

新潟脳神経研究会特別例会の御案内

日時：令和5年5月25日(木) 17:00～18:00

場所：脳研究所 A棟 1階 検討会室

バイオメディカル画像解析における ラベル効率的な機械学習



九州大学大学院システム情報科学研究院・
准教授

備瀬 竜馬 先生

様々な生命現象を捉えた画像を深層学習等の機械学習を用いて解析する際には、自動識別器を学習するために、画像に加えて正解ラベル（教師）を作成する必要がある。しかし、生物分野においては、正解ラベル作成には、専門知識が必須であり、一般物体認識タスクと比べ、教師データ作成コストが格段に高く、大量のデータを保有していたとしても、正解ラベルを大量に作成することが難しい。

本発表では、このような課題を解決するための効率的に付与されたラベルを用いた学習手法（Label Efficient Learning）とそのバイオメディカル画像解析への応用例について紹介する。例えば、Label Efficient Learning の代表例として、少数の教師データに加えてラベルが付与されていないその他のデータ（教師なし）を効果的に活用する半教師学習がある。他にも、直接的な教師データではないが、間接的にそのタスクの答えに関係がある情報（弱教師）を用いた学習を弱教師学習がある。生命科学及び医学分野においては、データ収集時にこのような関連する情報を取得済みであることが多く、完全な教師データ作成に比べてはるかに簡便に取得できることが多い。このような簡便に取得可能な弱い教師を用いて、関連するタスクを解くことを可能とする。これらの学習手法の問題設定の概要を紹介した後、具体例をいくつか紹介する。

弱教師学習の具体例として、癌種割合（弱教師）を用いた病理画像の癌種別のセグメンテーションを紹介する。セグメンテーションにおいて、通常の教師あり学習を適用しようとする、大量の病理画像に対して、癌種ごとの領域の境界を専門家がアノテーションする必要がある。一方、1枚の病理画像に対して、癌種Aは約70%、癌種Bは20%、癌種Cは10%というように癌種のおおよその割合が診断情報として既に得られている。この癌種割合情報は、領域のセグメンテーションという解きたいタスクの答えではないが関連した情報であり、弱教師として考えることができる。我々は、少量の教師データと大量の教師なしデータに付与されている癌種割合情報（弱教師）を活用することで、セグメンテーションの精度改善を可能とする手法を開発した。また、他にも細胞画像解析における細胞検出、セグメンテーション、追跡といった学習タスクを例として、弱教師や教師なしデータを有効に活用して学習する手法について紹介する。

どうぞ奮ってご参加ください。

(担当：脳研究所 細胞病態学分野)

