

3D CNN による学習データセット ModelNet10 の 分類実験の評価と改善方法の検討

永田研究室 F120011 金沖 巧望

1. 目的

近年, 3D オブジェクトの認識技術が急速に重要視されており, 自動運転車の環境認識やロボットによる物体操作など, 様々な分野での応用が期待されている. また, ディープラーニングの進展により, 3D CNN (3D Convolutional Neural Networks) が 3D オブジェクト認識でより高い性能を発揮することも期待されている. 特に, 3D CNN の応用例としてボクセルを利用したクラス分類が挙げられる. ボクセルは 3D ピクセルでオブジェクトを表現でき, 例えば, 様々な家具や車両, 動物など, 異なるカテゴリやクラスのオブジェクトを機械的に正確に分類する手法として有効である. しかし, 実際の分類問題に対処するためには, 良質で多様なデータセットが不可欠となる. 今回の評価に使用する ModelNet10 は 10 種類の異なるカテゴリの 3D モデルで構成され, 3D CNN モデルの汎化性を評価するのに適している. 本研究では, 今後, このような 3D CNN の技術をものづくりの現場に適応するにあたり, ModelNet10 を使用したボクセルベースのクラス分類実験を通じて, 3D オブジェクト認識性能と 3D CNN モデルの汎化能力に関する基礎的な知見を得ることを目的とする.

2. 研究内容

本研究では, 分類対象として ModelNet10 の 3 次元モデルを使用した. ModelNet10 の学習データセットは bathtub, bed, chair, desk, dresser, monitor, night_stand, sofa, table, toilet の計 10 カテゴリの家具関連のオブジェクトで構成されており, 訓練データ 3,991 と test データ 908 の計 4,899 個のデータが存在する. 今回ダウンロードした ModelNet10 のデータフォーマットは OFF 形式と呼ばれる. ModelNet10 データセットの分類実験に 3DCNN を適用するため, OFF 形式から VXL 形式へのデータ変換を行った. 具体的には, 変換ツール “IMAGEtoSTL” を使用して OFF 形式から PLY 形式への変換を行い, その後, 開発した MATLAB ソフト上で PLY 形式から VXL 形式 (32×32×32 ボクセル) への変換を行えるようにした. 2つの 3次元畳み込み層と 2つの全結合層を備えた, 全 11 層の 3D CNN (VoxNet) を設計し, 変換した VXL 形式データ (計 4,899 個) を使用して ModelNet10 の訓練を実施した. その後, test データに対する分類の汎化性能を評価し, さらに改善方法の検討を行った.

3. 結果

ModelNet10 による VoxNet の学習条件は, 次の 2 通りで設定した. ①エポック数 5000, バッチサイズ 64, 学習率 0.0001, ②エポック数 200, バッチサイズ 64, 学習率 0.01. 学習条件①は最初に設定した初期値であり, その分類精度は 77.75%であった. その後, 試行錯誤により条件を変えながら学習条件②では分類精度を 81.17%まで高めることができた. その結果の混同行列を図 1 に示す. 今回の 3D CNN は非常にシンプルな構造であったことと, 学習データセットとして使用した ModelNet10 の各カテゴリのデータ数が非常にインバランスな配分であったため, 全カテゴリに対する汎化性を均等に高めることは容易ではなかった. 今後の展望として, 畳み込み層やプーリング層の追加を考慮し, 3D CNN の深化に向けた設計方法を検討する. また, OFF 形式などの CAD モデルの回転による新たなデータ拡張法の開発を検討する. 更に, OFF 形式から VXL 形式への一括変換を可能にするファイル変換システムを開発することで, より大きなデータセットである ModelNet40 に含まれる OFF ファイルの一括変換が可能になり, それらの学習と分類実験を行う.

	bathtub	bed	chair	desk	dresser	monitor	night stand	sofa	table	toilet
bathtub	31	1	3		3	2	4	1	4	1
bed		93	3				1		1	2
chair		2	94	1		1		2		
desk	2	2		48	6	2	2	11	13	
dresser		2		4	68		8	3	1	
monitor	1	1		1	4	89	1		2	1
night stand			1	1	19		47	4	14	
sofa		2	1					96	1	
table			2		12		2	3	2	78
toilet	1	1	2						1	95

Fig.1 Classification results of ModelNet10 using 3D CNN named VoxNet.