



《第71回》室内環境理解とその医療分野への応用

杉 山 治

1. はじめに

学生時代、私は国際電気通信基礎技術研究所 (ATR) でインターンシップを行っており、コミュニケーションロボットの研究に携わっていた。その研究の一つとして、人とロボットの自然な対話を実現するための環境理解があり、この環境理解の技術には、レーザーレンジファインダを用いた SLAM (自己位置推定と環境地図作成) や、加速度計を搭載したセンサデバイスを用いた人間の行動予測が含まれていた^(注1)。当時のレーザーレンジファインダは、室内の正確な地図を作成するために不可欠なものであったが、その機器は非常に大きく、高価であるという課題があった。特に、大規模な環境をカバーするためには、複数の機材が必要となり、導入や運用におけるコストが非常に高くなるという制約があった。ポータブルなセンサデバイスとしては加速度センサを搭載したモジュールが、スマートフォンが登場する以前に実用化されていた^(注2)。このモジュールは持ち運びが可能なサイズで、加速度計で計測された時系列データ等から看護師の行動理解に活用されていた。加速度センサだけでは、動作の細部や環境との相互作用を完全に把握することが難しく、解釈には限界があったが、この技術は当時の病院環境において、看護師の動きをある程度理解するために貢献していた^(注3)。10年以上の年月が経ち、私は研究分野をコミュニケーションロボットから音環境理解へと変え、その後、医療情報の研究に携わることとなった。興味深いことに、10年以上前に存在した病院における室内環境理解のニーズは過去から現在に至るまで一貫して存在しており、むしろ現代の病院ではそのニーズがさらに多様化していた。医療現場における

効率の向上や安全性の確保のためには、環境理解技術の進化が不可欠であり、最新の技術をどのように活用できるかが重要な課題となっている。

2. 室内環境理解と病院 BCP

私が病院に配属されて間もない時期、2016年に熊本地震が発生した。その後、耐震工学や建築分野の専門家を含む多職種からなる調査チームが結成され、被災地で唯一の大学病院や病院避難を余儀なくされた医療機関を中心に、地震によって大きな被害を受けた10施設の病院建物の聞き取りを行った¹⁾。私も医療情報の担当者として調査に加わった。調査の結果、対象となった9つの病院のうち、4つの病院では構造上の安全性に懸念があるとして避難が行われた。しかし、その後の専門家による詳細な解析の結果、避難が行われた4つの病院のうち3つは、実際には問題がなかったことが判明した²⁾。このような必要がなかったかもしれない避難決定は、地震直後の混乱や限られた情報に基づく判断によるものであったと考えられる。もし、当時の現地スタッフがスマートフォンなどで撮影した写真や映像を基に病院の機能評価ができていたならば、不必要な避難による病院の経済的損失を回避できた可能性がある。その後、2017年には、災害拠点病院に事業継続計画 (Business Continuity Plan, BCP) の策定が義務化され、災害時にどのように事業を継続していくかという点について計画を策定し取り組むことが必要となった。

調査から得られた教訓をもとに、現在の室内環境理解技術を導入することで、同様の課題に対処できないだろうか考えた。15年前と現在の技術的な違いは、深層学習モデルの発展により、画像解析技術が格段に進化している点である。本研究では、セマンティックセグメンテーションの技術を用いて、建物の構造部材を特定し、それらが被災しているかどうかを判定するアプリケーションの開発を試みている。セマンティックセグメン

^(注1) <https://www.atr-p.com/products/HumanTracker.html>

^(注2) <https://www.atr-p.com/products/sensor.html>

^(注3) https://dil.atr.jp/ISL/research_project-2/project_archive/e-nightingale

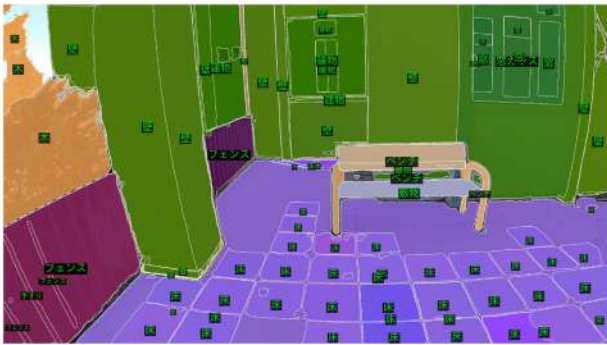


図1 トルコ地震で被災した建物のセグメンテーション
Segment Anything で領域抽出した画像。建築構造部材や周辺オブジェクトが細かく抽出されている。

ーションとは、画像内の各ピクセルがどのクラスに属するかを判定する技術であり、物体や構造物の詳細な識別が可能となる。2022年までに、セマンティックセグメンテーション技術は大量のオープンデータを用いて進化を遂げ、建築構造部材を含むさまざまなオブジェクトの検出が可能となっていた。さらに、2023年にはMeta社から“Segment Anything”^(注4)が発表され、ほぼ完全な領域抽出まで可能となった。図1はトルコ地震で被災した建物の入り口を撮影した画像にSegment Anythingを適用したもので、建物の各部材が領域も含めて抽出されていることがわかる。

本研究では、OpenAI社が発表したCLIP^(注5)を発展させたCLIPSeg³⁾を用いて、撮影された画像中の建築構造部材部分を抽出し、さらにその中で被災箇所を特定することを試みた。CLIPSegは、ゼロショットでセグメンテーションを行うことができるモデルであり、図2のように建築構造部材領域を抽出し、その領域を再度セグメンテーションすることで壁のひび割れや天井の崩落といった被災部分を抽出することが可能である。画像に多数のオブジェクトが映り込んでいる場合、セグメンテーションの精度が著しく低下することが確認されており、建築構造部材の判別を行い、それぞれの部材に焦点を当てて画像をクリッピングすることで建物被害をより正確に抽出できることが現在までの検証で明らかになっている。このように、画像解析技術の進展により、建物の被害を特定するタスクが専用の機材を必要とせず画像を撮影するだけで、ゼロショットモデルによって取り扱うことが可能となった。

3. 今後の医療現場における室内環境理解

本稿では、室内環境理解における過去と現在の技術を比較し、画像処理技術のこの10年間の飛躍的な進展により、従来は特殊な機材が必要とされていた室内環境理解が、スマートフォンに搭載されたカメラで撮影した映像

(注4) <https://segment-anything.com/>

(注5) <https://openai.com/index/clip/>

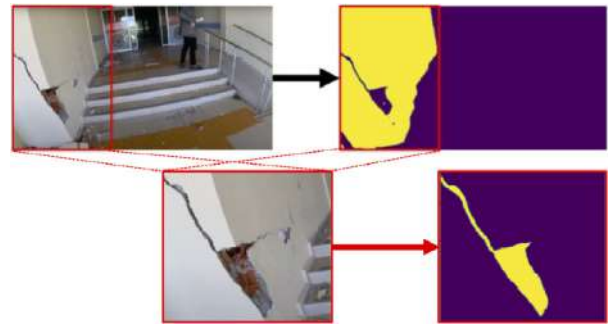


図2 被災箇所の抽出

CLIPSegに“Wall”(壁)というプロンプトを与え、画像中の壁の領域をトリミング、“damaged”(被災した)というプロンプトで被災箇所を抽出する。

だけで実現可能となる一例を示した。この技術の発展により、従来のように高価な専用機材を用いなくても、誰でも手軽に高度な環境理解を行えるようになったこと、また、さまざまな場面で画像認識ができるゼロショットモデルが公開されてきたことは、非常に画期的なことであり、これからさまざまな分野で応用が可能になると考えられる。前章で紹介したセマンティックセグメンテーションはその一部に過ぎず、たとえば、画像キャプションといった技術は、画像のコンテキストを文章で説明する能力をもつため、冒頭に言及した医療従事者(看護師)の行動記述にも役立つと考えられ、これらの画像処理技術が普及することで、災害時の迅速な状況把握や、病院内での業務支援システムへの応用が期待される。

(2024年9月10日受付)

参考文献

- 1) 石原 由菜ら：地震災害時の医療機能に対する影響評価の試み—大阪北部地震の影響分析から—, 日本災害医学会雑誌, 27-3, 195/200 (2022)
- 2) M. Shimoto, K. Cho, M. Kurata, M. Hitomi, Y. Kato, S. Aida, O. Sugiyama, N. Maki, and S. Ohtsuru: Hospital Evacuation Implications after the 2016 Kumamoto Earthquake, *Disaster medicine and public health preparedness*, 16-6, 2680/2682 (2022)
- 3) T. Lüddecke and A. Ecker: Image Segmentation Using Text and Image Prompts, In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 7086/7096 (2022)

[著者紹介]

すぎやま おさむ
杉山 治君

2009年慶應義塾大学大学院理工学研究科後期博士課程修了。修士修了後、ソニー株式会社にてパソコン・テレビのソフトウェア開発に従事し、博士取得を機に国際電気通信基礎技術研究所(ATR)で研究員として研究分野に転向、コミュニケーションロボットの研究に携わる。その後、東京工業大学の研究員を経て、京都大学医学部附属病院にて、機械学習・IoTの技術を医療に応用する研究を行う。現在、近畿大学情報学部准教授として医療情報の研究に携わる。

E-mail: sugiyama@info.kindai.ac.jp

所属：近畿大学情報学部 大阪府東大阪市小若江3-4-1