

創造的な生物・情報知識融合型の研究開発

平成19年度採択新規研究開発課題

○課題名：酵素反応分類に基づく酵素反応予測システムの開発

代表研究者：長野希美（産業技術総合研究所生命情報工学研究センター）

S-1

酵素活性部位探索に対する機械学習の適用

長野希美¹、加藤毅^{1,2}

(1 産業技術総合研究所生命情報工学研究センター、2 お茶の水女子大学生命情報学教育研究センター)

酵素反応を比較する際には、酵素蛋白質の触媒部位を比較することが不可欠で、蛋白質の局所構造を比較する方法（テンプレート解析法等）が使われる。Wallaceらのテンプレート解析法は、原子レベルの局所構造テンプレートを作成し、大規模な構造データベースに検索をかける手法である。しかし、この方法は次のような問題点がある：(a) テンプレート構築(どの原子はテンプレートに含め、どの原子を外すかという選別作業)は、アノテータによる職人的手作業で行うことを前提とする。このため、予測性能はアノテータの能力に大きく依存する；(b) テンプレートに含まれるそれぞれの原子には酵素反応に重要なものと重要でないものがあるにも関わらず、機能未知の蛋白質（クエリ）との重みなし平均原子間距離を距離尺度として評価している。以上の欠点により既存のテンプレート解析法は膨大なFalse positive が検出されてしまう。そこで、原子毎に重み付けして平均原子間距離を測ることにより、False positiveが減少することが期待される。本研究開発課題では、この重み係数を自動的に求める機械学習システムを開発している。

○課題名：進化型計算と自己組織化による適応的画像分類法の開発

代表研究者：馳澤盛一郎（東京大学大学院新領域創成科学研究科）

S-2

クラスタリングと組み合わせ最適化を用いた生体画像の適応的自動分類法の開発

馳澤盛一郎¹、朽名夏磨¹、桧垣匠¹、松永幸大²、藤井博史³、山口雅之³

(1 東京大学大学院新領域創成科学研究科、2 大阪大学大学院工学研究科、3 国立がんセンター東病院臨床開発センター機能診断開発部)

バイオイメージングは幅広い生命現象をカバーする基本的なアプローチであり、各分野で膨大かつ多彩な画像データが蓄積しつつあるが、その一方、解析方法の整備は遅れている。私たちは多岐にわたる生命科学分野において画像データ解析の効率化を図るために、画像の性質や解析の目的に特化しない、画像の分類や検索のためのフレームワークを開発している。本研究では(1) 画像の評価尺度となる適切な特徴量の探索と(2) 自動分類システム等の構築に必要な多量のアノテーション情報の入力支援とをあわせて行なうために、画像群のクラスタリング結果を組み合わせ最適化の対象とした部分教師付き学習法を考案した。また、選択された特徴量と付加されたアノテーション情報をもとに分類基準の決定を行なう学習アルゴリズムに、ハイパーパラメタが少なく計算時間の短い集団学習法を採用することで設定工程を迅速化した。現在、動植物細胞の細胞周期の分類、植物組織からの気孔の検索と開閉評価、MRI画像およびSPECT画像からのガンの種別や重量の推定等について本システムによる自動化を遂げており、今後、応用範囲のさらなる拡大が期待される。