

主 論 文 要 旨

| 報告番号 | 甲 ㊦ 第 号 | 氏 名 | 鎌 谷 高 志 |
|--|---------|-----|---------|
| 主 論 文 題 名 Construction of a system using a deep learning algorithm to count cell numbers in nanoliter wells for viable single-cell experiments (生存シングルセル実験におけるナノリットルウェル中の細胞数測定を行うディープ・ラーニングアルゴリズムを用いたシステムの構築) | | | |
| (内容の要旨) 共著者の白崎らは細胞集団の不均一性を解析する手段としてナノリットルウェル内でリアルタイムに生存細胞の動態やサイトカイン分泌を確認するSingle cell解析法を開発した。私が研究している喘息における重要な細胞である2型自然リンパ球は、ヒトにおいては末梢血から採取できる量が少ない希少細胞であり、同実験系を利用し機能解析を行っていた。しかしナノリットルウェル内の生存細胞の数を確認するために、経時的に得られた大量の顕微鏡画像を用いて目視下で数える必要があった。そのための消費時間は著しい上に、時間経過に伴う疲労など技術者の負担が大きく、速度、精度も下がってしまうことが難点であった。 そこで顕微鏡画像として得られた大量のデジタルデータをオートメーション化することを考え、ニューラルネットワークの1つであるディープ・ラーニングを用いることを考えた。ディープ・ラーニングの1つである畳み込みニューラルネットワーク (Convolutional Neural Network: CNN) を用い、18層のネットワークを構築しナノリットルウェル内の生存細胞数を解答データとして学習させ、同ネットワークを使用することで顕微鏡画像からナノリットルウェル内の適切な生存細胞数を短時間かつ高い精度で出力することを検証した。具体的には顕微鏡画像は1ウェルあたり511x511ピクセルのJPEG2000データで構成される。そのデータを計103019サンプル集め、まず同サンプルの細胞数を4つのカテゴリ(0, 1, 2, および3以上の細胞数)に分類した。二人の技術者が目視下によって解答データを作成し、CNNの出力結果と比較した。その結果、同サンプルによる訓練後、訓練したデータには使用していない独立したテストデータで検証したところ、解答データと比較してCNNは98.3%の精度でサンプルを4つのカテゴリに分類することができ、消費時間も1/80 (3960サンプルの解答作成に30分から22.5秒) に減少した。またウェルの形状が異なる顕微鏡画像を用いて、同じ構成のCNNを用いて学習させたところ同様の精度となり、ウェルの形状によらない汎用性が示された。 次に、CNNが不正解であった画像に注目したところ、間違えた画像の多くは目視で確認しても解答を出すことが難しい不明瞭な画像であることがわかった。その特徴を利用してさらに精度を上昇させるために、CNNの出力結果である4つのカテゴリの事後確率を利用し、事後確率の中で最も高い値が、特定のカットオフを下回ったナノリットルウェルを目視で再確認するシステムを構築し、精度が99%以上に向上した。 今回用いたCNNシステムは大量のデジタル画像データから高い精度かつ短時間で細胞数を分類することを可能にした。今後大量の画像データを用いた単一生存細胞の解析実験などへ汎用できる可能性が考えられた。 | | | |