

山岳地帯での遭難者救助を想定した、赤外線カメラを使用した人物検出に使用するデータセットの検討

Study on a Dataset for Detecting Persons with Infrared Camera Supposing to rescue in Mountainous Area

○塩崎 雄晴^{*1}, 入江 寿弘^{*2}, 小林 伸彰^{*3}, 新宮 清志^{*4}
Yusei Shiozaki^{*1}, Toshihiro Irie^{*2}, Nobuaki Kobayashi^{*3} and Kiyoshi Shingu^{*4}

*1 日本大学大学院理工学研究科精密機械工学専攻 大学院生

Graduate Student, Precision Machinery Engineering, Graduate School of Science and Technology, Nihon University.

*2 日本大学理工学部精密機械工学科 教授 博士(工学)

Professor, Department of Precision Machinery Engineering, College of Science and Technology, Nihon University, Dr. Eng.

*3 日本大学理工学部精密機械工学科 准教授 博士(工学)

Associate Professor, Department of Precision Machinery Engineering, College of Science and Technology, Nihon University, Dr. Eng.

*4 日本大学 名誉教授, 総合資格学院 特別顧問 工学博士

Professor Emeritus, Nihon University, and Executive Senior Adviser, Sogo-Shikaku-Gakuin, Dr. Eng.

Summary: In this study, we have attempted to build a dataset of images captured using infrared cameras to detect humans.

Detecting the location of victims as soon as possible in mountainous regions and other disaster sites at night is required for the rescue teams. Therefore, we have considered creating a dataset using the infrared camera images of a human in mountainous regions. In addition, the effectiveness of the YOLOv3 and Tiny-YOLOv3 object detection has also confirmed.

キーワード: 人工知能; 深層学習; 物体検出; 害獣検出; 赤外線カメラ

Keywords: artificial intelligence (AI); deep learning; object detection; infrared camera.

1. はじめに

近年、登山者の増加に伴い、山岳遭難者の数が著しく増加している。救助隊による救助活動が行われる際には、遭難現場の情報が全くない場合、広大な山岳地をくまなく捜索する必要があり、捜索は困難を極め、莫大な時間と費用が生じる。さらに、救助隊員の負担も大きくなり、二次災害の危険が増加する可能性がある。この問題は、建築・都市空間における大規模地震や集中豪雨などの自然災害による被害者の捜索にも存在する。特に、目視による瓦礫の下敷きになった被害者の捜索は非常に困難で、時間を要してしまい、生存可能時間内に救出できない可能性が高くなると考えられる。このような状況下ではより短時間に土砂等の足場の状況に左右されずに上空から捜索することが有用となる。一方で、昼間の捜索はカメラ画像や肉眼での捜索が可能であるが、夜間は視認や撮影ができないため、定点カメラやドローンでの捜索を想定すると、熱画像等での自動認識が必要になると考えられる。しかし、遭難者の位置情報が乏しい状態での探索は複数台のドローンを広範囲にわたって稼働させる必要がある。そのため、ドローンを監視する人員を削減する

ために、我々は深層学習による物体検出アルゴリズムによって自動的に遭難者を検出するシステムを検討した。

本研究では、ドローンに搭載した赤外線カメラを用いて、深層学習による熱画像中の人物検出を試みた。リアルタイムでの人物検出を想定するため、処理速度が速い物体検出器が求められる。我々が本研究で用いたYOLOv3^{1),2)}、Tiny-YOLOv3 と呼ばれるアルゴリズムは画像の領域推定と分類を同時に処理するため短時間で物体を検出することができる。この双方のモデルに、山岳を模した場所で、赤外線カメラで空撮した熱画像を学習させた。人物や動物を目線程度の高さで撮影した赤外線データセットは自動運転の分野などで多く検討されており³⁾⁴⁾、また、俯瞰から赤外線カメラで人間を空撮したデータセットは災害救助などの分野で多く検討されているが、人物検出などを想定した俯瞰視点から人物を赤外線カメラで撮影したデータセットは報告されておらず、独自のデータセットを作成し検討した。

また、使用する検証用動画に対し、コントラスト補正処理を加えた。これによって人間とその他の物体、背景の明暗の差を大きくして差別化を図った。そのうえで、

学習済みモデルを用いた人物検出を行い、コントラスト補正前後の状態を比較、検討を行った。

2. YOLOv3

本研究では、物体検出に YOLOv3 という手法を用いた。YOLOv3 とは、2018 年に Joseph Redmon と Ali Farhadi によって提案された手法であり、Faster R-CNN⁵⁾のような従来の手法と比較してリアルタイム性に優れた物体検出モデルである。この手法の特徴は各領域推定とクラス分類を一つの畳み込みニューラルネットワークで並列に行う点である。これによって処理速度の高速化を実現している。

今回は、このアルゴリズムを基準として作成したデータセットの検討を行った。

2.1. YOLOv3 と Tiny-YOLOv3

YOLOv3 で使用した Darknet-53 に対し、Tiny-YOLOv3 は Convolution レイヤの数が少なく構成されている⁶⁾。この特徴により、Tiny-YOLOv3 は YOLOv3 に比べて精度は落ちるが、処理速度を向上させている。

3. 物体検出実験

3.1. データセット 1

ドローンを使用して空撮した熱画像から人物検出を行うため、独自の熱画像データセットを作成した。人物撮影では熱画像カメラに FLIR 社製 FLIR ONE を用いて撮影を行い、ドローンは Holy Stone Toys HS700 を使用した。撮影場所は、図 1 に示すドローンリゾート南熱海⁶⁾の飛行場での撮影を行った。

また、得られた熱画像に対し、回転・反転のデータ拡張処理を行い、合計 2280 枚の熱画像を作成した。その後、1 枚ずつアノテーションを行い、クラス名を”person”と定めた。ここでのアノテーションとは、教師データに対象となる物体が指定したクラスとして認識できるようにタグ付けを行う行為である。テスト用のデータは同条件にて撮影した動画の一部を使用した。

3.2. 比較用データ

比較のため、異なる場所で撮影した比較用データを用意した。データセット 1 は山岳で撮影されたものであり、比較用データは市街地等の異なる条件下での人物の俯瞰撮影データである。比較用データの撮影カメラは FLIR 社の FLIR C3 である。また、各データの情報を表 1 に示す。



図 1 「ドローンリゾート南熱海」
飛行場 天候 曇り



図 2 モデルに学習させた熱画像例



図 3 比較用データ

3.3. 学習

学習に使用した人物の熱画像の例を図 2 に示す。COCO(Common Objects in Context⁷⁾)による学習済みモデルを使用して、独自データセットを YOLOv3、Tiny-YOLOv3 双方に学習させた。図 4、図 5 のように 1~50 エポック (学習用データセット全てを学習する回数) 間では、最終畳み込み層のみを学習させ、図 6、図 7 のように 51 エポック以降では、学習率を下げて全ての層を学習させた。図 4~図 7 の loss と val_loss はそれぞれトレー

ニングセットの学習についての損失関数のスコアとバリ
デーションセットを用いたチューニング時の損失関数の
スコアを表している。

3.4. コントラスト補正処理

データセット1は全体的に白く、人間と背景が識別し
づらいことも検出精度に影響していると考え、テスト用
の動画にコントラスト補正処理を加え、明暗の差を大き
くすることで、検出精度の向上を試みた。

図8、図9に示すヒストグラムの pixel value は、ピク
セルの輝度値を、縦軸の frequency はその輝度値に属する
ピクセルの数を示している。図8より、コントラスト補
正前の画像では、画像に含まれるピクセルの輝度値がお
おそ100~150の間に多く存在し、画像全体の明暗の差
が乏しく、輝度値の分布が狭いことがわかる。また、図
9にコントラスト補正後の熱画像を示した。これによっ
て画像全体の輝度の分布が広がり、背景と人物との明暗
の差が大きくなったことがわかる。

表1 各データの情報

データセット1	
カメラ1	FLIR ONE
ドローン	HOLY STONE HS700S
端末	iPhone6S
撮影場所	ドローンリゾート南熱海
日時・天候	2018/11/13 14:00 曇り
画像数	2280枚
解像度	1080x1440
テストデータ	
本数	2本
動画時間	5m16sec,1m07sec
比較用データ	
カメラ2	FLIR C3
撮影場所	神奈川県 東海道線 辻堂駅周辺
日時・天候	2020/09/27 18:00 曇り
動画時間	64sec
フレームレート	9Hz
解像度	640x480

4. 実験結果

4.1. 熱画像中の人物検出

作成したデータセット1を学習させたYOLOv3とTiny-
YOLOv3で、赤外線カメラで空撮したテストデータから
人物検出を行った。出力結果を図10に示し、YOLOv3、
Tiny-YOLOv3で誤検出が発生した出力結果を図11に示
す。また、テストデータに対してのF値とFPSの出力結
果を表2に示す。ただし、F値は適合率と再現率の調和
平均で、計算に用いるTP, FP, FNのカウントは出力され
たテストデータの人物が映っている各フレームに対しカ

ウントした。また、図11~図13中のA~Dは誤検出対象
を示している。

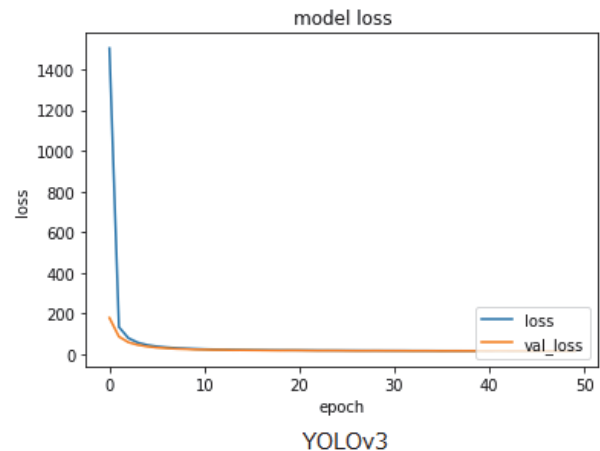


図4 1~50エポック間の学習推移(YOLOv3)

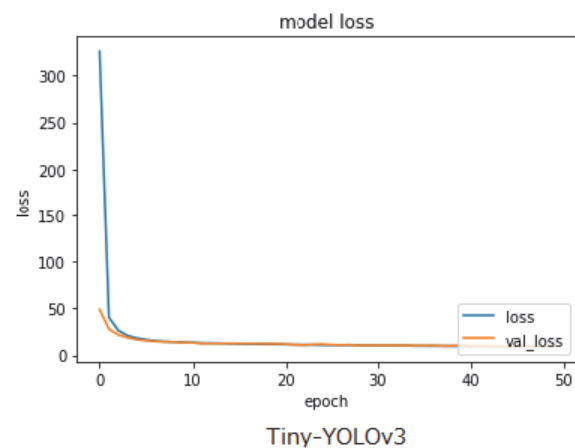


図5 1~50エポック間の学習推移(Tiny-YOLOv3)

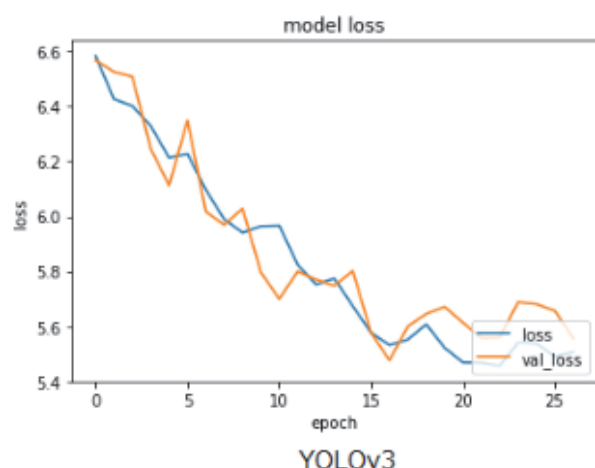


図6 51エポック以降の学習推移(YOLOv3)

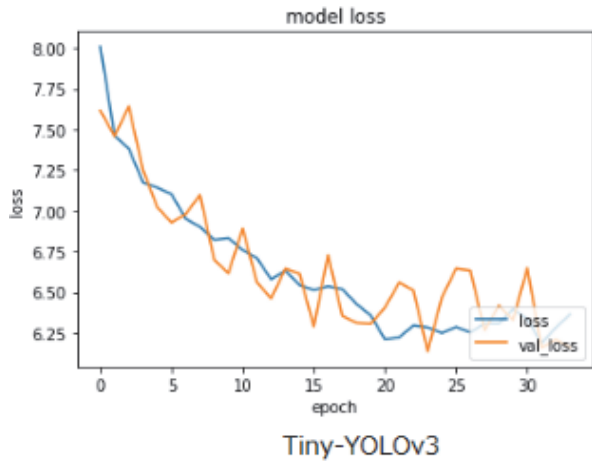


図 7 51 エポック以降の学習推移(Tiny-YOLOv3)

動作させ、汎用性を確認した。出力結果を図 14 と図 15 に示す。また、学習前の YOLOv3 の図 14 と図 15 の右画像の同フレームに対する出力結果を図 16 に示す。ただし、図 14~16 中の E,F は誤検出対象とし、検出クラスはすべて"person"、数値はスコアとする。



図 10 出力結果(左:YOLOv3、右:Tiny-YOLOv3)

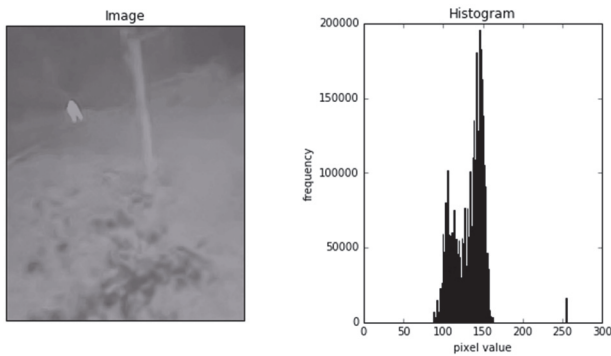


図 8 補正前画像とそのヒストグラム

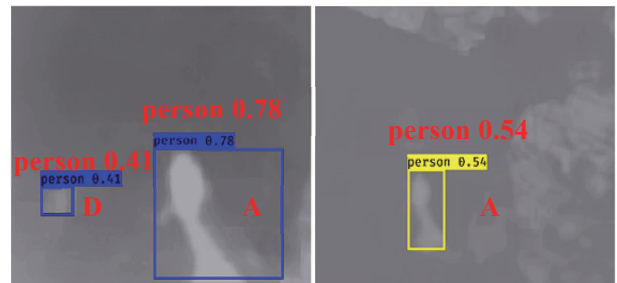


図 11 誤検出が発生した出力結果 (左:YOLOv3、右:Tiny-YOLOv3)

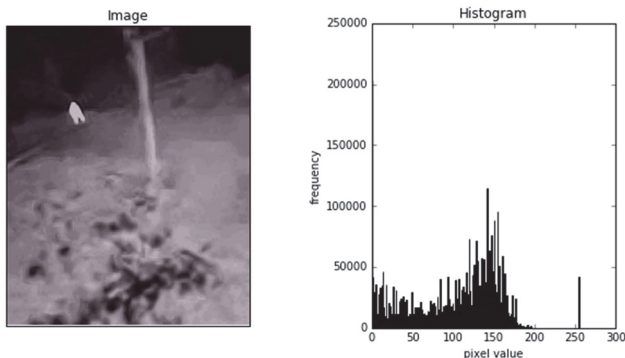


図 9 補正後画像とそのヒストグラム

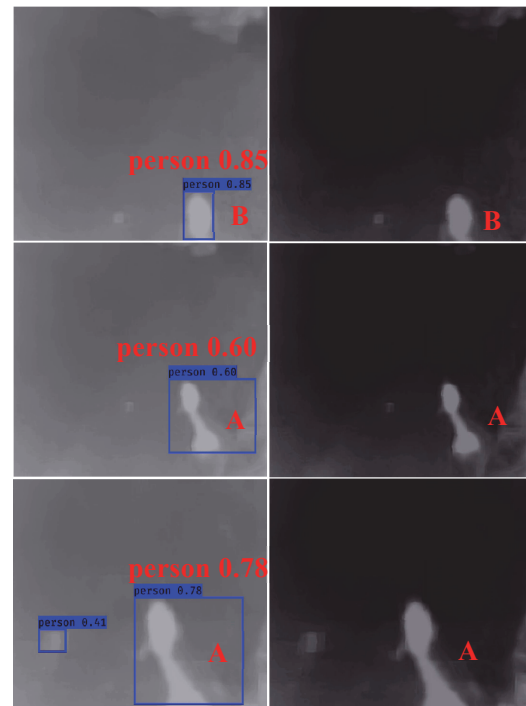


図 12 テスト用動画にコントラスト補正を加えた補正前(左)補正後(右)の出力結果(YOLOv3)

4.2. コントラスト補正

テストデータにコントラスト補正を加えて YOLOv3 と Tiny-YOLOv3 で検出を行った。誤検出した画像に対してコントラスト補正を加えた出力結果を図 12 と図 13 に示す。

4.3. 比較用データとの比較と汎用性の検討

異なる環境で同じような角度から人物撮影した比較用データに対してデータセット 1 を学習させた YOLOv3 を

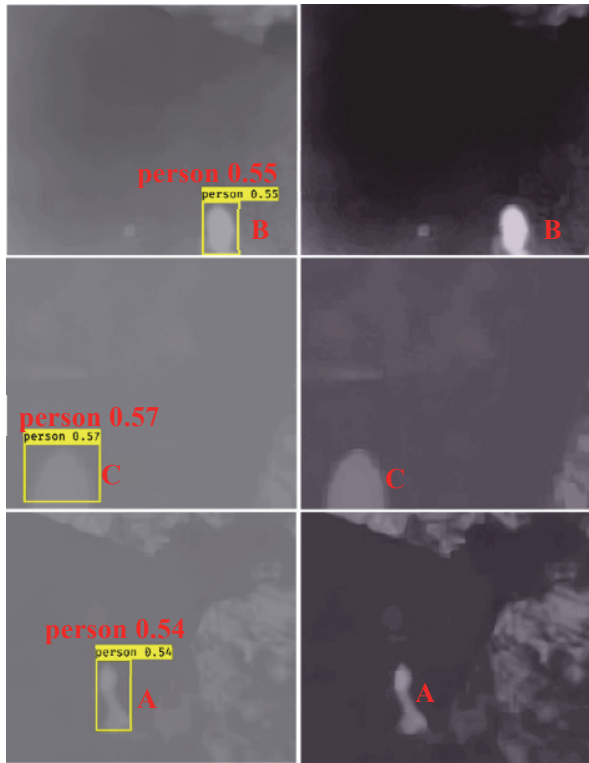


図 13 テスト用動画にコントラスト補正を加えた補正前(左)補正後(右)の出力結果(Tiny-YOLOv3)

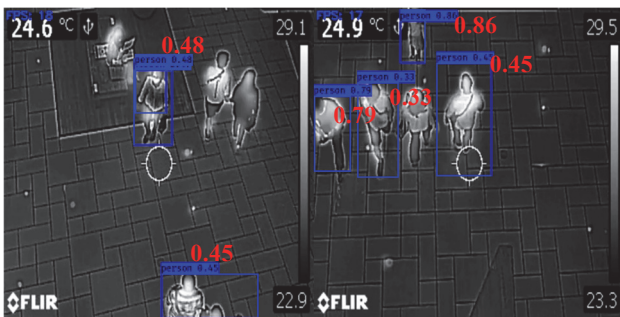


図 14 比較用データの出力結果

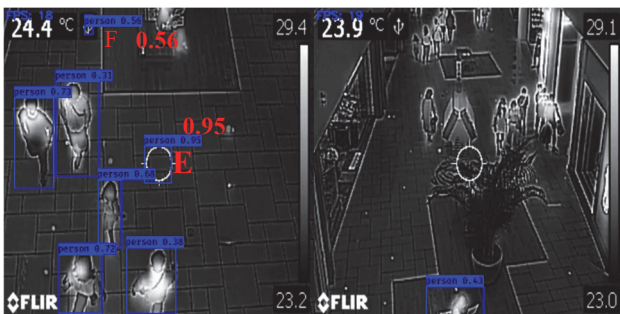


図 15 誤検出(左 E,F)と未検出(右)がある比較用データの出力結果

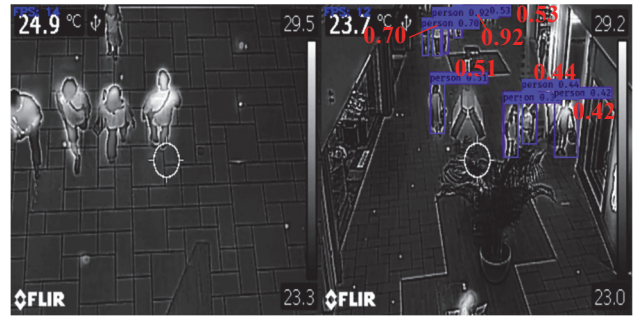


図 16 比較用データに対する学習前の YOLOv3 の出力結果

5. 考察

作成した熱画像データセットを学習させた YOLOv3 と Tiny-YOLOv3 を用いて、山岳を模した場所での人物検出を行った。表 2 より、再現率は YOLOv3 が 94.6[%]、Tiny-YOLOv3 が 94.4[%]と双方のモデルとも高い値であった。しかし、出力動画のフレームレートに大きな差が生じた。各フレームレートは表 2 より YOLOv3 が 4~5[FPS]、Tiny-YOLOv3 が 11~13[FPS]であった。この結果より、Tiny-YOLOv3 の方が YOLOv3 よりも高速に処理ができることがわかる。Tiny-YOLOv3 は 2.1 節で述べたように convolution レイヤ数が YOLOv3 に比べ少なく構成されているため、再現率や適合率が低くなると予測されるが、本研究のように遭難者を想定した人物の 1 クラスのみを設定した場合は十分な再現率と適合率を示すと考えられる。

また、本研究では災害遭難者の検出を目的としているため、リアルタイム性が求められる。処理速度の面では、ハードウェアの性能などにも大きく依存するが、その分費用が高額になる。したがって、処理速度が速く、リアルタイム性に優れた Tiny-YOLOv3 を探索システムに使用することが有用であると期待できる。今回の出力結果では、図 11 に示すような誤検出が見られた。この誤検出のパターンは YOLOv3 と Tiny-YOLOv3 双方のモデルで発生している。この理由として被写体の輪郭、背景との明暗の差が不明瞭であることが原因であると考えた。そこで誤検出を取り除くために、入力するテスト用動画に対して、コントラスト補正処理を加えて検証した結果、図 12、図 13 に示すように誤検出を取り除くことに成功した。この結果は、コントラスト補正処理を施したことによって物体と背景の明暗差を大きくし、物体と背景の境界が鮮明に表示されたためと考えられる。一方で、表 2 を見るとコントラスト補正後の Tiny-YOLOv3 の再現率が低下している。ここで、検出されなかった画像には、人間ではあるが、赤外線カメラとの距離により輪郭の丸い白い物体のようになってしまった人間が写ったものが多くみられた。

表 2 人物検出における実験結果と性能指数

	適合率[%]	再現率[%]	F値[%]	フレームレート[FPS]
YOLOv3	99.2	94.6	96.9	4~5
Tiny-YOLOv3	99.3	94.4	96.8	11~13
YOLOv3(補正後)	100	91.9	95.8	4~5
Tiny-YOLOv3(補正後)	100	83.1	90.8	11~13

これらの画像は学習データにおいて”person”のクラスを割り当てられているが、コントラスト補正処理後の検証用動画ではその外形から、人間だと検出されなかったためであると考えられる。

汎用性については、図 14 に示すとおり、人物の検出ができているが、図 15 のように中央のレティクルに対して person と誤検出してしまっている場合もある。これはデータセット 1 に含まれる遠距離の人物が丸くぼやけた輪郭をもってしまったため、白く輪郭が丸いものに反応してしまっていると考えられる。今回撮影に使用した FLIR ONE と FLIR C3 はいずれも遠距離撮影ができる性能をもっておらず、より性能の良い素子を備えた機材を用いて撮影、検出を行えばより確実に人物を検出できる可能性がある。

また、今回は人物を俯瞰から撮影した場合を想定しているため、人物の横側から撮影したデータセットを用意しておらず、図 15 の右の画像のように、カメラと人物との距離が開くと、検出から外れることが確認できる。

最後に、学習前の YOLOv3 は COCO データセットを用いて事前に学習されており、人間に対しての検出が可能であるが、図 16 左図に示すように、人物の頭上からの検出に対しての検出率は高くないと思われる。しかし、データセット 1 による学習により、図 14 のように人物の俯瞰からの検出が可能となった。

6. まとめ

本研究では、前例が少ない空撮を想定した俯瞰の熱画像を元に作成した独自のデータセットを YOLOv3 や Tiny-YOLOv3 に学習させ、検証用動画へのコントラスト補正処理を加えることで、夜間での山岳遭難者の検出を想定した熱画像を利用した人物の検出に世界に先駆けて成功した。今回は山岳遭難者を想定してデータセットを作成し検討したが、これらは都市災害時の被害者捜索への拡張が期待できる。人の視線程度の高さから人間を撮影し、検出した前例は数多くあるが、俯瞰からみた人物はその輪郭が大きくことなるため、改めてデータセットを作成し、その有用性を検討した。

また、今後、ドローンや定点カメラを活用することで夜間での遭難者の位置情報も取得可能になり、災害救助

の支援となることが期待できる。

【参考文献】

- 1) Joseph Redmon, Ali Farhadi : YOLOv3: An Incremental Improvement pp.1-2 arXiv:1804.02767v1 2018
- 2) ” YOLO:Real-Time Object Detection”
<https://pjreddie.com/darknet/yolo/>
(参照日 2020-04-02)
- 3) Hart AG, Rolfe RN, Dandy S, Stubbs H,MacTavish D, MacTavish L, et al. (2015) CanHandheld Thermal Imaging Technology ImproveDetection of Poachers in African Bushveldt?. PLoSONE 10(6): e0131584. doi:10.1371/journal.pone.0131584 pp-6
- 4) <https://www.electronicdesign.com/markets/automotive/article/21807456/thermal-imaging-will-make-autonomous-vehicles-safer-and-more-affordable>
(参照日 2020-09-30)
- 5) Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks” arXiv:1506.01497v3, 2016.
(参照日 2020-09-25)
- 6) “ドローンリゾート南熱海”
<https://drone-resort.com/>
- 7) Yi Lin, Tsung, et al. “Microsoft COCO:Common Objects in Context” Computer Vision-ECCV 2014 pp.740 - 755. arXiv:1405.0312v3 2015