

ニューラルネットワークの特徴マップの解析と活用
(Analysis and Applications of Feature Map of Deep Neural Networks)

氏名 井手 秀徳

脳の視覚野を模倣して作られた畳み込みニューラルネットワーク (CNN) は、画像分類の分野などで注目されている手法である。階層構造を持つ CNN による情報処理の本質は、画像などの入力データに内在する重要な情報 (特徴) を、データからの学習によって獲得することである。本質に近い特徴を獲得するためによく試みられるアプローチは、CNN の層を増やす (深くする)、変数の数を増やすなどである。しかし、それらは、学習に用いられるデータ (学習サンプル) を CNN が過度に学習してしまうことで逆に性能が悪化する、値の発散により学習が失敗するといった多くのリスクも伴う。他にも、CNN の実行に要する莫大な計算コストも問題になる。本論文では、CNN が獲得すべき特徴について議論し、その獲得を促進することで学習の安定化や分類精度向上を試みるだけでなく、獲得した特徴を活用することで、CNN の計算コスト削減も目指す。

第 2 章では、CNN の中間層が獲得すべき特徴について議論し、その獲得を促進し、学習を安定させる正則化手法を提案する。CNN は、複数の層からなる中間層と、中間層で得られた特徴を用いて分類問題を解く出力層 (分類器) で構成されている。CNN の各層、特に、入力に近い層は脳の視覚野に近い機能を持つため、学習時に視覚野の形成を促進できれば、より良い特徴を獲得できると示唆されている。スパースコーディングを用いると V1 視覚野の受容野を自己組織化できることから、本章では、中間層で獲得する特徴に対してスパース制約を課すスパース正則化を提案した。スパース正則化を様々な CNN に適用することで、学習が安定し、分類精度も向上することを実験的に確認した。また、学習サンプルが極端に偏り、ネットワークの学習がうまくいかない条件下でも、スパース正則化ならばその悪影響が少ないことも確認した。これらの結果は、CNN の分類精度を改善する新たな正則化手法の有効性を示しただけでなく、未だ多くの部分が未知である CNN を理解するのに役立つと考えられる。

第 3 章では、CNN の分類器にとって好ましい特徴について議論し、その獲得を促進する正則化手法を提案する。K クラスの分類問題を解く分類器は、K 個のニューロンで構成され、ニューロンごとに分類すべきクラス (ターゲットクラス) が割り当てられる。サンプルが CNN に入力されると、ソフトマックス関数を介して、そのサンプルがターゲットクラスに属する事後確率をニューロンごとに出力し、その確率が最も高いニューロン (クラス) を分類結果とすることで、分類器は K クラス分類問題を解いている。分類器全体では K クラス分類問題を解くのにに対し、各ニューロンはターゲットクラスかそれ以外 (非ターゲットクラス) かの 2 クラス分類問題を解く。このとき、各ニューロンにとって好ましい特徴は、ターゲットクラス、もしくは非ターゲットクラスの情報を含んだ 2 つのガウス分布 (分散は等しく、平均は異なる) である。本章では、そのような特徴の獲得を促進するために、各ニューロンに対する正則化として判別基準を適用する判別正則化を提案した。判別正則化により、そのような特徴が獲得されやすくなり、分類精度も向上することを確認した。また、分類器の活性化関数として一般的なソフトマックス関数の代わりに、シグモイド関数を活性化関数として用いることで、より高い分類精度を示す場合があることも確認した。これらの結果は、ニューラルネットワークの出力層 (分類器) を改善する重要性を示しただけではなく、未だ本質的な部分の多くが未知である CNN を理解するのに役立つと考えられる。

第 4 章では、CNN が獲得した特徴マップを活用することで、その計算コストを削減する新たな Pruning 手法を提案する。様々な Pruning 手法が提案されているが、その本質は、CNN が獲得した特徴の、分類問題に対する貢献度合い (重要度) を評価し、重要度が低い特徴、ならびに、関連する変数の剪定 (Pruning) である。これは、Pruning で活用すべき情報は獲得された特徴であることを示唆している。各特徴の分類に対する重要度

(Pruning Score) は、正解が割り当てられた学習サンプルを用いて経験的分類損失を計算し、それを特徴マップに関してテイラー展開することで近似できる。しかし、経験的分類損失は一部の極端なサンプルに大きく影響されるため、Pruning Score が信頼できない可能性がある。本章では、テイラー級数を代数的に変換して導出した関数を用いる Pruning 手法を提案した。これにより、一部の極端なサンプルによる悪影響を減らしつつ、特徴の有無が分類損失に与える影響を直接評価 (予測) することができた。また、その延長として、Pruning Score を正規化する新たな正規化手法も提案した。他にも、Pruning を補助変数の連続最適化問題に置き換え、それを安定して解く Pruning 手法も提案した。CNN を用いた画像分類の実験では、どちらの提案手法も、他の類似した Pruning 手法と比較して良好な性能を示した。これらのことから、CNN が獲得した特徴マップは、CNN の計算コスト削減に活用できることを示せた。これらの結果は、単に特徴マップの情報を活用することの重要性を示しただけでなく、ブラックボックスである CNN 内部の挙動を解析するのに役立つと考えられる。