

## 少ない遺伝子発現データを活用するための畳み込みニューラルネットワークと転移学習による毒性予測システム

高橋弘樹<sup>†1</sup> 山根順子<sup>†1</sup> 藤渕航<sup>†1</sup>

**概要**：機械学習分野の発達により、新薬開発においても機械学習による分類・回帰が薬効・毒性予測に利用されるようになってきた。我々の研究室でも、薬剤暴露させたヒト幹細胞の遺伝子発現データを用いて、薬剤の毒性予測を行うための機械学習システムの開発を目指している。しかし、機械学習は大量のデータからパターン認識を行うのが本来の使い方であるが、一方で遺伝子発現データを大量に収集するのは容易ではない。したがって、少ないデータを用いて学習できるような機械学習システムの開発が今後必要になると考えられる。本研究では、畳み込みニューラルネットワーク（CNN）の学習済みモデルを用いる。このモデルは ImageNet の大規模画像データセットを事前に学習させているため、既学習の重みを共有することで少データの欠点を解消する。我々は、この転移学習モデルと自己組織化マップ法で二次元画像化された遺伝子発現データを用い、ヒト肝細胞に対する化合物毒性の有無を予測する。本研究を通じて、転移学習モデルのマルチラベル分類の予測正答率、ハイパーパラメータ決定のための最適化戦略、遺伝子選択法の導入による予測精度の変化等を議論する。

**キーワード**：毒性予測、転移学習、畳み込みニューラルネットワーク、遺伝子発現データ

---

<sup>†1</sup> 京都大学 iPS 細胞研究所  
Center for iPS Cell Research and Application, Kyoto University