

Appendices

A. Details Of Collected Imagery

This section provides supporting details of the sUAS imagery that was collected in support of Patricia Wu-Murad. Figure 4 contains the geographic information associated with the collected imagery. Figure 5 contains the runtimes of the sUAS videos that were collected during the search.

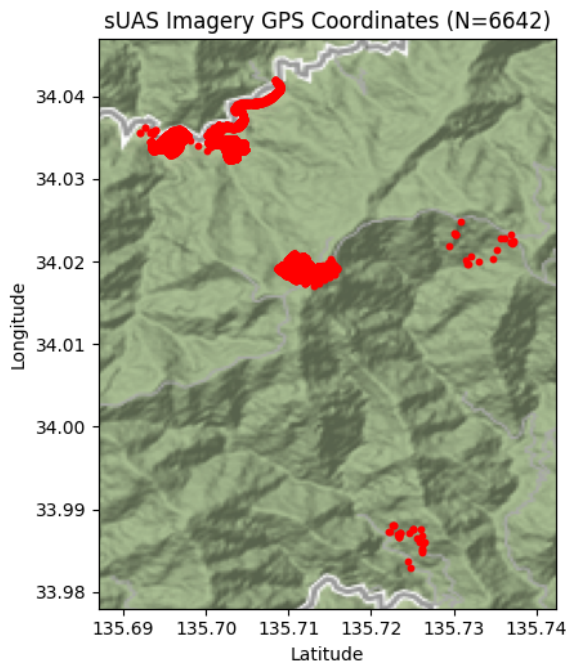


Figure 4. The GPS coordinates of imagery collected in the search for Patricia Wu-Murad denoted in red. Imagery that did not have an associated GPS coordinate has been omitted.

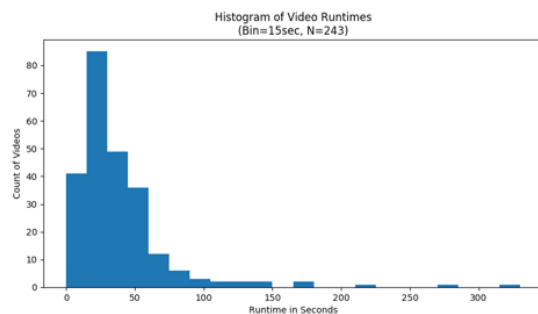


Figure 5. Histogram of the runtimes of the 243 collected videos from the search for Patricia Wu-Murad.

B. RX Classifier Hyperparameters

The following are the hyperparameters used when running the RX spectral classifier model.

- **Image Resize:** All input images are resized to be 1024x1024
- **P-Value:** Pixels are considered anomalous if they have a z-score and associated p-value of < 0.0001
- **DBSCAN eps:** Anomalous pixels were clustered together if their distance was < 14.4815
- **DBSCAN Min Samples:** Clusters were dropped if they had < 209 pixels.
- **Cluster Count Limits:** Images that resulted in < 1 or > 4 clusters were ignored.

Sample inputs and outputs, and the intermediate pixel classification step is included in Figure 6 for awareness.

C. EfficientDET Training & Hyperparameters

The input resolution of the EfficientDET model was 512x512 and it was trained on the HERIDAL dataset[28]. The tf_efficientnetv2_l model [46] was used as the backbone which was then fine tuned using HERIDAL data. The HERIDAL has a published train and test set. The train set was randomly split into into 90% training samples and 10% validation samples.

Training samples were generated in the following manner. The functionality of the Albumentations library [13] was utilized and extended.

1. Select a random image from the HERIDAL training set
2. Randomly resize the image between 0.7 and 1.1 of the original image size while maintain aspect ratio.
3. If that image contains no bounding boxes select a random 512x512 crop from the image.
4. If that image contains at least one bounding box, with equal likelihood
 - Select a random 512x512 crop from that image
 - Select a random bounding box and then select a random 512x512 crop of that image that contains that selected bounding box.

Training samples were augmented in the following manner. The Albumentations library [13] was used to perform these augmentations.

1. Randomly flip the image along its x or y axis with 50% likelihood
2. Randomly rotate the image at 90 degrees increments between 0-270 around its center
3. Perform embossing at 25% likelihood
4. Synthetically add snow to at 15% likelihood [43]

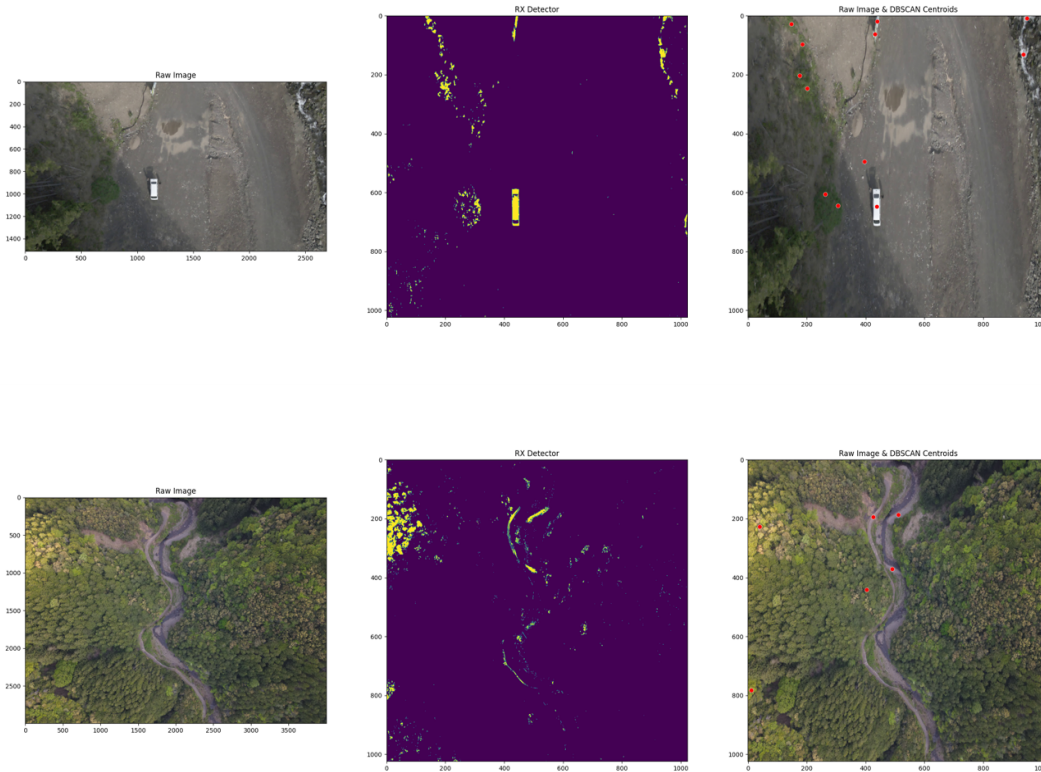


Figure 6. Sample inputs and outputs of the Unsupervised RX Spectral classified as applied to imagery collected from the field.

5. Synthetically add fog to at 15% likelihood (fog_coef_lower=0.01, fog_coef_upper=0.4) [43]
6. Apply the sepia filter to the image at 10% likelihood
7. Add random Gaussian noise to every pixel
8. Perform one of the following
 - Add random brightness and contrast to the image at 50% likelihood
 - Jitter the color of the image at 50% likelihood

Finally, the color values of the image were normalized according to the following color channel mean values [R: 0.485, G: 0.456, B: 0.406] and standard deviation values [R: 0.229, G: 0.224, B: 0.225].

The default set of hyperparameter set forth in the EffDet python library [49] were used. The relevant hyperparameters are detailed below.

- **Learning Rate:** 0.0002
- **Optimizer:** Adam
- **backbone_drop_path_rate:** 0.2
- **backbone_indices:** None
- **min_level:** 3
- **max_level:** 7
- **num_levels:** 5
- **num_scales:** 3
- **anchor_scale:** 4.0
- **pad_type:** same
- **act_type:** swish
- **norm_layer:** None
- **norm_eps:** 0.001
- **norm_momentum:** 0.01
- **box_class_repeats:** 3
- **fpn_cell_repeats:** 3
- **fpn_channels:** 88
- **separable_conv:** True
- **apply_resample_bn:** True
- **conv_bn_relu_pattern:** False
- **downsample_type:** max
- **upsample_type:** nearest
- **redundant_bias:** True
- **head_bn_level_first:** False
- **head_act_type:** None
- **fpn_name:** None
- **fpn_config:** None
- **fpn_drop_path_rate:** 0.0
- **alpha:** 0.25
- **gamma:** 1.5
- **label_smoothing:** 0.0
- **legacy_focal:** False
- **jit_loss:** False

- **delta:** 0.1
- **box_loss_weight:** 50.0
- **soft_nms:** False
- **max_detection_points:** 5000
- **max_det_per_image:** 100

The bounding boxes that are generated by the model are fused together using the Weighted Box Fusion technique described in [44]. When performing tiled inference on the HERIDAL dataset, or on data collected operationally, any bounding boxes that are overlapping are merged together by generating a larger bounding box that contains exactly the two overlapping bounding boxes.

The *Tiled_EfficientDET*₈₄ and *Tiled_EfficientDET*₁₇₄ models were trained for 84 and 174 epochs respectively, where one epoch was defined as having generated one sample from each image in the training set. The training process for *Tiled_EfficientDET*₁₇₄ took approximately 16 hours and training was conducted on a Nvidia 1080TI.

Inference on the 18759 images generated in the search for Patricia Wu-Murad took approximately 4.74 hours and on average 0.91 images per sec were processed.

D. Four Candidate Regions of Interest

The four candidate regions of interest that were handed off to the ground search teams are contained in Figures 7, 8, 9, and 10. None of these regions of interest appear to show a person, instead they show an abnormality that cannot be immediately explained.



Figure 7. Candidate Image #1 that was handed off to the ground search teams.



Figure 8. Candidate Image #2 that was handed off to the ground search teams.



Figure 9. Candidate Image #3 that was handed off to the ground search teams.



Figure 10. Candidate Image #4 that was handed off to the ground search teams.