

ロボットビジョン技術を搭載したドローン実演機の試作と RTF でのフィールド 実証 (第3報)

Prototype of drone demonstration machine equipped with robot vision technology and field demonstration at
Fukushima Robot Test Field (3rd report)

南相馬技術支援センター 機械加工ロボット科 三浦勝吏 塚本遊 齋藤宏
電子機械技術部 電子・情報科 鈴木健司 ロボット・制御科 近野裕太 松本聖司
共同研究者 公立大学法人会津大学 コンピュータ理工学部 復興支援センター 山田竜平

ロボットビジョン技術である三次元構造復元、AI、5G 通信の3つの要素技術を搭載したドローン実演システムを試作し、福島ロボットテストフィールドにおいて実証試験を行った。実証試験では、5G 通信で映像を伝送し、伝送先で建物、瓦礫、道路、人を AI により領域検出することを確認できた。

Key words: ロボットビジョン、ドローン、AI、物体検出、領域検出、5G 通信、速度計測、VPN、映像伝送

1. 緒言

福島ロボットテストフィールド (以下、福島 RTF) は陸・海・空で活用するロボットの国内最大級の実証フィールドであり、無人航空機をはじめ自動走行車や水中ドローンなどの実証試験が行われている。特に無人航空機では、インフラ点検や気象観測など、産業利用に向けた実証試験が行われている。

本事業では、ハイテクプラザが保有する AI、5G 通信、三次元構造復元技術の3技術と市販ドローンを組み合わせた実演システムを構築し、福島 RTF の市街地フィールドで実証試験を行った。AI 技術では災害現場での家屋や人の発見、土砂エリアなどの検出を行い、災害現場の情報収集及び被害状況を分析する簡易型災害対応ドローンの実現を、三次元構造復元技術ではドローンが取得する映像からの建造物等の3D データ化を目的とした。また、ドローンで撮影した高精細映像や機械学習用の大容量データを遠隔にある災害対策本部と共有することを想定した映像伝送システムの開発も行った。

本事業は令和3年度から令和5年度の3年間、会津大学との共同研究として実施している。2年目までは、AI による物体検出、三次元構造復元、5G 通信それぞれの要素技術を統合し、ドローンの実演システムの試作に取り組んだ。今年度は4K 映像をリアルタイムに伝送できるよう伝送システムを構築した。また、実演システムに、YOLACT¹⁾ により救助者を発見できる機能や被害状況の分析を行う機能、GoogleEarth を用いた被災地情報の管理機能などを追加した。さらに、AI による物体検出及び領域検出、三次元構造復元、映像伝送すべてを組み合わせた実証実験を福島 RTF にて行い、実演システムの機能を検証した。

2. 実験と結果

2. 1. 映像伝送

2. 1. 1. 4K 映像の伝送

令和3年度の実験結果より、構築した VPN 装置では4K 映像をリアルタイムに伝送するためには通信速度が足りないことが分かった²⁾。そこで、VPN 装置を用いずに YouTube のライブ配信機能を用いることで4K 映像を伝送することとした。

映像伝送には、4KカメラをPCに取り込むため、カメラ映像を録画や配信可能にするキャプチャーデバイス (AverMedia 社製 LIVE Gaming) を用いた。PC から5G ルータを用いて4K 映像をアップロードし、インターネットに接続されている別の PC から映像を確認することで4K 映像を遠隔地でリアルタイムに確認できる。4K 映像を30[fps]で配信するにはアップリンクで30[Mbps]の通信速度が推奨されている³⁾が、令和3年度の研究において、5G 通信可能な docomo ルータでは通信速度を満たしていることを確認している²⁾。実際に4Kカメラの映像をLIVE配信したところ、問題なく配信できることを確認した。4K 映像配信の様子を図1に示す。

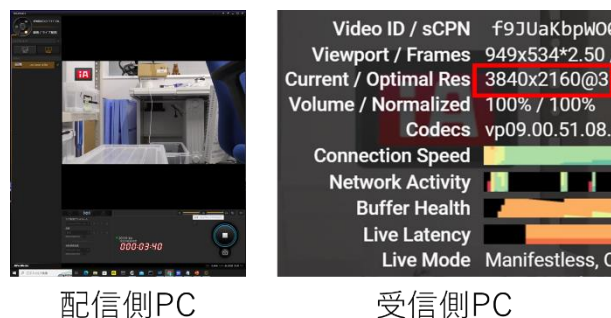


図1 YouTubeによる4K映像LIVE配信の様子

2. 1. 2. docomoMEC の通信速度計測

4K 映像を伝送するための通信速度を確保する他の手段として、docomoMEC⁴⁾ (Multi-access Edge Computing) を用いることで映像伝送が可能か通信速度の測定を行い検討した。docomoMEC とは docomo が提供している MEC システムである。MEC とはスマートフォンなどのモバイル端末や IoT 機器などからのアクセスを考慮して、ETSI (欧州電気通信標準化機構) が標準化を進めるエッジコンピューティングに関する規格で、エッジ端末でそれぞれ分散して処理を行うため、応答時間を削減でき、より低遅延の通信が可能になる。

この docomoMEC サービスを用いて 5G 端末間で P2P 接続を行い、通信速度を計測した。通信速度の計測はソフトウェアの JPerf を使用した。JPerf は速度計測ツール iPerf⁵⁾ を GUI で操作できるソフトウェアであり、クライアント側からサーバ側へ、指定したデータ量を指定した時間に連続して送信することで、ネットワークの帯域幅や実効速度を計測することができる。通信はストリーミング配信などの映像伝送で使用される UDP 通信について計測した。計測は 10 回ずつ行い、その平均値を記録した。結果を図 2 に示す。構築した VPN 装置による通信速度と比べて速い速度で通信できることを確認した。

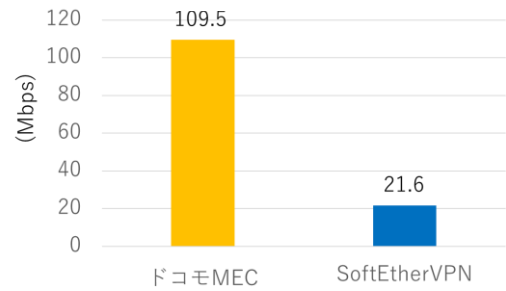


図 2 docomoMEC の通信速度

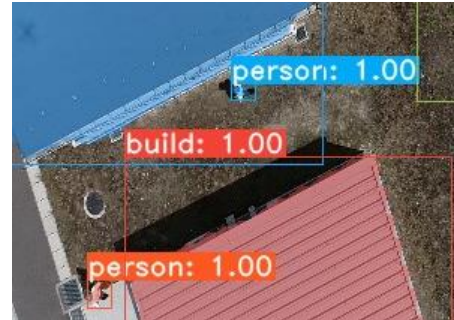


図 3 YOLACT による人の検出

2. 2. YOLACT による領域検出

2. 2. 1. 検出物の追加

昨年度にはリアルタイムでセグメンテーションを行う YOLACT を利用して、建物 (building)、道路 (road) 及び瓦礫 (rubbles) の 3 クラスの領域検出を可能とする AI 領域検出システムを開発した⁶⁾。今年度は 3 クラスに加え、人 (person) を検出できるよう再度学習を行った。構築した学習モデルを用いて福島 RTF 市街地フィールドで実証試験を行った。ドローンは DJI 社製 Phantom 4 Pro V2.0 を使用した。ドローンのカメラを真下に向けた状態で、高度約 30[m] を飛行させた。YOLACT の実行結果を図 3 に示す。高度 30[m] 地点では人を検出できることを確認した。

表 1 開発環境

CPU	Core-i9 10920X
メモリ	128GB
GPU	Geforce RTX 2080Ti
学習モデル	YOLACT
入力画像の解像度	640×480

2. 2. 2. 学習モデルの修正

前年度構築した学習モデルは道路や建物が斜めに映ったときの検出の精度が悪かったため、追加学習を行い修正した。学習した計算機の仕様を表 1 に示す。学習には解像度 640×480 の画像を使用し、アノテーション作業を済ませた画像を約 250 枚用意して学習した。画像は福島 RTF の市街地フィールドでドローンを飛行させ収集したものをを用いた。学習には表 1 の環境で約 10 時間を要した。修正結果を図 4 に示す。画像を比較すると修正前に道路や建物のエッジからはみ出して領域検出されていたものが、修正後には道路や建物のエッジの通り領域検出されていることが分かる。

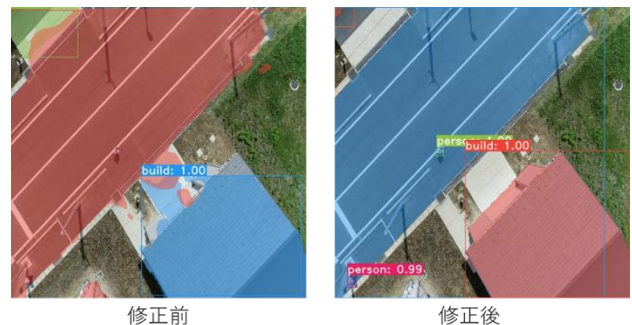


図 4 YOLACT 修正結果

また、図 4 の画像において、正しい領域検出の結果との一致率を IoU (Intersection over Union) を用いて比較した。IoU は、物体検出や領域分割のタスクにおいて、予測された領域と正解領域の一致度を測る指標の一つで、(1) 式のように表される。この時、 A_{true} は正しい領域検出部分で A_{pred} は予測した領域検出部分を示す。IoU が 1 に近いほど、比較した 2 つの検出結果が一致していることを示す。

$$IoU(A_{true}, A_{pred}) = \frac{|A_{true} \cap A_{pred}|}{|A_{true} \cup A_{pred}|} \quad (1)$$



図5 市街地フィールドのオルソ画像

結果は修正前が 0.831 に対して修正後は 0.968 であった。この時の正しい画像とは手動でアノテーションしたものを用いて正しい領域検出結果とし、領域検出した部分をそれぞれマスク画像に変換し比較している。

2. 3. 三次元構造復元

2. 3. 1. オルソ画像の構築

前年度には SfM (Structure from Motion) と MVS (Multi View Stereo) の技術を用いて福島 RTF 市街地フィールドの三次元構造を復元した⁶⁾。

今年度はドローンから取得した映像から、オルソ画像を構築した。オルソ画像とは、写真上の像の位置ズレをなくし、真上から見たような傾きのない正しい大きさと位置に表示される画像に変換した画像である。オルソ画像は位置、面積及び距離などを正確に計測することができる。福島 RTF 市街地フィールドのオルソ画像復元結果を図5に示す。復元にはドローンで撮影したフル HD の画像約700枚を用いた。オルソ画像を構築することによって、被災した現場の位置や距離、面積を測定することができる。

2. 4. 実演システムの機能追加

2. 4. 1. 実装済みの機能

前年度には、VPN 装置と ROS を用いて、ドローンの映像を遠隔地に伝送し、伝送先で AI による物体検出を行うシステム及び伝送された映像からリアルタイムに復元用の画像を生成するシステムを構築した⁶⁾。

このシステムを用いることでドローンの映像をリアルタイムに伝送し、伝送先である遠隔地から被災地の状況確認や分析を行うことができるが、新たに、ドローンの映像から救助者を発見し知らせる機能や現場の被災状況を分析し情報を管理する機能を追加した。

2. 4. 2. 救助者発見機能

2. 2. 1. 項で示したように YOLACT で人 (person) の検出が可能となった。そこで、YOLACT の表示プログラムを改良し、人が検出された場合は「Rescuee found!!」の文字が画面に表示されブザー音を鳴らして知らせてくれるシステムを構築した。人が検出され



図6 人を検出した際の YOLACT の表示結果



図7 人を検出した際の YOLACT の表示結果
オルソ画像の領域検出結果 (左上) マスク画像 (右上)
被害状況の分析結果 (下)

た際の YOLACT の様子を図6に示す。これにより、この実演システムを搭載したドローンが被災地を飛行した場合、リアルタイムで AI が人を発見し、監視者に知らせることができる。

2. 4. 3. 被害状況の分析機能

2. 3. 1. 項で構築したオルソ画像を YOLACT により領域検出することによって、構築した範囲全体を道路、建物瓦礫ごとの領域に分けることができる。そこで、YOLACT で検出した領域をもとに、瓦礫の箇所や瓦礫の領域の割合を算出するプログラムを構築した。図7に瓦礫の領域の割合を算出の様子を示す。オルソ画像に YOLACT を行い (図7中左上の画像)、検出された領域のマスク画像 (図7中右上の画像) を抽出し、そのうち瓦礫の部分の割合を算出する。マスク画像の赤色で示した部分が瓦礫の領域である。市街地フィールドでは瓦礫の箇所が2か所あり、建物、道路、瓦礫として検出された領域のうち、約 16.8[%]が瓦礫であった。オルソ画像は映した場所の距離や位置、面積を測定できることから、YOLACT と組み合わせることによって、被災状況の範囲などの解析を行うことができる。

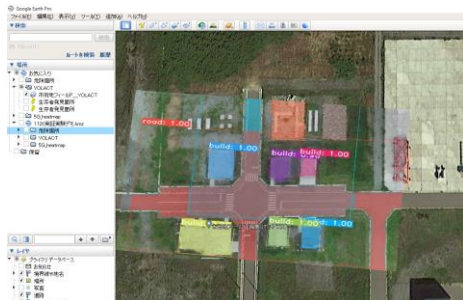


図8 GoogleEarthを用いた情報管理
領域検出結果(上) 危険箇所などの記録(下)

2. 4. 4. GoogleEarthを用いた被災地情報の管理

ドローンで得られた情報を管理する手段としてGoogleEarthを用いた。GoogleEarthのレイヤ機能を用いることで地図上に様々な情報を重ねて表示することができる。一例を図8に示す。図8の上の画像は2. 4. 3. 項で示したオルソ画像で領域検出した画像をGoogleEarthに埋め込んだものである。このように領域検出の結果を定期的に記録しておくことで災害の前と後の変化を重ねて表示し比較するといったことが可能である。また、図8の下画像のようにドローン映像から確認できた危険箇所や生存者発見箇所などといった特記事項をGoogleEarth上にプロットして記録することも可能である。そのほか、ドローンの飛行経路や、1年目に計測した5G通信速度のヒートマップ²⁾などもGoogleEarth上に記録することができる。

2. 5. 実演システムの実証試験

2. 5. 1. 福島RTF市街地フィールドでの実証試験

2. 4. 項で記した実演システムを用いて実証試験を行った。実証試験は福島RTF市街地フィールドの上空をドローンが飛行し、ドローンの飛行経路を福島RTF研究棟からオペレートするというものである。市街地フィールドにはドローン操縦者や撮影補助者、救助者役の職員を配置し、研究棟には伝送されたドローンの映像を見ながらドローンのオペレートを行う職員を配置して実験を行った。ドローンの映像伝送は構築した実演システム⁶⁾を用いて行い、操作の指示などの情報伝達はTeamsのビデオ会議機能を用いて行った。システムのイメージを図9に示す。また、実際の研究

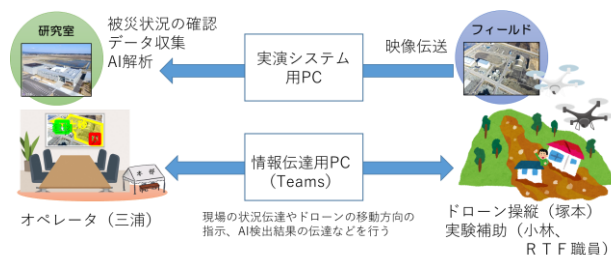


図9 システムのイメージ

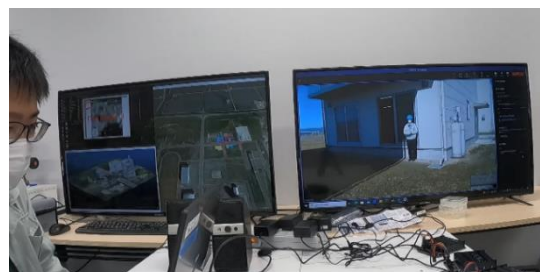


図10 研究棟でのモニターの様子

棟での様子を図10に示す。図10に示すようにモニターから市街地フィールドの様子を確認でき、操作の指示やGoogleEarth上に情報を記録できる。

実験の結果、ドローンの映像がリアルタイムに研究棟まで伝送されていることを確認した。伝送された映像を見ながら、ドローン操縦者にドローンの高度や進行方向の指示を行ったところ、支障をきたすことなく情報伝達でき、ドローンの操縦を行うことができた。

また、被災地での使用を想定し、市街地フィールドに救助者役の職員を配置し、研究棟から救助者を確認できるか実験を行った。結果、YOLACTにより人を検出し、研究棟から救助者役の職員を確認することができた。この実験をドローンの高度を変えて行ったところ、高度40[m]までは、人を検出することができた。

さらに、伝送された映像からリアルタイムに三次元復元用の画像を生成するシステムも問題なく動作していることを確認した。これにより復元処理に時間のかかる作業を伝送先で行うことができる。

3. 考察

3. 1. 映像伝送

YouTubeのライブ配信機能を用いることで4K映像を伝送することができたが、現在は5G(Sub6帯)の上空利用は認められておらず⁷⁾、このシステムをドローンに搭載し、実験を行うことができなかった。5Gの上空利用も検討段階にあるため、上空利用が可能になった際は実証試験を行っていきたい。

また、docomoMECの通信速度の評価を行った結果、UDP通信の速度が109[Mbps]であった。この結果は4K映像の配信に必要な通信速度よりも早く、十分映像伝送システムに有用だと考えられる。こちらも5Gの上空利用が可能になった際は実証試験を行っていきたい。

3. 2. YOLACT による領域検出

YOLACT を利用して、建物 (building)、道路 (road) 及び瓦礫 (rubbles)、人 (person) の 4 クラスの領域検出を可能とした。今回「人」を検出するよう学習を行ったが、学習データの収集は、市街地フィールドに職員を配置し、各々適当に移動した様子をドローンで上空から撮影することで収集した。ドローン飛行時はヘルメットの着用が必須であるためヘルメットを着用してデータ収集を行ったため、ヘルメットを着用していない場合、「人」と検出するか検証する必要がある。人と検出できなかった場合、上空から見た人の映像をどのように集めるかが課題である。

3. 3. 三次元構造復元

ドローンで撮影した画像から市街地フィールドのオルソ画像を構築することができた。復元には約 1 時間かかったため、復元時間の短縮が課題である。

3. 4. 実演システムの機能追加

YOLACT により救助者を発見できる機能や被害状況の分析を行う機能、GoogleEarth を用いた被災地情報の管理機能などを追加した。これらの機能を用いることでドローンによる災害現場での捜索や分析の一助とすることができる。今後もニーズ調査を行い、機能のブラッシュアップを行っていききたい。

3. 5. 実演システムの実証試験

実演システムの実証試験を福島 RTF 市街地フィールドで行い、ドローンの映像をリアルタイムに遠隔地に伝送できることを確認した。また、YOLACT により伝送先から「人」を検出できることを確認した。実演システムを用いることでドローンの映像を遠隔地に伝送できるが、この伝送システムは 5G 通信または LTE 通信が可能な場所でなければ使用できないため、通信できない地域での映像伝送が課題である。

4. 結言

3 年目は、4K 映像の伝送方法の検討や YOLACT の追加学習を行った。また、実演システムの機能を追加し、福島 RTF 市街地フィールドで実証試験を行った。

映像伝送では、YouTube のライブ配信機能を用いることで 4K 映像を伝送できることを確認した。また、docomoMEC を用いて P2P 通信した時の通信速度を計測し、前年度構築した VPN 装置による通信速度と比べて速い速度で通信できることを確認した。

YOLACT による領域検出では、建物 (building)、道路 (road)、瓦礫 (rubbles) に加え、人 (person) を検出できるよう再度学習した。また、検出精度を上げるため、追加学習した結果、正しい領域検出の結果と

の一致率を比較すると修正前の IoU が 0.831 に対して修正後の IoU は 0.968 と精度を上げることができた。

三次元構造復元について、ドローンで撮影した画像から市街地フィールドのオルソ画像を構築することができた。

ドローンの実演システムに、ドローンの映像から救助者を発見し知らせる機能や現場の被災状況进行分析し情報を管理する機能を新たに追加した。このシステムを用いて RTF 市街地フィールドで実証試験を行った結果、ドローンの映像がリアルタイムに伝送先である研究棟まで伝送されていることを確認した。また、YOLACT により人を検出し、ブザー音で知らせてくれることによって、伝送先から人を発見することができた。

3 つの要素技術 (三次元構造復元、AI、5G 通信) を搭載したドローン実演システムを試作し、ドローンの高機能化を図ることを目的に 3 年間研究を行ってきた。「ドローンの映像を AI で物体検出及び領域検出する」、「ドローンの映像から三次元構造を復元する」、「5G 通信を用いてドローンの映像をリアルタイムに遠隔地に伝送する」といったシステムを開発し、RTF 市街地フィールドで実証することができた。

これまで、ハンズオンセミナーを開催し 3 つの要素技術について技術普及を行ってきた。令和 5 年度をもって本課題は終了となるが、今後も継続してセミナーや講習会などを開催し、技術普及に努めていきたい。

参考文献

- 1) Daniel Bolya, Chong Zhou, Fanyi Xiao, and Yong Jae Lee. Yolact: real-time instance segmentation. In ICCV, 2019
- 2) 太田悟, 他「ロボットビジョン技術を搭載したドローン実演機の試作と RTF でのフィールド実証 (第 1 報)」。福島県ハイテクプラザ, 2022
- 3) “YouTube ヘルプ YouTube でライブ配信する”, https://support.google.com/youtube/answer/2853702?hl=ja&ref_topic=9257892&sjid=6476660956880835367-AP (参照 2024-2-22)
- 4) “docomo 公式ウェブサイト”, <https://www.mec.docomo.ne.jp/index.html> (参照 2024-2-22)
- 5) “iPerf プロジェクト 公式ウェブサイト”, <https://iperf.fr/> (参照 2022-2-22)
- 6) 三浦勝吏, 他「ロボットビジョン技術を搭載したドローン実演機の試作と RTF でのフィールド実証 (第 2 報)」。福島県ハイテクプラザ, 2023
- 7) “総務省電波利用ホームページ”, <https://www.tele.soumu.go.jp/j/sys/others/drone/uv/> (参照 2022-2-22)