

学位論文要旨

論文題目：生体試料の TOF-SIMS 分析における教師なし機械学習の応用

理工学研究科 理工学専攻 青柳研究室
学籍番号：D196101 氏名：松田 和大

飛行時間型二次イオン質量分析法 (TOF-SIMS) は、固体試料表面にイオンビームを照射することで最表面かつ微小部の質量分析が可能な分析手法であり、特に最近では生体組織の組成分布を可視化するための質量イメージング手法として活用が進んでいる。TOF-SIMS では原子イオンや分子イオンのほかに、分子が断片化されたフラグメントイオンが多く発生するため、大規模かつ複雑なデータとなることが一般的である。特に生体試料のような組成が複雑な試料では、その傾向が顕著である。そのため、TOF-SIMS データに対して網羅的な分析を行うことは、人力のみでは困難になりつつある。そこで本研究では、近年、発展が著しい機械学習、具体的には人工ニューラルネットワークを用いた特徴抽出法であるオートエンコーダーについて、生体試料の TOF-SIMS データ解析への応用を検討した。

毛髪データに対するオートエンコーダーの適用検討では、学習によって得られる二種類の重みパラメーターの違いや、中間層サイズが抽出される特徴に与える影響について調べた。それによりオートエンコーダーで TOF-SIMS データを解析するための基礎的な指針を得ると共に、生物学的知見と矛盾しない成分分布を、オートエンコーダーが特徴として元データから抽出できる性能を有していることを示した。

次に、特徴抽出性能を向上させるための改良として、中間層をスパース (疎) にするために正則化の概念を導入した、スパースオートエンコーダーの適用検討を行った。皮膚の二次元イメージデータを用いて、正則化項の種類や正則化の程度が、抽出される特徴にどのような影響を与えるのかについて調べた。それにより、適切な条件の下では、スパースオートエンコーダーは単純なオートエンコーダーに比べて特徴抽出性能が顕著に向上することを示した。

さらに、マトリックス効果の影響で濃度に対し非線形な応答を示すデータに対し、オートエンコーダーを実施することで、マトリックス効果の影響を軽減した (濃度依存的な) シグナル強度を持つ特徴が抽出できることを示した。これは非線形近似が可能な人工ニューラルネットワークの使用の有意性を示したデータと考えられる。

本研究は、生体試料の TOF-SIMS データに対し、人工ニューラルネットワークを基にした解析手法の適用を本格的に検討した初の研究である。そのため、本研究で得られた知見は今後の大規模分析データへの機械学習の適用に役立つことが期待される。