

# 奈良先端大において取り組みたい研究テーマについて

試験区分: 情報科学区分

氏名: 梶田修慎

希望研究室: 生体医用画像研究室

現在の専門: 深層学習

## 1 これまでの修学内容

学部では1年次に線形代数学・解析学など、専門的な学習を始めるために必要な基礎を学び、2年次からは人工知能基礎・応用、データベース、自然言語処理など専門的な講義が開講され、積極的に受講した。そしてこの時期から、将来取り組みたい分野を見つけるきっかけ作りとして、資格取得に向けて学習をはじめ、基本情報技術者試験・応用情報技術者試験を取得した。4年次の研究室配属では第一志望であった、脳型情報処理や人工知能を専門としている研究室に配属され、Vision Transformerを用いた画像解析の研究をしている。

## 2 取り組みたい研究テーマ

私が、奈良先端大において取り組みたい研究テーマは「Self Attention を用いた生体医用画像のマルチモーダルセグメンテーション」である。本稿では、この研究テーマの研究背景、研究課題、提案手法、期待される結果について述べる。

## 3 研究の背景

近年、ディープラーニング技術は実務レベルの利用が急速に拡大しており、医療分野にも多く応用されている。中でも「診断支援」の領域では、ディープラーニングの持つ特徴抽出能力を用いることで、生体医用画像から病変のある箇所を抽出する技術が注目されており、コンピュータ診断支援 (Computer-Aided Diagnosis : CAD) と呼ばれている。この技術は「医師への負担を緩和させ、病変の見落としリスクを低減させる」ことが目的であるが、病変箇所と解剖学的構造物が重なった場合や、病変の規模が小さい場合は抽出することが困難である。そのため、誤診断や医師による二重チェックなどが発生し、根本的な解決には至っていない。

病変部位を抽出する手法としてセグメンテーションが用いられることが多く、最近ではU-netを利用した研究が頻繁に行われている [5]。U-net では、畳み込み層

を重ねることで失われていく画像構成要素の位置情報を、Skip Connection によって保持することができるが、セグメンテーションしたい対象が重なっている場合や複数ある場合に効果を発揮しづらい。その問題を解決するために、インスタンスごとにセグメンテーションを行うモデルが開発され、Mask R-CNN と呼ばれている [1]。魚住らの研究によると Mask R-CNN のモデルは小児胸部 X 線画像における肺領域の自動抽出において U-net より低い抽出精度を示していたものの、肺炎による炎症の広がりがある症例において正確に肺野を抽出していた [4]。このことから、Mask R-CNN の抽出精度を向上させることができれば、過剰抽出や欠損が無いロバスト性の高いモデルができると予想される。

## 4 研究課題

上述の通り画像による診断支援を実務で利用するには、精度の向上を目指すべきである。そのためには先に記した通り、解剖学的構造同士のリニアリティや、病変部位との重なりによる誤診断をどのように解決するかを考える必要がある。

## 5 提案手法

私は二つの視点から医用画像を解析する手法 (図 1) を提案する。一つ目は自然言語処理の Self Attention [3]。二つ目は画像処理の Mask R-CNN。これらを組み合わせ、文章と画像双方からアプローチを行うことにより、正確に病変セグメンテーションができると私は考える。

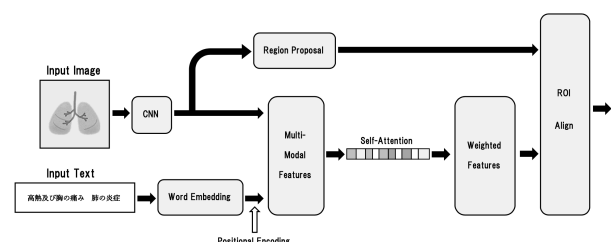


図 1: マルチモーダルなセグメンテーションの流れ

図1はMask R-CNNのアーキテクチャの一部を抜粋し改変したものである。

ここでは、画像と文章の入力から、すでに提案されているMask R-CNNのアーキテクチャに接続するまでの処理を説明する。初めに、患者の症状や想定される疾患を文章として与えてWord Embedding[6]により固有のベクトルに落とし込み、Positional Encodingにより位置情報を付加する。これにより得られたベクトルを特徴量として、CNNで得られた特徴量と組み合わせ、Multi modal featuresとして扱う。次はMulti modal featuresを元にSelf Attentionを求める。図1の例では、入力文章として「高熱及び胸の痛み 肺の炎症」を与え、それぞれのSelf Attentionを求めている。この場合Attentionは「肺」、「炎症」などで大きくなり、「及び」、「の」で小さくなると考えられる。以上により得られたAttentionを元の特徴量に掛け合わせてWeighted featuresを算出。また、RPN (Region Proposal Network) [2]によって得られたRegion ProposalをWeighted featuresと共にROI Align (Region Of Interest Align) に入力し特徴量を得る、という流れだ。その後、通常のMask R-CNNと同様に、Mask branchとFully connected layerを経てセグメンテーションされた出力が得られる。

## 6 期待される結果

私が提案するSelf Attentionを用いた手法は、CNN単体で特徴量を抽出した場合に比べ、病気に対する情報を多く保持していると予想されるので、過剰抽出や欠損の少ない結果が安定して得られると考える。また、Self Attentionを求めていることから、画像・文章のどこに注目しているかを可視化できるので、患者に対する病症の説明も比較的容易になると予想される。

## 7 最後に

本稿では、私が奈良先端大で取り組みたい研究テーマである「Self Attentionを用いた生体医用画像のマルチモーダルセグメンテーション」の研究概要について述べてきた。私が奈良先端大を志望するのは、医用画像工学を学べる数少ない大学院の一つであり、学部を持たず

に様々なバックグラウンドをもった人間を受け入れていて、そのサポートも充実しているからだ。また、全ての講義が英語で行われており、早期からグローバルに活躍するための力を養うことができ、同時に高度な技術力を身につけられる環境が備わっているのも志望動機のひとつである。

## 参考文献

- [1] Kaiming He, Georgia Gkioxari, Piotr Dollár, and Ross Girshick. Mask r-cnn. In *Proceedings of the IEEE international conference on computer vision*, pp. 2961–2969, 2017.
- [2] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster r-cnn: Towards real-time object detection with region proposal networks. *arXiv preprint arXiv:1506.01497*, 2015.
- [3] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Lukasz Kaiser, and Illia Polosukhin. Attention is all you need. *arXiv preprint arXiv:1706.03762*, 2017.
- [4] 魚住春日, 松原尚輝, 寺本篤司, 日木あゆみ, 本元強, 河野達夫, 齋藤邦明, 藤田広志. Mask r-cnn を用いた小児胸部 x 線画像における肺領域の自動抽出. *Medical Imaging Technology*, Vol. 38, No. 3, pp. 126–131, 2020.
- [5] 近藤堅司, 小澤順, 清野正樹, 藤本真一, 田中雅人, 安達登志樹, 伊藤春海, 木村浩彦. U-net を用いた胸部 x 線画像からの解剖学的構造の領域抽出. 人工知能学会全国大会論文集 第 32 回全国大会 (2018), pp. 2J404–2J404. 一般社団法人 人工知能学会, 2018.
- [6] 堅山耀太郎. Word embedding モデル再訪. オペレーションズ・リサーチ, Vol. 11, pp. 717–724, 2017.