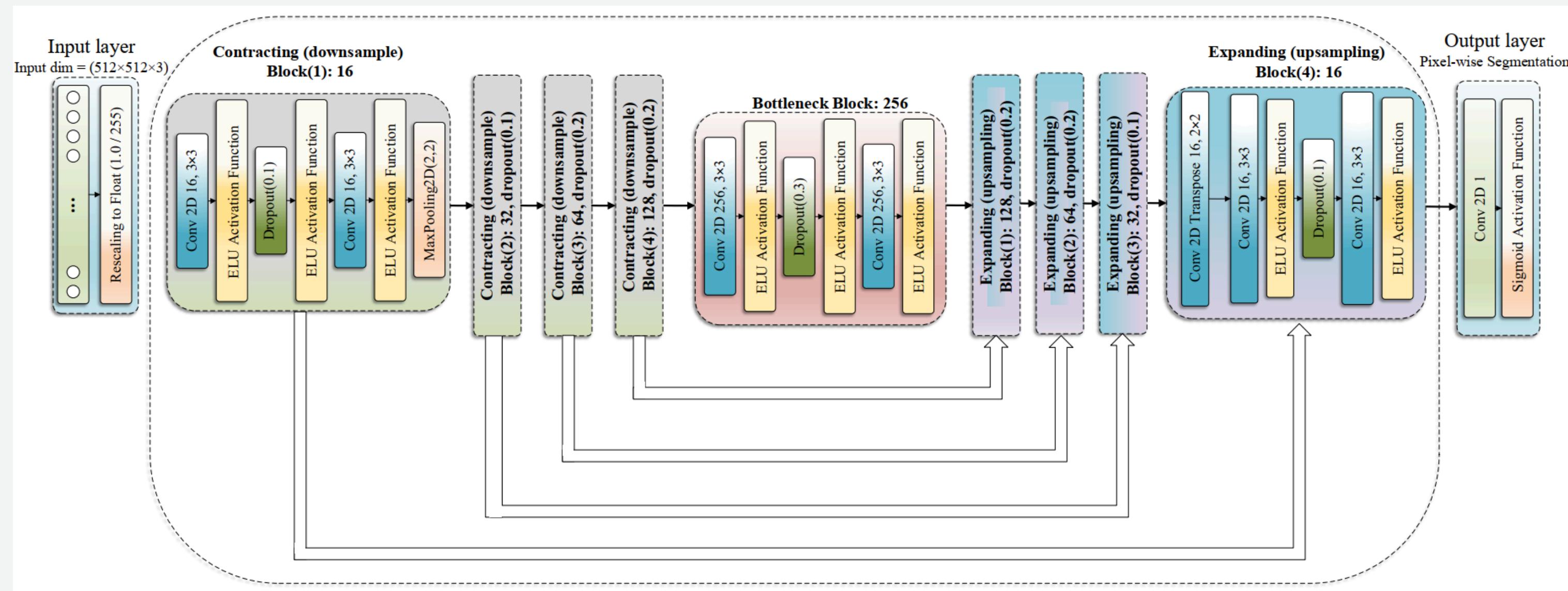


RandomFog: Kỹ thuật then chốt giúp mô hình học cách phân biệt vùng cháy xuyên qua các lớp khói dày đặc hoặc sương mù rừng núi.



RandomShadow: Tạo ra các vùng tối giả lập (bóng mây, bóng cây cổ thụ). Điều này cực kỳ quan trọng để hệ thống không bị báo động nhầm (giảm False Positive) khi gặp các vùng tối có hình dạng phức tạp.

II.4/ MODULE 2: FIRE DETECTION AND LOCALIZATION



Hình 4. Kiến trúc mô hình U-Net & ResNet-34

Tại sao chọn U-Net?

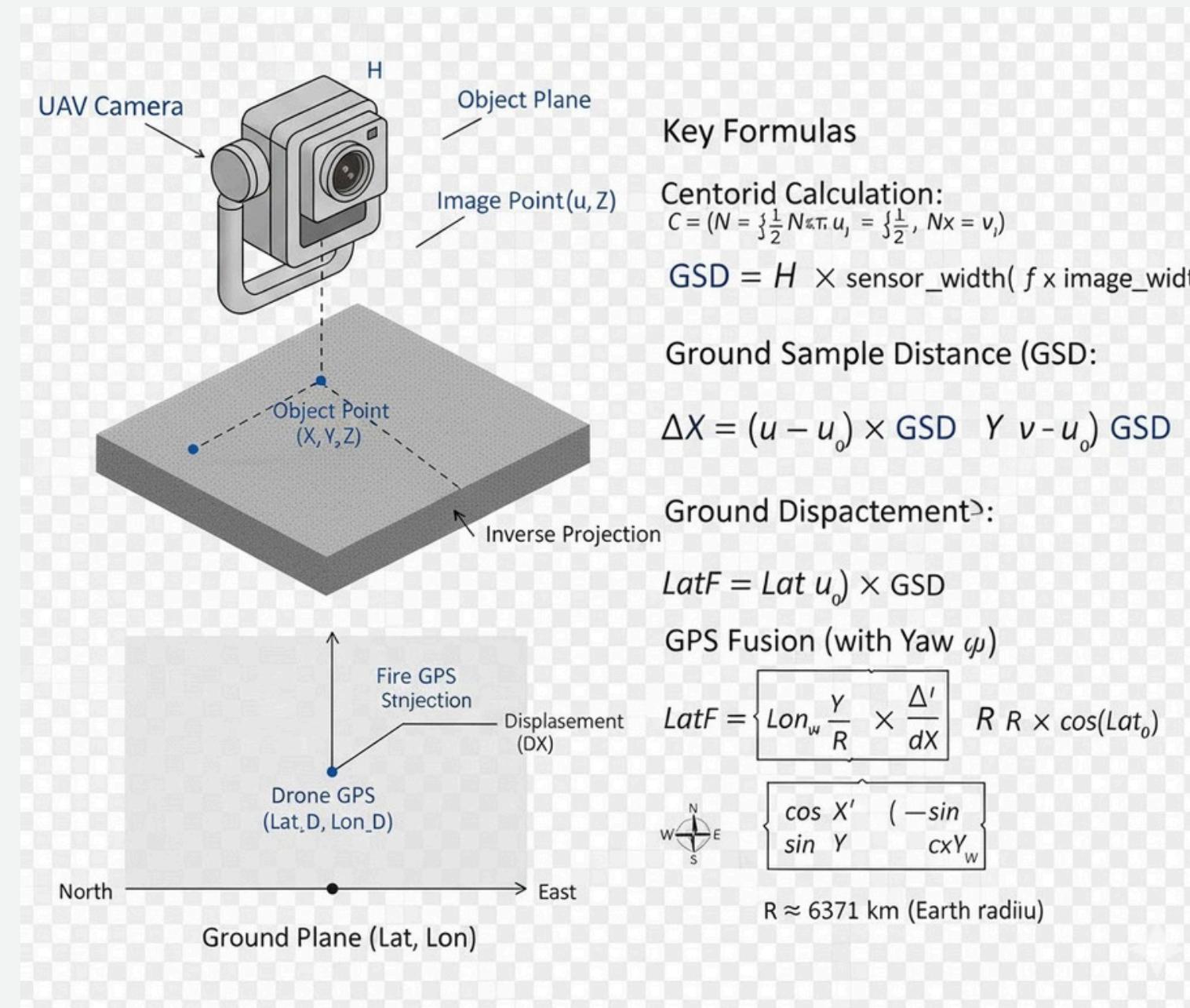
- Phát hiện các đám cháy nhỏ, phân tán tốt hơn các mạng FCN thông thường.
- **Skip Connections:** Giúp khôi phục chi tiết các biên giới đám cháy phức tạp bị mất đi trong quá trình nén ảnh.

Cấu trúc đối xứng (Encoder-Decoder):

- **Encoder (Nhánh co bóp):** Sử dụng ResNet-34 backbone để trích xuất đặc trưng ngữ nghĩa mạnh mẽ nhưng vẫn đảm bảo tốc độ xử lý real-time trên thiết bị nhúng.
- **Decoder (Nhánh mở rộng):** Khôi phục độ phân giải không gian để đưa ra mặt nạ phân vùng mức pixel.

II.4/ MODULE 2: FIRE DETECTION AND LOCALIZATION (tt.)

TOÁN HỌC ĐỊNH VỊ TỌA ĐỘ GPS (GEOSPATIAL LOCALIZATION)



Hình 5. Mô hình Camera Pinhole: Chuyển đổi tọa độ pixel trên ảnh thành tọa độ địa lý.

The mathematical framework is executed through the following stages:

- **Centroid Calculation:** The system first identifies the geometric center of the detected wildfire region. The centroid $C(u, v)$ is calculated by averaging the coordinates of all pixels classified as fire (N):

$$C = \left(\frac{1}{N} \sum_{i=1}^N u_i, \frac{1}{N} \sum_{i=1}^N v_i \right) \quad (4.2)$$

- **Ground Mapping via Ground Sample Distance (GSD):** As illustrated in Figure 4.4, the pixel displacement is converted into metric distance on the ground. The **GSD**, representing the real-world size of a single pixel, is defined as:

$$GSD = \frac{H \times \text{sensor_width}}{\text{f} \times \text{image_width}} \quad (4.3)$$

- **Coordinate Transformation and Fusion:**

- **Rotation Transformation:** The relative metric offsets from the image center ($\Delta X, \Delta Y$) are rotated based on the UAV's **Yaw** angle to align with the North-East-Down (NED) geographic coordinate system.
- **GPS Fusion:** The calculated metric displacements are converted into decimal degrees and fused with the drone's current position (Lat_D, Lon_D). This yields the absolute geographic coordinates (Output 4) of the wildfire.

Hình 6. Các bước tính toán chính

Ý nghĩa: Giúp lực lượng cứu hỏa biết chính xác vị trí đám cháy thay vì chỉ nhìn qua video.

HÀM MẤT MÁT ĐA MỤC TIÊU (MULTI-OBJECTIVE LOSS FUNCTION)

4.4.4 Multi-Objective Loss Function

To address the severe class imbalance and ambiguous fire boundaries in aerial wildfire imagery, we adopt a **Multi-Objective Loss** that combines Focal Loss, Dice Loss, and Soft Binary Cross-Entropy (Soft BCE). This composite loss encourages robust region overlap, emphasizes hard-to-classify fire pixels, and stabilizes training.

The overall loss function is defined as:

$$\mathcal{L}_{total} = \lambda_1 \mathcal{L}_{Focal} + \lambda_2 \mathcal{L}_{Dice} + \lambda_3 \mathcal{L}_{SoftBCE}$$

where λ_1 , λ_2 , and λ_3 are weighting coefficients.

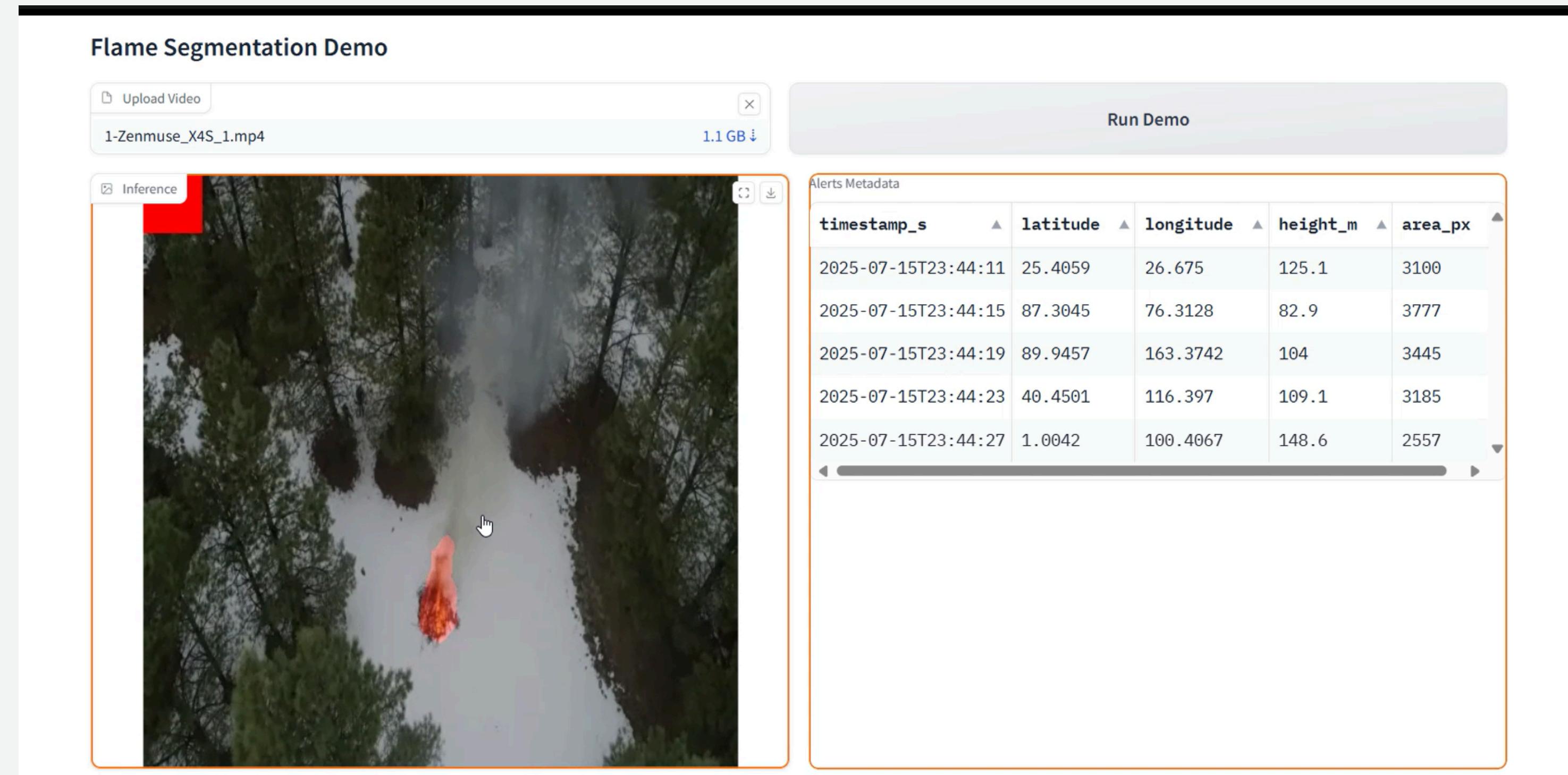
- **Thách thức:** Sự mất cân bằng lớp cực lớn (vùng cháy rất nhỏ so với nền rừng).
- **Công thức Loss tổng hợp (L_{-total}):** Kết hợp 3 thành phần để bổ trợ cho nhau:
 - **Dice Loss:** Tối ưu hóa độ chồng lấp giữa vùng dự đoán và vùng thật (IoU).
 - **Focal Loss:** Tập trung xử lý các pixel "khó" (vùng biên đám cháy, khói mỏng).
 - **Soft BCE:** Giúp ổn định quá trình hội tụ của gradient.
- **Kết quả:** Mô hình không bị "lười" (chỉ dự đoán toàn màu đen) mà tập trung bắt chính xác từng đốm lửa nhỏ.

II.5/ MODULE 3: LOGIC RA QUYẾT ĐỊNH VÀ KHỞI TẠO CẢNH BÁO

- **Vai trò (Functional Role):** Là lớp quản trị cuối cùng, đóng vai trò "bộ lọc" thông minh để ngăn chặn tình trạng "Alert Fatigue" (quá tải cảnh báo giả).
- **Kiểm chứng tính nhất quán Không-Thời gian (Spatio-temporal Consistency):**
 - **Vấn đề:** Các vật thể phản chiếu ánh sáng mặt trời, lóa ống kính (lens flare) hoặc nhiễu cảm biến có thể đánh lừa AI trong vài khung hình.
 - **Giải pháp:** Hệ thống áp dụng Temporal Persistence Filter (Bộ lọc kiên trì theo thời gian).
 - **Cơ chế:** Một sự kiện chỉ được xác nhận là "Đám cháy" nếu các pixel lửa duy trì sự nhất quán trên một cửa sổ gồm N khung hình liên tiếp.
- **Đóng góp:** Giảm đáng kể tỷ lệ phát hiện giả (FDR), đảm bảo chỉ những mối đe dọa thực sự mới được gửi đến đơn vị ứng cứu.
- **Cơ chế kích hoạt:** Khi điểm tin cậy tổng hợp vượt ngưỡng an toàn, hệ thống tự động đóng gói dữ liệu.
- **Gói tin cảnh báo (Telemetry Package) bao gồm:**
 - Tọa độ chính xác: Vĩ độ, kinh độ của tâm (centroid) vùng cháy.
 - Hình ảnh trực quan: Ảnh snapshot RGB được chồng lớp (overlay) mặt nạ phân vùng để người vận hành dễ dàng đánh giá quy mô lửa.
 - Dấu thời gian (Timestamp): Theo chuẩn ISO-8601 để ghi nhận chính xác thời điểm bùng phát.
- **Khả năng tương tác:** Dữ liệu được đóng gói chuẩn JSON, dễ dàng tích hợp vào các hệ thống hạ tầng quản lý thảm họa hiện đại.

II.5/ MODULE 3: LOGIC RA QUYẾT ĐỊNH VÀ KHỞI TẠO CẢNH BÁO

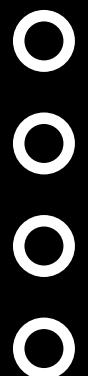
FINAL SYSTEM OUTPUT: RGB FRAME WITH SEMANTIC SEGMENTATION MASK OVERLAY AND REAL-TIME GEOSPATIAL ALERT METADATA.



Hình 7. demo hệ thống



III/ THỰC NGHIỆM



III.1/ CÁC ĐỘ ĐO

1. IoU: Đo lường mức độ chồng lấp giữa vùng cháy dự đoán và thực tế. Nhóm đặt mục tiêu $\text{IoU} \geq 0.60$ để đảm bảo bản đồ hỏa hoạn có độ trung thực cao.

$$\text{IoU} = \frac{\text{Area of Overlap}}{\text{Area of Union}} = \frac{|A \cap B|}{|A \cup B|} = \frac{TP}{TP + FP + FN}$$

2. F2-Score

- Tại sao dùng F2 thay vì F1?** Trong cứu hỏa, việc bỏ sót một đám cháy (False Negative) nguy hiểm hơn nhiều so với việc báo động nhầm (False Positive).
- Cơ chế:** Đặt trọng số $\beta = 2$ để ưu tiên chỉ số Recall (Độ nhạy).
- Mục tiêu:** $F2 \geq 0.75$

5.1.2 Alert Robustness and Safety Bias (F2-Score)

In wildfire surveillance, the cost of a *False Negative* (missing an active fire) far outweighs that of a *False Positive* (false alarm). Therefore, we utilize the **F2-Score**, a variant of the F-beta score that places greater statistical weight on **Recall** than on **Precision**:

$$F_\beta = (1 + \beta^2) \cdot \frac{\text{Precision} \cdot \text{Recall}}{(\beta^2 \cdot \text{Precision}) + \text{Recall}} \quad (5.2)$$

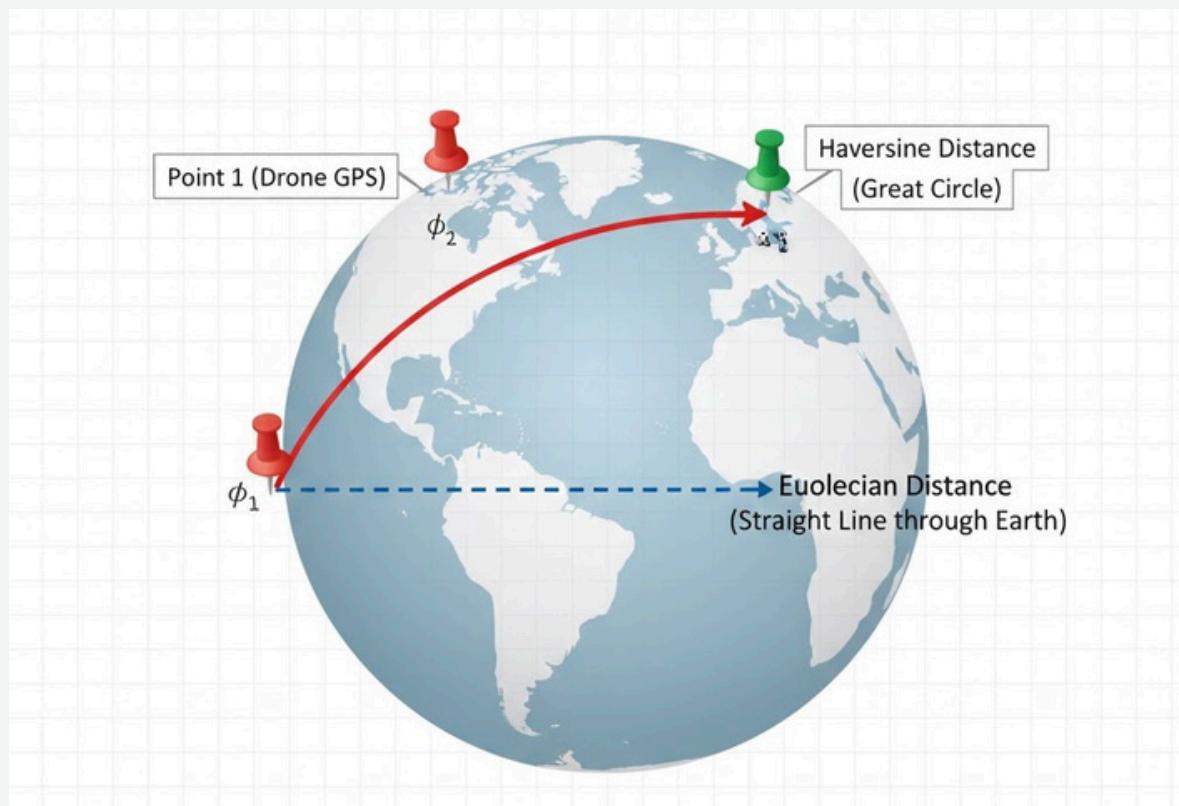
By setting $\beta = 2$, we prioritize the system's ability to detect all potential ignition points. The underlying components are defined as:

- Precision** ($\frac{TP}{TP+FP}$): The accuracy of the "fire" predictions.
- Recall** ($\frac{TP}{TP+FN}$): The sensitivity of the system in capturing all fire pixels.

Our objective is to maintain an F_2 -Score of at least 0.75 to ensure maximum safety.

III.1/ CÁC ĐỘ ĐO (Ht.)

3. Độ chính xác Định vị (Haversine Distance):



The distance d is calculated as follows:

$$d = 2r \cdot \arcsin \left(\sqrt{\sin^2 \left(\frac{\Delta\phi}{2} \right) + \cos(\phi_1) \cos(\phi_2) \sin^2 \left(\frac{\Delta\lambda}{2} \right)} \right)$$

where:

- ϕ_1, ϕ_2 : Latitudes of the ground truth and predicted points.
- λ_1, λ_2 : Longitudes of the ground truth and predicted points.
- r : Earth's radius ($\approx 6,371$ km).

- **Vấn đề:** Khoảng cách Euclid thông thường không tính đến độ cong của Trái Đất, gây sai số khi UAV bay xa.
- **Giải pháp:** Sử dụng công thức Haversine để tính khoảng cách theo đường tròn lớn (Great Circle).
- **Mục tiêu:** Giữ sai số định vị trung bình dưới 50 mét, giúp lực lượng cứu hộ tiếp cận đúng vị trí tâm dịch.

4. Global Pixel Accuracy

Đây chỉ là chỉ số phụ vì có sự mất cân bằng lớp (vùng cháy rất nhỏ so với nền rừng) → tập trung vào IoU và F2-Score để đánh giá hiệu quả thực tế hơn.

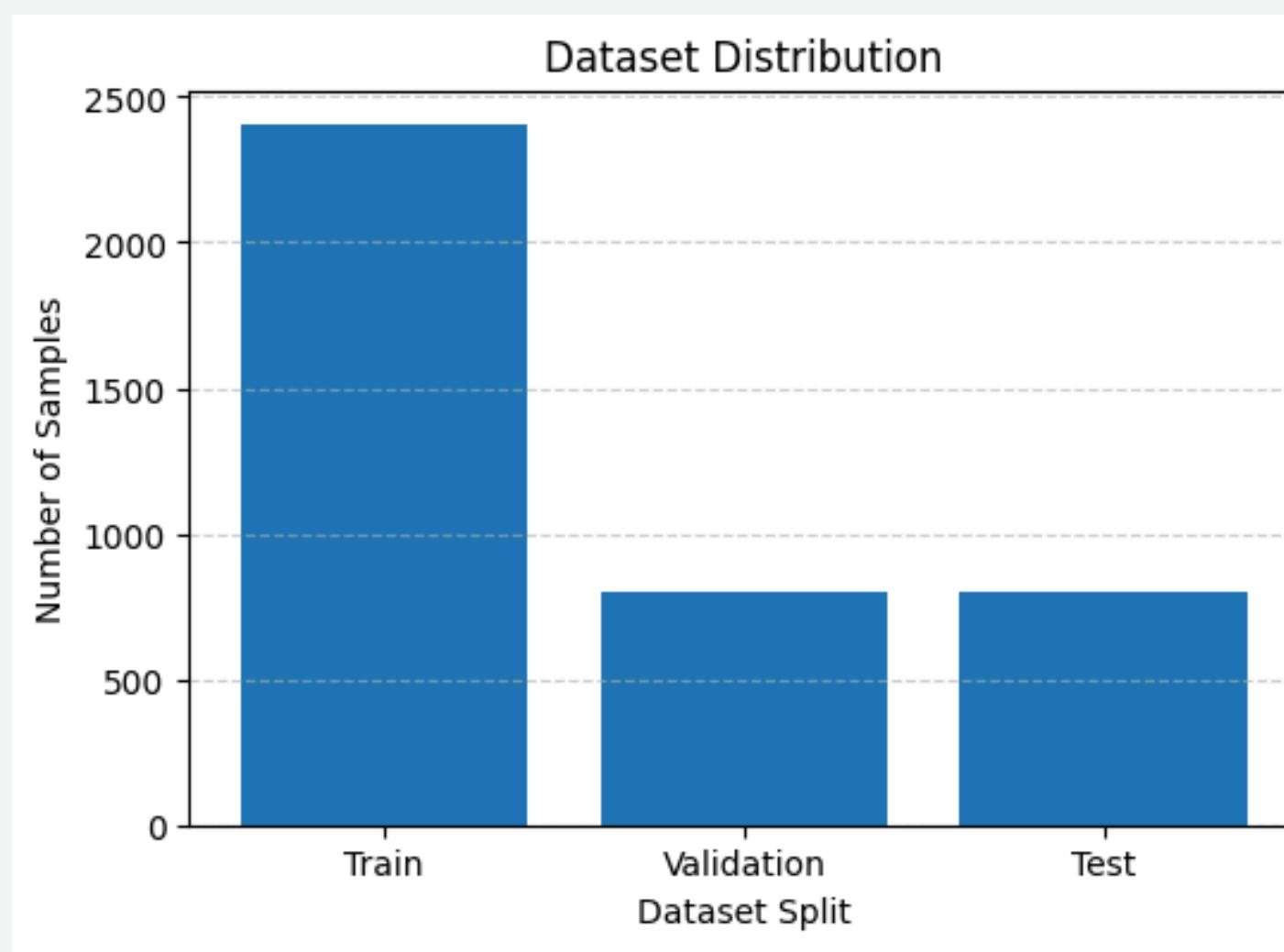
$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

III.2/ DATASET

Nguồn dữ liệu: Bộ dữ liệu FLAME Dataset (từ IEEE DataPort).

Phân chia dữ liệu: tỷ lệ 6:2:2, số lượng:

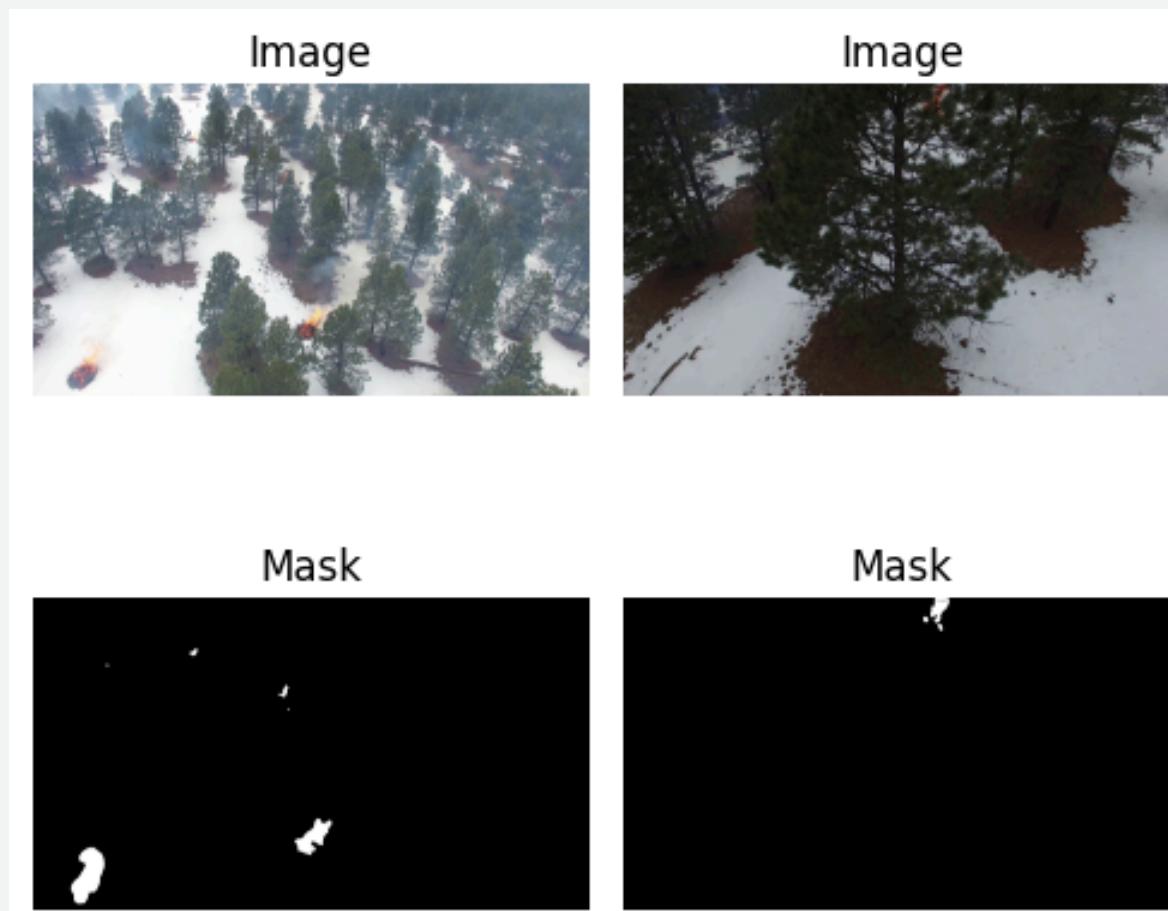
- TRAIN -> images: 1201, masks: 1201
- VAL -> images: 400, masks: 400
- TEST -> images: 402, masks: 402



Hình 8. Plot dataset distribution

```
train/  
  images/ (1,201 samples)  
  masks/ (1,201 samples)  
val/  
  images/ (400 samples)  
  masks/ (400 samples)  
test/  
  images/ (402 samples)  
  masks/ (402 samples)
```

Hình 9. Hierarchical directory structure



Hình 10. Sample Training Images

TĂNG CƯỜNG VÀ TIỀN XỬ LÝ DỮ LIỆU

- Tăng cường dữ liệu (Stochastic Augmentation):** Sử dụng thư viện Albumentations để tăng entropy của tập dữ liệu, chống quá khớp (overfitting).
- 3 Nhóm biến đổi chính:**
 - Hình học (Geometric):** Flip, Rotate, Perspective, RandomScale - Mô phỏng drone bay ở các độ cao và góc nhìn khác nhau.
 - Quang học (Photometric):** Brightness, Contrast, Hue, RGBShift - Xử lý sự thay đổi ánh sáng mặt trời và sai lệch màu cảm biến.
 - Nhiều môi trường (Environmental):** RandomFog (Khói), RandomShadow (Bóng cây), Blur - Giả lập điều kiện quan sát khắc nghiệt nhất trong rừng.
- Quy chuẩn hóa:** Mọi ảnh được đưa về Tensor kích thước 512 x 512 để tối ưu hóa tính toán trên thiết bị nhúng.

3. Data Augmentation

```

1 transform = A.Compose([
2     A.HorizontalFlip(p=0.5),
3     A.VerticalFlip(p=0.5),
4     A.RandomRotate90(p=0.5),
5     A.Rotate(limit=30, p=0.3),
6     A.Perspective(scale=(0.05, 0.1), p=0.3),
7     A.RandomScale(scale_limit=(-0.2, 0.2), p=0.3),
8
9     A.RandomBrightnessContrast(0.3, 0.3, p=0.5),
10    A.HueSaturationValue(20, 30, 20, p=0.5),
11    A.RGBShift(20, 20, 20, p=0.3),
12
13    A.RandomFog(alpha_coef=0.1, p=0.4),
14    A.RandomShadow(shadow_roi=(0, 0, 1, 0.5), p=0.4),
15    A.GaussianBlur(blur_limit=(3, 7), p=0.3),
16
17    A.Resize(512, 512),
18 ], p=1.0)

```

Hình 11. Cài đặt trong bài toán

MÔI TRƯỜNG HỆ THỐNG & KIẾN TRÚC MÔ HÌNH

- **Môi trường Phần cứng & Phần mềm:**
 - **GPU:** Sử dụng NVIDIA GeForce RTX (CUDA-enabled) - Đảm bảo xử lý song song, tối ưu cho huấn luyện và chạy real-time trên drone.
 - **Framework:** PyTorch Lightning (đảm bảo tính module và dễ mở rộng) kết hợp với TorchMetrics để tính toán IoU, Recall chuẩn xác

- **Kiến trúc Mô hình (U-Net + ResNet-34):**
 - **Backbone (Encoder):** Sử dụng **ResNet-34** đã được huấn luyện sẵn (**Pre-trained on ImageNet**).
 - **Lợi ích:** Tận dụng kỹ thuật **Transfer Learning** giúp mô hình hội tụ nhanh hơn, đồng thời các khối **Residual Blocks** giúp trích xuất đặc trưng sâu mà không bị mất tín hiệu gradient.
 - **Skip Connections:** Giúp khôi phục chi tiết các biến hỏa hoạn phức tạp từ Encoder sang Decode

III3/ EXPERIMENTAL SETUP AND ENVIRONMENTS

CẤU HÌNH HUẤN LUYỆN & THAM SỐ TỐI ƯU

- **Dữ liệu & Tiền xử lý:**
 - Kích thước ảnh: 512 x 512 pixel.
 - Chuẩn hóa: Z-score normalization giúp ổn định dòng dữ liệu.
- **Hàm mất mát đa mục tiêu (Multi-Objective Loss):**
 - Tổng hợp 3 thành phần: $L_{total} = L_{Focal} + L_{Dice} + L_{SoftBCE}$
 - Mục tiêu: Vừa bắt chính xác các đốm lửa nhỏ (Focal), vừa đảm bảo độ chồng lấp mặt nạ cao (Dice), vừa ổn định quá trình học (BCE).
- **Chiến lược tối ưu hóa (Optimization):**
 - Optimizer: Adam (Learning Rate: 2×10^{-4}).
 - Scheduler: Cosine Annealing giúp LR giảm dần một cách mượt mà về 1×10^{-5} giúp mô hình hội tụ vào các cực tiểu tối ưu.
 - Tham số khác: Batch size = 4; Epochs = 50.

III.4/ KẾT QUẢ THỰC NGHIỆM

Method	Accuracy	F ₂ -score	IoU	Precision	Recall
U-Net + BCE	0.9978	0.7228	0.6396	0.8992	0.6890
U-Net + BCE + Aug	0.9981	0.7672	0.6892	0.9129	0.7377
U-Net + 3-Loss	0.9980	0.7740	0.6847	0.8870	0.7501
U-Net + 3-Loss + Aug	0.9982	0.8162	0.7139	0.8627	0.8054

Hình 12. Kết quả trên tập test

Những cải tiến mang tính quyết định:

- Sức mạnh của 3-Loss:** Việc chuyển từ BCE sang hàm mất mát đa mục tiêu (Focal + Dice + SoftBCE) giúp chỉ số **Recall tăng mạnh từ 68.9% lên 80.5%**. Điều này giúp giảm thiểu tối đa hiện tượng bỏ sót đám cháy (False Negatives).
- Hiệu quả của Augmentation:** Các phép biến đổi giúp mô hình xử lý tốt nhiều môi trường, cải thiện chỉ số IoU thêm **5-8%**.
- Độ an toàn (F2-score):** Đạt **0.8162**, vượt xa mục tiêu đề ra ban đầu (0.75), khẳng định độ bền bỉ của hệ thống.

○ ○ ○

IV/ DEMO

○ ○ ○

DEMO

LINK VIDEO DEMO

Flame Segmentation Demo

Upload Video

Drop File Here
- or -
Click to Upload

Run Demo

Inference

Alerts Metadata

timestamp_s	latitude	longitude	height_m	area_px

○ ○ ○

V/KẾT LUẬN

○ ○ ○

KẾT LUẬN

- **Hiệu năng hệ thống:**

- Ứng dụng thành công mạng U-Net với ResNet-34 backbone kết hợp hàm mất mát đa mục tiêu.
- Đạt chỉ số IoU cao: 0.7139 và F2-score: 0.8162, khẳng định khả năng phân đoạn biên hỏa hoạn chính xác.

- **Đặc điểm nổi bật:**

- Phân đoạn mức Pixel: Cung cấp thông tin chi tiết về diện tích và ranh giới lửa.
- Độ tin cậy cao: Nhờ bộ lọc nhất quán theo thời gian và định vị địa lý chính xác, tạo ra một khung làm việc mạnh mẽ cho can thiệp sớm.

- **Ý nghĩa:** Chứng minh tiềm năng to lớn của việc tích hợp Deep Learning và Drone trong bảo vệ môi trường.

- **Vấn đề Đạo đức & Xã hội:**

- Quyền riêng tư: Đảm bảo ranh giới thu thập dữ liệu và ẩn danh hóa thông tin tại các khu vực dân cư lân cận.
- Trách nhiệm: Cần cơ chế kiểm toán hệ thống minh bạch trong các quyết định của AI.
- Việc làm: Hệ thống đóng vai trò công cụ hỗ trợ quyết định; cần đào tạo lại nhân sự để quản lý và vận hành công nghệ mới.

- **Hướng phát triển tương lai (Future Work):**

- Tối ưu hóa phần cứng: Nén mô hình (Quantization/Pruning) để chạy trực tiếp trên chip nhúng của Drone (NVIDIA Jetson).
- Fusion dữ liệu đa phổ: Kết hợp thêm ảnh Nhiệt (Infrared) để nhìn xuyên khói và hoạt động ban đêm.



THANKS FOR
LISTENING