PiDP：処方箋に対するGCNの対照学習に基づく錠剤検出

PiDP: Pill Detection based on GCN-based Contrastive Learning on Medical Prescription

グエン チュン タイン [1]　グエン フィ レ [2]　川西 康友 [3;1]　駒水 孝裕 [1]　井手 一郎 [1]
Trung Thanh Nguyen　Phi Le Nguyen　Yasutomo Kawanishi　Takahiro Komamizu　Ichiro Ide

名古屋大学 [1]　ハノイ工科大学，ベトナム [2]　理化学研究所 [3]
Nagoya Univeristy　Hanoi University of Science and Technology, Vietnam　RIKEN

1　はじめに

病気の治療過程で，錠剤の取り違えは患者にとって身体的や精神的な影響だけでなく，時には致命的な問題も生じる．そのため，服用される錠剤が処方箋どおりか否か確認することが重要である．先行研究として，我々はPIMA[1] という，錠剤画像と処方箋テキストを照合する手法を提案した．この手法では，錠剤画像と処方箋テキストの組を入力として，処方箋に記載された錠剤名と錠剤画像を一致させるよう学習する．しかし，単一の錠剤画像を対象としているため，複数の錠剤が写った画像には対応できなかった．そこで，本発表ではこの問題に対処した，錠剤の検出手法：PiDP を提案する．

2　提案手法

図 1 に本研究の提案モデルの概要を表す．提案モデルは以下の 3 つのモジュールから構成される．

1. 錠剤検出：Faster R-CNN[2] を画像全体に適用して得られる特徴量マップから物体領域候補を提案する．その後，注目領域（RoI）プーリングにより，特徴量マップ上の候補領域から特徴量を抽出する．
2. 処方箋情報抽出：処方箋からテキスト情報を抽出する．具体的には，テキストから Bidirectional Encoder Representations from Transformers（BERT）モデルによりテキスト埋め込みを生成する．さらに，処方箋内のテキストの位置から処方箋グラフを生成し，Graph Convolutional Network（GCN）を用いて特徴量を表現し，それによって錠剤名を表すテキストを強調する．
3. 錠剤・処方箋の対応付け：上記 1 と 2 で得られた特徴量から，次の 3 つの損失関数を計算する：1）各候補領域が物体か背景かを分類，2）錠剤名に対応する錠剤画像を照合，3）テキストが錠剤名か否かを分類．

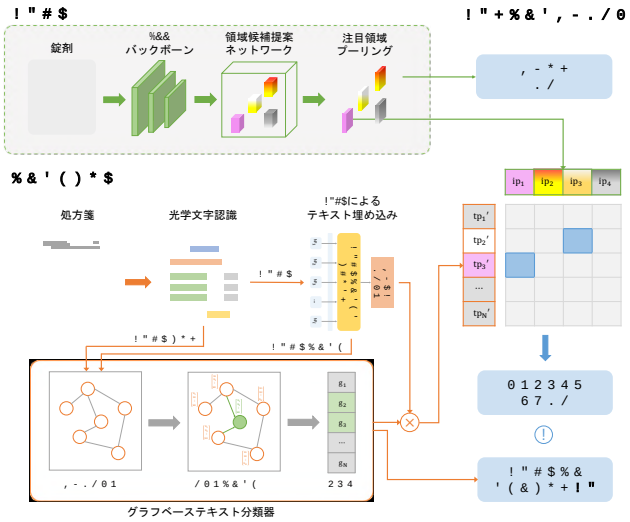

図 1　提案手法（PiDP）の処理手順

表 1　検出結果の平均適合率［%］

Model	mAP
Faster R-CNN	47.59
Faster R-CNN + CLIP	56.60
PiDP（GCN なし）	52.69
PiDP（GCN あり；提案）	65.39

3　実験

提案手法の有効性を確認するため，公開データセット[1] を用いて学習・評価し，錠剤検出の平均適合率（mAP）を評価した．表 1 に提案手法と比較手法の評価結果を示す．PiDP は従来の Faster R-CNN 検出手法と比較して mAP の値が 17.8% 向上した．また，Contrastive Language-Image Pre-training（CLIP）[3] による画像とテキストの対応付けを Faster R-CNN と組み合わせることにより，Faster R-CNN より高い適合率が得られたが，PiDP よりは低かった．また，提案手法において GCN を利用しない場合は，適合率が大きく低下した．

4　むすび

本発表では，PIMA 手法を拡張し，複数の錠剤が写った画像に対応できる新しい検出手法：PiDP を提案した．実験により，提案手法の有効性を確認した．今後は，未知の錠剤クラスを検出する手法を検討する．

謝辞　本研究の一部は Vingroup Innovation Foundation 研究費 VINIF.2021.DA00128 による．

参考文献

[1] T. T. Nguyen et al., “A novel approach for pill-prescription matching with GNN assistance and contrastive learning,” *PRICAI*, 2022.

[2] S. Ren et al., “Faster R-CNN: Towards real-time object detection with region proposal networks,” *ICCV*, 2015.

[3] A. Radford et al., “Learning transferable visual models from natural language supervision,” *MLR*, 2021.

Copyright © 2023 IEICE