

## 7.3 モデルを最大限に活用 するために

## 7.3.1 高度なアーキテクチャパターン

- バッチ正規化

### 正規化

→機械学習モデルが学習する様々なサンプルの類似度を高める事で、モデルが新しいデータを学習したときにうまく汎化できるようにすることを目的とした手法

データからデータの平均値を引き、標準偏差で割ると中心0、標準偏差が1になる

Normalized\_data =

$(data - np.mean(data, axis=\dots)) / np.std(data, axis=\dots)$

## 7.3.1 高度なアーキテクチャパターン

- バッチ正則化

→層の一種(kerasではBatchNormalization)

→訓練中に平均と分散が変化するとしても、データを適応的に正則化できる

→効果としては(残差接続と同じように)勾配の伝播を手助けする

一般的にはBatchNormalization層は畳み込み層か全結合層の後に使用される

#畳み込み層の後

```
conv_model.add(layers.Conv2D(32,3,activation='relu'))
```

```
conv_model.add(layers.BatchNormalization())
```

#全結合の後

```
dense_model.add(layers.Dense(32,activation='relu'))
```

```
dense_model.add(layers.BatchNormalization())
```

## 7.3.1 高度なアーキテクチャパターン

- dw畳み込み

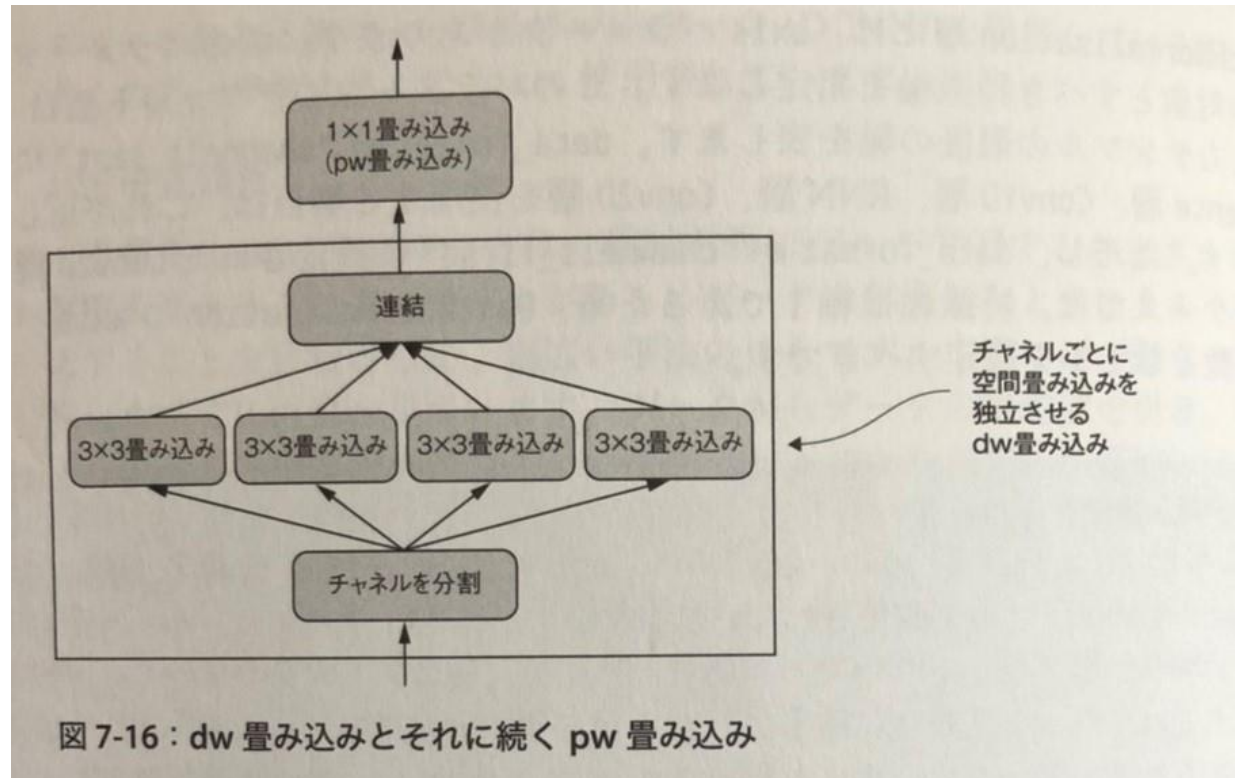
→入力の各チャンネルで空間畳み込み演算を別々に実行した後、pw畳み込み(1\*1畳み込み)演算を通じて出力チャンネルを連結

→入力の空間的な位置同士は高い相関関係にあるが、異なるチャンネル同士はほぼ独立していると想定される場合

## 7.3.1 高度なアーキテクチャパターン

- dw畳み込み

→限られた量のデータで小さなモデルを1から訓練する場合とくに重要



## 7.3.2 ハイパラメータの最適化

- ディープラーニングモデルを構築するときどのように作ればよい？  
→層の数、種類

アーキテクチャレベルのパラメータはハイパラメータと呼ばれる

ハイパラメータを最適化

1. ハイパラメータの集まりを選択
2. 対応するモデルを構築
3. モデルに訓練データを学習させ、検証データで性能を測定
4. 次に試すハイパラメータの集まりを選択する
5. 2~3を繰り返す
6. 最後にテストデータでモデルの性能を測定する

## 7.3.2 ハイパラメータの最適化

プロセスのポイント

- 検証データでの性能履歴をもとに、次に評価するハイパラメータの集まりを選択する。
- ベイズ最適化、遺伝的アルゴリズムなどの手法

ハイパラメータの更新

- フィードバックの計算がかなり高つく可能性がある
- ハイパラメータ空間は離散的な決定で構成されており、連続的でも微分可能でもない
  - 勾配降下法を実行できない

解決するために…ランダムサーチ, Hyperopt

## 7.3.3 モデルのアンサンブル

- 結果を得るために
  - モデルのアンサンブル(model ensembling)
  - よりよい予測値を生成するために、様々なモデルの予測値をプーリングする

### アンサンブルの成功

- 一連の分類機の多様性
- すべてのモデルが同じように偏っていると、アンサンブルも偏ったものになる。

## 7.3.4 まとめ

- 高性能なディープ畳み込みニューラルネットワーク(DCNN)
  - 残差接続、パッチ正則化、dw畳み込みを使用する
  - 将来的には、表現力の高いdw畳み込みが畳み込みに完全にとって代わる可能性がある。
- ディープニューラルネットワークを構築する
  - ハイパラメータやアーキテクチャに関する選択
  - モデルが良いものとなるようにするには、ハイパラメータ空間を体系的に探索
- 成果を上げる
  - モデルの大きなアンサンブルを作成
  - 似通ったモデルからアンサンブルを作成しても無意味

# 7章のまとめ、理解したこと

- 層のグラフ、層の再利用、pythonの関数としてのモデルの使用に基づき、モデルを構築する方法
- Kerasのコールバックを使って訓練中のモデルを監視し、モデルの状態に基づいてアクションを実行できること
- TensorBoardを利用することで、成果指標、活性化のヒストグラム、さらには埋め込み空間まで可視化できること
- バッチ正則化、dw畳み込み、残差接続
- ハイパラメータの最適化とモデルのアンサンブルを使用すべきなのはなぜか