

畳み込みニューラルネットワークによる 3 種類の道路の識別

2015SC067 村瀬康介

指導教員：陳幹

1 はじめに

現代では、多くの画像が処理されそれを認識したり画像から情報を抽出したりすることで多くの情報を自動で理解する必要が増えてきています。

同時にニューラルネットワークの画像認識の有効性が注目され、多く研究されている。

本研究は無人自動車などの道路を利用するデバイスに道路の形状を正しく認識することによってデバイスの正しい運用させ事故のない世界に近づける可能性があります。

は畳み込みの処理の後に行う。

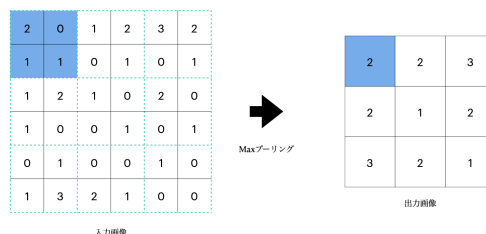


図 3：プーリング層

プーリング層は、図 3 のように画像を分割し各領域の中にある代表的な値を取り出して再び並べることによって新しい画像を作成。このような処理のことをプーリングと呼ぶ。わかりやすく言えば、プーリングは画像をぼかす処理。これにより、画像の位置から情報の正確性が低下して、位置の変化によるエラーが低下。

プーリングを行うことで画像のサイズは元の画像よりも小さい。

2 畳み込みニューラルネットワーク (Convolutional neural network)

CNN とは、人間の「視覚」の仕組みをモデルとしており、人の視覚のように画像認識を得意である。

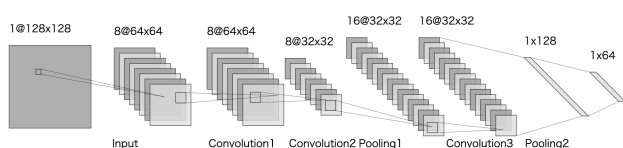


図 1：畳み込みニューラルネットワーク

図 1 は CNN の例になりますが、出力層の各ニューロンが各画像に対応しその画像が何を表しているか確率で示します。例えば、犬の画像を学習済みの CNN に入力すると、75 %で犬、15 %で猫、6 %でウサギ、4 %でネズミ、のようにその画像が最も高い確率を出す。

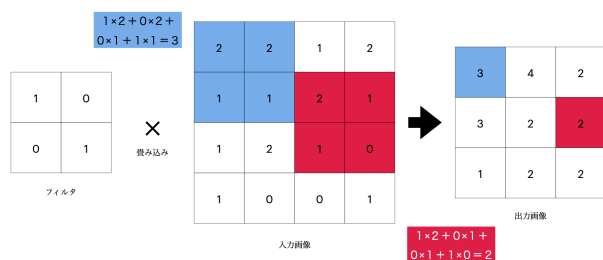


図 2：畳み込み層

畳み込み層は、図 2 のように入力画像にフィルターをかけることによって輪郭を強調する。

畳み込みを行うことで画像のサイズは元の画像よりも小さい。

プーリング層は例外がある場合もありますが、基本的に

3 使用する環境とデータセット

使用する環境は google colaboratory を使用する。ノート PC の速度では実現できない GPU が搭載されているため、学習スピードを速く進める点が利点である。

学習のためのデータセットには、カリフォルニア大学のバークレー校が公開した「Berkeley DeepDrive 100K」と呼ばれるデータセットを使用し合計 10152 枚のデータセット A を作成し、トレーニングデータを交差点、直線、T 字路の 3 つに分けてそれぞれの画像にラベル付与させ使用する。



図 3：データセット (左：交差点、真ん中：直線、右：T 字路)

4 モデル

畳み込み層の 2 層と Max プーリング層の 1 層を 2 回繰り返す、最後に、全結合層が最後に 2 層の合計 8 層の CNN を使用する。

活性化関数には、ソフトマックス関数を使用し、誤差関数

には、多クラス交差エントロピーを使用する。

5 シミュレーション

訓練データを1回使い切って学習することをエポック (Epoch) といい、今回は20回エポックを行う。図5.1は横軸をEpoch数、縦軸を精度(訓練用データやテスト用データを入力しCNNの出力結果とそのデータのラベルの正答率)とし、横軸は右に行くほど学習した回数が多くなっており、縦軸は0~1の間で推移し1に近いほど精度の高いCNNである。

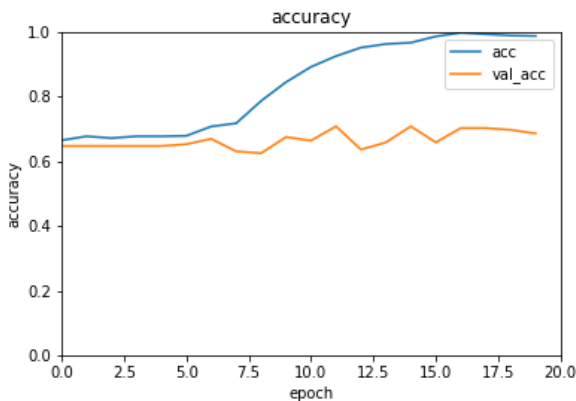


図 5.1 : CNN での画像認識

図 5.1 は訓練用のデータでは精度が向上しているが、テスト用のデータでは精度が向上しているようには見られない。なので、過学習が起こっているように見える。

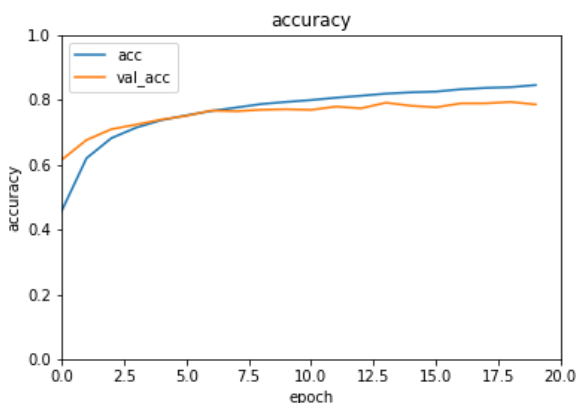


図 5.2 : Dropout を使った画像認識

図 5.2 は Dropout を使い、図 5.1 と比べると過学習は改善され、テスト用データの方で学習によって精度の向上が見られるが、今度は逆に訓練用のデータの方で精度が落ちてしまった。

6 データセットの影響

データセットの偏りがあるので、トレーニングデータとテストデータを混ぜて偏りの影響がないかを確認する。

トレーニングデータは BDD のデータ 9850 枚 + 自分で撮った画像を半分にした 150 枚の合計 10000 枚のデータ

を使い、テストデータは残りの自分で撮った画像 150 枚のデータを使用します。

モデルは 4 章で用いたモデルを使用。

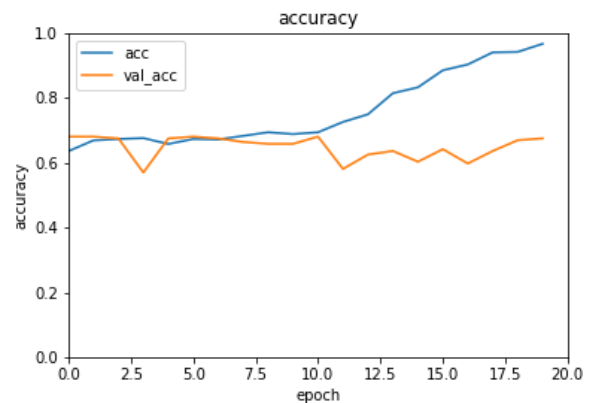


図 7.1 : BDD + 自分で撮った画像 (合計 10000 枚) の画像認識

データの種類は多い方が精度は向上していくが、データ量が少なすぎると学習が行われなくなる。

7 日本の道路の特徴

5 章で行った CNN のモデルから自分で撮ってきた画像 (テスト用画像) を不正解だった 10 枚取り出し、海外のデータでもテストを行い不正解になった 10 枚の画像を取り出し、海外の道路の不正解と日本の道路の不正解を見比べ日本の道路にしかない特徴を見つける。

日本の道路では突出して細い道で撮った画像に不正解が多く見られた。

8 結論

本研究では、画像認識の精度向上を目的とした手法を提案し、1つは Dropout を用いて精度の向上を達成し、2つ目にデータ数と画像の種類の比率を変化させ精度の向上を達成し、従来のモデルよりも高い精度を達成した。精度を上げるためには、データセットの種類を増やすこと、過学習をできる限り防ぐことが必要であるとわかった。

9 参考文献

参考文献

- [1] チームカルポ: 「物体・画像認識時系列データ処理入門」. 秀和システム.
- [2] 我妻幸長: 「人工知能技術の教科書」. 翔泳社.
- [3] 吉井健人, 「畳み込みニューラルネットワークによる曲がり角の画像の識別」, 南山大学院理工学研究科 2019 年度修士論文.
- [4] Nitish Srivastava, Geoffrey Hinton; "Dropout: A Simple Way to Prevent Neural Networks from Overfitting", Journal of Machine Learning Research 15 (2014) pp.1929-1958