

## ドローン空撮映像からの安全領域判定手法の実装と性能評価 Implementation and Evaluation of Safety Area Judgment Method from Drone Aerial Video

江村 優吾<sup>†</sup>      小嶋 和徳<sup>†</sup>      伊藤 慶明<sup>†</sup>      馬淵 浩司<sup>†</sup>  
Yugo Emura      Kazunori Kojima      Yoshiaki Itoh      Hiroshi Mabuchi

### 1. はじめに

近年ドローンは幅広い分野で用いられる。足場が組めない高所を撮影できたり、地上では時間がかかる地点を短時間で運搬できる。近年老朽化したプラント内での点検や検査が注目される一方、引火物へのリスク配慮も求められている。現在のドローンは、障害物の回避や着陸地点の安全確認は通常赤外線センサーもしくは操縦者の目視で行われているが、プラントのような複雑な施設では、見失い衝突または墜落の危険がある。そのため、墜落直前もしくは墜落中にドローンが撮影した画像・映像を利用して自律的に着陸地点の安全性を確認できる機能が求められる。

この問題に対し本研究では以前、物体検出ツールYOLO(You Only Look Once)[2]に加えて、エッジ抽出とSVMを利用して安全領域の判定を行った[1]。結果として精度の一番高いモデルはエッジ抽出処理後に安全領域をSVMで判定した手法であった。本稿ではその手法を、USBカメラを接続したシングルボードコンピュータ上に実装し、ドローンに搭載した。そのドローンを岩手県立大学構内で飛行させ、上空から撮影した動画に対して性能を評価する。

### 2. 研究概要

本研究ではプラントでの飛行を想定する。自動着陸のためにドローンに搭載された単眼カメラから取得した俯瞰動画に対し、YOLO, SVM, エッジ抽出を組み合わせて安全領域の検出を行っている。学習画像、検証画像ともにプラントの衛星画像を用いた場合、最も性能が良いモデルはエッジ抽出+SVMであり精度はPプラントの画像ではF値1.00, Qプラントでは0.91であることが確認できている。

本稿ではエッジ抽出+SVMに対して、衛星画像、大学構内実画像、衛星画像+大学構内実画像で学習した3つのモデルを利用し、大学構内で撮影した別の動画を評価対象として精度を比較する。

#### 2.1 安全領域

本研究では、安全領域と呼ばれるドローンの着陸地点の名称を定義する。安全領域の条件は以下の通りである。

1. プラントの施設が領域内に含まれてない。
2. 5m×5mの領域。
3. 車、人など衝突時の損害が大きいものが領域内に含まれてない。

以上の条件を満たす領域を候補として、プラント内の道路をクラス「道」として学習する。

### 2.2 ドローン飛行想定

ドローンがプラント内で飛行中、何らかの原因でプロペラが停止しパラシュートが展開した状況を想定する。気圧センサーから飛行高度を取得し、ドローンに搭載された単眼カメラを使用して俯瞰動画を取得する。

### 2.3 エッジ抽出+SVM

取得した画像にCanny法によるエッジ抽出を行い、さらに二値画像を取得する。二値画像に対して安全領域の候補を検出する。高度毎に決められた大きさの矩形領域を用意し、画像左上から領域の白の割合が3/28以下の場所を探索する。その後カラー画像の同じ範囲のHSLヒストグラムを特徴量としてSVMにより「道」かどうかを判定する。

### 3. 実動画による評価

#### 3.1 検証データ

検証に利用した動画は、岩手県立大学構内でカメラを接続したシングルボードコンピュータをドローンに搭載し、飛行させた上で撮影した俯瞰動画であり、高度35mで撮影した640×480pix, 2490フレームの動画である。

#### 3.2 正解データの作成

検証にあたり、検証動画に対してマスク画像を作成し、正解データとした。例として図1に示す。マスク画像はHSLの特徴量の色相を用い、芝生の部分が入らないように抜き取った後、モノフォロジ変換を用いて隙間を埋め、OpenCVのconnectedComponentsを使用することで作成した(図1(右))。

### 4. 検出実験

#### 4.1 評価方法について

SVMが判定した領域の全てがマスク画像の白の領域なら正解とする。Precisionは正解した領域数÷選択した領域数で表し、Recallは1つでも正解したフレーム÷正解が存在する総フレームで表す。複数フレームでのRecallは決められたフレーム数内で1つでも正解したフレーム÷決められたフレーム数内でマスク画像が存在するフレームの総数で表す。

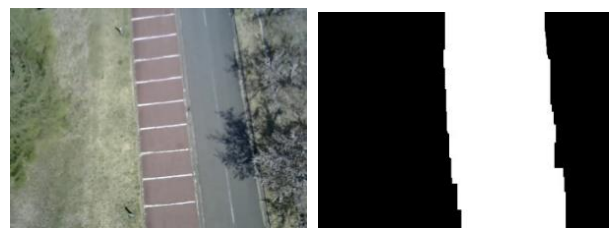


図1 動画の1フレーム(左) マスク画像(右)

<sup>†</sup> 岩手県立大学 Iwate Prefectural University

## 4.2 実験結果

実験で得られた Precision, Recall, F 値を図 2, 図 3 に示す. また表 1 に処理速度を示す.

図 2 より, 全体的な性能を考えると学習画像に大学構内の画像を使用したモデルが一番 F 値は高かったが, Precision を見ると大学構内の画像と衛星画像を使用したモデルが最も高かった. これは衛星画像のデータを使うことで大学構内の「道」を学習していたものから一般的な「道」を学習したと考える. それに対し衛星画像のみの結果は一番低い値となっている. これは衛星画像の「道」と大学構内の「道」は特徴が異なるためであると考え.

図 2 ではエッジ抽出の Canny 法の閾値を, 下限 50 上限 110 で実行していた. ここで Canny 法の閾値を下げ, より多くのエッジを検出することで確かな「道」を検出できると考え, Canny 閾値を下限 40 上限 50 で実行した. その結果が図 3 である. 全体的な性能を考えると学習画像に大学構内の画像を使用したモデルが一番精度は高かったが, Precision が他の学習モデルと比べると最も低い結果となった. 図 2 と比べるとエッジ抽出の段階で候補領域数が減り, 図 2 で検出できた領域が検出できなくなり, Recall が低下した. 向上すると考えていた Precision についてはエッジ抽出が関係ない領域で誤検出があると考えられる. しかし衛星画像のみの結果では Precision が向上しており, 期待した効果を得ることができた. 図 2, 3 の結果を考えるに, 実験環境のデータを学習していない場合においては, Canny 法の閾値を下げることによる効果はあると考える.

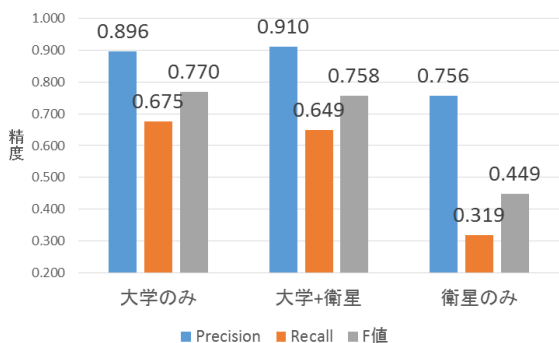


図 3 下限 50 上限 110 の 1 フレーム毎の結果

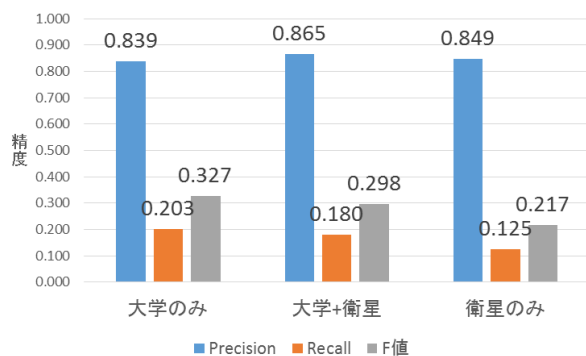


図 4 下限 40 上限 50 の 5 フレーム毎の結果

表 1 1 フレームあたりの処理速度の結果(平均)

	下限 50 上限 110	下限 40 上限 50
大学のみ	0.00598[sec]	0.00423[sec]
大学+衛星	0.00610[sec]	0.00428[sec]
衛星のみ	0.00541[sec]	0.00400[sec]

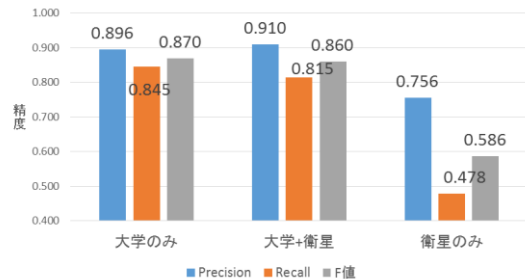


図 2 下限 50 上限 110 の 5 フレーム毎の結果

表 1 より Canny 法の閾値を下げることで, 多くのエッジが抽出され SVM の候補数が少なくなったことにより処理速度が向上した結果となった. 実際の飛行時は映像の取得があるため, より処理速度が遅くなる. 例えば 30FPS のカメラの場合は 0.033 秒かかるため, ここに処理を加えると約 0.04 秒となり 25FPS ほどの処理速度となる. ドローンのパラグライダーの展開時の落下速度を考えると判定に必要なフレーム数を増やす余地があると考え. 図 4 に図 2 の結果を 5 フレーム毎に検出した場合を示す.

図 4 より全ての場合で Recall 値が上昇している. 処理時間を 0.04 秒から 5 フレーム毎の 0.2 秒にする代わりに精度が上がることを考えると, 複数フレームでのより良い検出法を考えたい.

## 5. おわりに

本研究では学習画像に衛星画像, 大学構内画像, 衛星+大学構内画像を用い, SVM+エッジ抽出手法で安全領域の検出を行った. 実験の結果, 実動画に対しても検出が可能であった. 全体的な性能を考えると大学構内画像で学習したものが一番精度は高かったが, Precision だけで見ると衛星+大学構内画像で学習したものが最も精度が高かった. しかし, 1つの地点での動画だけだったため, 様々な地点での動画で評価する必要がある.

また, 利用できるフレーム数についても考察し, 現状は 5 フレームを利用したものが精度, 処理速度のバランスを考慮すると優れていると考える. それに伴い複数フレームの利用にあたり処理速度の高速化, 複数フレーム利用した領域抽出箇所の変更を今後の課題としたい.

謝辞

本研究を行うにあたり協力いただいた, 東亜非破壊検査株式会社様に感謝します.

## 参考文献

- [1] 江村 優吾, 他, “製油所におけるドローン空撮画像からの安全領域判定手法の比較”, FIT2020 第 19 回情報科学技術フォーラム, H-019 (2020).
- [2] J. Redmon, S. Divvala, R. Girshick, A. Farhadi, “You only look once: Unified real-time object detection”, Proc. of the IEEE on CVPR, pp. 779-788, (2017).