

深層学習における画像の拡張について

1. はじめに

深層学習モデルは数多くのパラメータを持つため、適切な学習には多種大量のデータが必要です。しかし、実際の状況では、十分なデータを収集することが難しい場合もあります。データが不足すると、学習したデータだけに正解する過学習が生じやすくなります。

画像（データ）を拡張することで、既存の画像に変形をはじめとした様々な手法を適用して学習に利用するデータセットの多様性が高まり、過学習を回避し、モデルの汎化性能が向上します。

2. 主な画像（データ）の拡張手法

主な手法として、回転、拡大縮小、平行移動、鏡像反転、明るさやコントラストの調整、ノイズの追加、せん断などがあげられます。

どの画像拡張手法を適用するかについては、データや目的から判断します。例えば、文字の画像であれば、回転させても意味が変わりませんが、鏡像反転は避けるべきです。また、服や布地であれば撮影時に歪んだり、明るさが変動したりすることがあるので対応した拡張が適用できます。

公開データセット Fashion-MNIST¹⁾に対して、画像拡張を適用した例を図1に示します。



図1 拡張前の画像と拡張した画像

さらに高度な画像拡張として Mixup という手法があります。この手法では、異なる2つの画像をブレンドして新しい画像を生成し、例えばシャツの確率30%、バッグの確率70%として学習させます。また、一つの画像からランダムな領域（パッチ）を切り取り、別の画像にそのパッチを貼り付ける CutMix という手法があり

ます。これらの画像拡張手法を用いることで、曖昧さを含む中間的なデータや現実世界の複雑なバリエーションに対しても適切な予測をするように、モデルを学習させる効果が期待できます。

3. 画像拡張の効果

画像拡張の有無による学習の様子（図2）を紹介します。モデルは畳み込みニューラルネットワークである ResNet50V2、データセットは繊維を顕微鏡で撮影した画像です。画像サイズは224×224ピクセル、8種類の繊維各300枚で分類問題として学習しました。画像拡張は、上下左右反転、回転および平行移動をランダムに適用しました。

グラフ上段は学習用画像での正解率で、データ拡張なしは過学習により高い正解率を示しています。一方下段は学習に利用していない検証用画像での正解率で、学習が進むにつれ拡張した方が正解率が高くなっています。

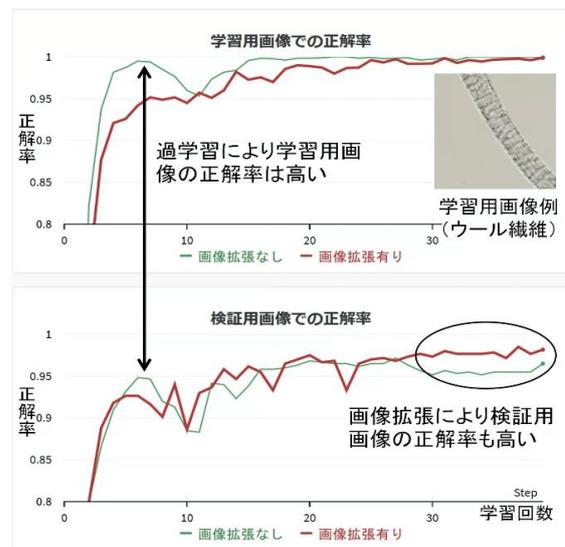


図2 画像拡張の有無による学習

4. おわりに

当センターでは単繊維の顕微鏡画像の鑑別について、AIを利用した取り組みを進めています。

AIを用いた画像分類などをお考えの方はお気軽にお問い合わせください。

参考文献

1) <https://github.com/zalandoresearch>