

硬質菌類の孔口拡大写真からの自動特徴量抽出

中島淳志

2018.1.13 千葉菌類談話会スライド会要旨

【背景・目的】

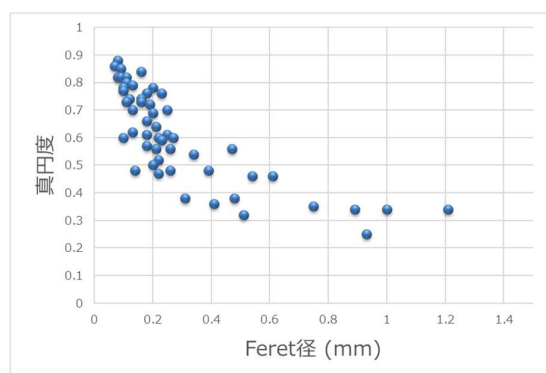
画像に基づく種同定は非侵襲的かつ再現性が高く、適切な画像処理手法を確立すれば高速化および自動化が可能である。硬質菌は標本作製過程での劣化が軟質菌よりも少なく、乾燥標本からの識別形質情報の抽出が比較的容易と考えられる。硬質菌の孔口のサイズおよび形状はグループごとに大きく異なり、本菌群の識別形質として一般的に用いられる。従来、孔口のサイズは主に「1 mmあたりの個数」で表現され、形状は「円形」「多角形」などの定性的記述がなされてきたが、より定量的なデータが同定および分類に有用である可能性がある。前者については、拡大像を用いれば正確な面積や直径の計測が比較的容易であり、後者は真円度 (circularity) 等の指標により定量的表現が可能である。本研究では孔口拡大写真の画像処理による測定値データの自動抽出手法を検討するとともに、得られたデータを特徴量とする機械学習による種判別を試みた。

【方法】

一定の光源・倍率・角度で撮影可能な USB 顕微鏡「Dino-Lite Premier E」を用いて、100 種の硬質菌類標本（主に千葉県立中央博物館収蔵の乾燥標本）の孔口面を「55 倍」の倍率で撮影した。種ごとに最低 25 枚の写真撮影した。ImageJ で自動粒子計測プログラムを作成し、孔口の「面積」「Ferret 径（最大、最小）」「真円度」「重心の座標」を抽出した。各種の 1 枚目の写真の抽出結果を目視で確認し、自動抽出が適用可能かどうか判定した。偽陽性データを減らすため、種ごとに面積と真円度の大きな範囲を設定し、その範囲に収まったデータのみを採用した。また、二次的に「Ferret 径の最大/最小比（形状の指標）」「最近傍 5 孔口重心との平均距離（密度の指標）」を算出し、誤同定を起こしやすいいくつかの組み合わせについては、「最大 Ferret 径」「真円度」と合わせた 4 変数を用い、Python の scikit-learn ライブラリでサポートベクトルマシン (SVM) による機械学習を試みた。

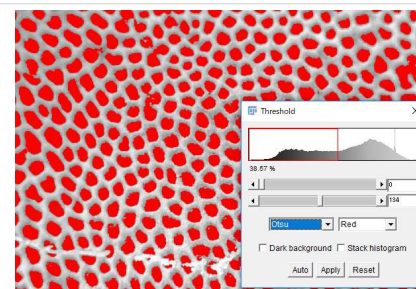
【結果】

本手法の適用は 100 種中 53 種で妥当と判断した。計 33,122,527 個（平均：624,953 個、範囲：2,342 個[アヒラタケ]～4,886,951 個[ウチワタケ]）の計測値データを得た。Ferret 径中央値最小はコブサルノシカケモドキ (0.07 mm)、最大はアヒラタケ (1.21 mm) であり、真円度中央値最小はタマシヨレイタケ (0.25)、最大はオオスルメタケ (0.88) であった。真円度と Ferret 径には有意な負の相関を認めた (pearsonr = -0.78, $p < 0.001$)。SVM の判別精度は「ツウチワタケ」と「～モドキ」で 90.4%、「アシグロタケ」と「キ～」で 97.5%であった一方、「コブサルノシカケ」と「カシ～」(和名が異なるが同種) では 72.4%に留まった。



【考察】

種ごとに適用可否の判定や抽出数値範囲の絞り込みを要したものの、画像処理により複数の計測値を大量かつ迅速に自動取得可能であった。本手法は濃色で境界明瞭の孔口を有する種には適していたが、白色～淡色の孔口や皺状・皺状子実層托はほぼ検出不能であり、コントラスト調整などの前処理や倍率の変更が必要と考えられた。Ferret 径と真円度との負の相関は、孔口が径を増すと放射方向に伸長する一般的傾向に由来すると考えられたが（アミスギタケやアミヒラタケに顕著）、フルイタケのように大型でも真円に近い孔口を有する種もあり、分類群や地理的分布（北方系・南方系）、子実体発達様式等との関連に検討の余地があるだろう。



【結論】

乾燥標本の孔口拡大写真から画像処理により大量の計測値を自動取得することができた。これは従来硬質菌類の分類に用いられていた定性的・半定量的形質よりも客観性・再現性に優れる定量データであり、頻度分布も容易に把握可能であった。また、取得したデータを特徴量として高精度の種判別器を作成可能であった。一方、約半数の種については汎用プログラムでは孔口を検出不能であり、手法の改良が今後の課題である。