# 生成AI技術を用いた肝細胞癌における腹部CT画像診断支援システムの開発

高橋健吾
東北大学大学院 医学系研究科 医科学専攻 医用画像工学分野 博士後期課程

2024年11月

## 当該分野の状況

### 臨床背景

肝細胞癌（HCC；Hepatocellular carcinoma）は，世界で腫瘍による死亡原因の第4位であり，原発性肝癌の75%～85%を占める[1]．HCCは進行期疾患として診断されることが多く，米国肝臓病学会や欧州肝臓学会のガイドラインによれば，血管浸潤は進行期HCCの特徴的な所見とされる．血管浸潤は，腫瘍細胞の播種とそれに伴う再発の危険を高めるだけでなく，術前の治療戦略にも影響を及ぼすため臨床的に非常に重要な危険因子と言える．今日では，HCCの血管浸潤を予測する人工知能（AI；Artificial Intelligence）研究が，罹患率の多い東アジアや欧米を中心に研究されている．しかし，乳癌や肺癌など他疾患に比べ，HCCを含む肝疾患はデータ数に限りがあるため，臨床意思決定支援システムとして血管浸潤を予測するAI研究が不足している現状がある[2]．さらに，既存のAI手法では疾患の不均一性を含む，臓器の解剖学的特徴に十分注目した学習ができていなく，腫瘍周辺に影響を及ぼす血管浸潤の効果的な学習手法の開発が求められる[3]．

### 技術背景

Self-Attention機構[4]によるTransformerは，生成系AIの分野で成功を収めた革新的AI技術である．OpenAI社により開発されたAI言語サービス「Chat GPT」は，Transformerを応用することで人間社会の創作活動の領域にまで影響を及ぼしている．近年では，Google社がTransformerを画像に応用したVision Transformer（ViT）を構築し，画像認識分野において従来型AIモデルであるConvolutional Neural Networks（CNN）よりも有望な結果を収めた[5]．ViTの大きな特徴は，元の画像を分割し，Self-Attention機構を用いた複雑な計算処理（行列の内積やsoftmax関数による非線形処理，Skip connection）によって分割画像全ての類似度をベクトル化し位置関係を定量化していることである．しかし，Self-Attention機構は注目領域と背景領域の類似度をも計算することで，大きな計算コストに見合わず注目領域に焦点を当てた学習に繋がっていない可能性がある．この課題は，文章の単語の位置情報を網羅的に学習するという，Transformer本来の計算アルゴリズムに起因している．

## 研究目的

Self-Attention機構によるViTの限界を解決すべく，Microsoft社はRetentive Network（RetNet）と呼ばれるAI技術を開発した[6]．RetNetには，Self-Attention機構にはない2つの構造的特徴がある：類似度行列式と再帰表現処理．現在，Manhattan距離に応じた加重因果マスクが自然画像の物体認識性能に優れており，その有効性はGrad-CAMと呼ばれる推論根拠の可視化処理によって証明されている[7]．医療画像においても，物体認識性能向上に伴う高い診断精度はAIの臨床導入に向けた一歩となる．本研究の目的は，肝臓の形態的特徴に応じた因果マスクを数理モデリングし，造影CT画像のHCCにおける血管浸潤を術前に予測するRetNetモデルを構築/検証することである．

## 研究手法

多重対角特性をもつ巡回行列や幾何学行列により肝臓の構造学的特徴を表現した因果マスクを開発する［右図］．さらに，cosine変換やRELU関数を用いた非線形処理を組み合わせることで，多様な強調表現をモデリングしていく．また申請者は，医療画像という限られたデータセットに対して効果的な学習・推論を行うため，「Sequence Pooling」と「Convolutional Embedding」を組み込む．これらの手法により，画像全体の位置情報を用いた予測が可能になるだけではなく，視覚パターンの抽出能力が強化され，標準のViTにおける線形写像では捉えきれない低次元特徴量を学習に用いることが可能になる．因果マスクにこのような手法も組み合わせることで，小規模データに対する脆弱性を克服したAI学習の実現を目指す．

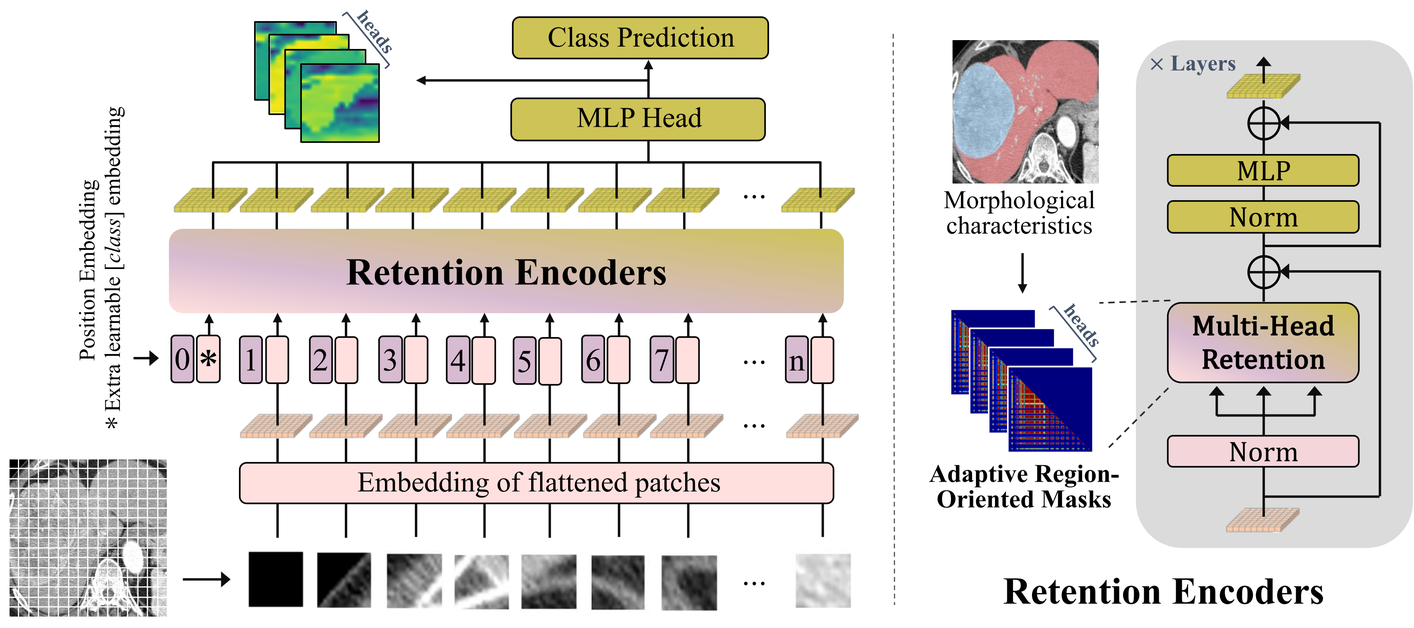


**Figure.** Proposal architecture

## 本研究の位置付け

現在の医療画像解析における大きな課題の一つは，AIが臓器や疾患に焦点を当てた学習を十分に実現できていないため，その診断結果の根拠を示すことが難しく，解釈可能性が欠如していることである．世界保健機関（WHO）が定めた「Ethics and governance of artificial intelligence for health」[8]では，AI判断が開発者，医療者，患者にとって理解可能であり，十分な情報公開と意味のある議論の必要性が求められている．

その上で，Transformerに関連するモデルには，Multi-Head Attention機構 の学習が複雑化する点や[9, 10]，Transformerにおける可視化手法が未だ明確化されていない点[11]など，推論の解釈性を高める方法論に議論の余地がある[12]．そこで本研究では，統計的な不足を補うための大規模モデルだけに依存するのではなく，解釈性を高めるための根本的な学習原理の解明に挑戦している点で学術的に重要な意義を持つ．

[1] Waisberg et al., *World J Hepatol*, 2015
[2] Xia et al., *Radiology*, 2023
[3] Cirui et al., *EClinicalMedicine*, 2021
[4] Vaswani et al., *ArXiv*, 2017
[5] Maithra et al., *Arxiv*, 2022
[6] Yutao et al, *Arxiv*, 2023
[7] Weixuan et al., *Arxiv*, 2023
[8] https://www.who.int/
[9] Li et al., *Pattern Recognition*, 2024
[10] Elena et al., *Arxiv*, 2019
[11] Hila et al., *CVPR*, 2021
[12] Kohitij et al., *Nat Mach Intell*, 2022