

研究室ローテーション 第6回

担当：和田山 正・中井彩乃

本講義の内容

- **畳み込みニューラルネットワーク**

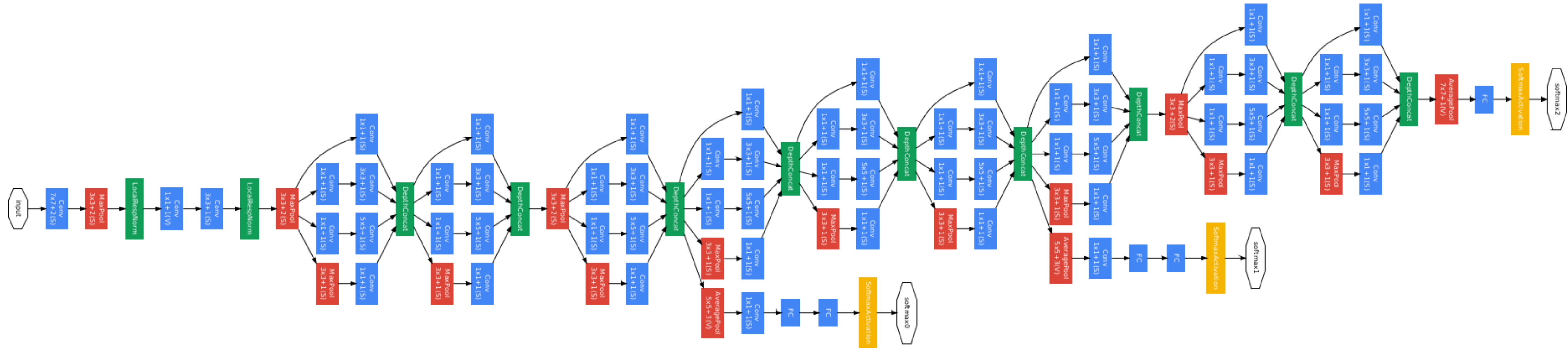
ImageNetデータ例とTOP 5 出力



cited from: ``ImageNet Classification with Deep Convolutional Neural Networks'', Alex Krizhevsky et al.

<https://www.nvidia.cn/content/tesla/pdf/machine-learning/imagenet-classification-with-deep-convolutional-nn.pdf>

GoogleNet



- Googleが開発した深層ニューラルネットワークの一種
- 畳み込みニューラルネットワーク
- ILSVRC2014 トップ

MNIST数字認識

MNISTデータセット



- **0-9の手書き数字**のデータセット
- 28 x 28ピクセル (深さ8bit, モノクロ)
- **訓練データ6万枚/テストデータ 1万枚**
- パターン認識アルゴリズムのテストにしばしば利用される (パターン認識の”Hello World” と呼ばれている)

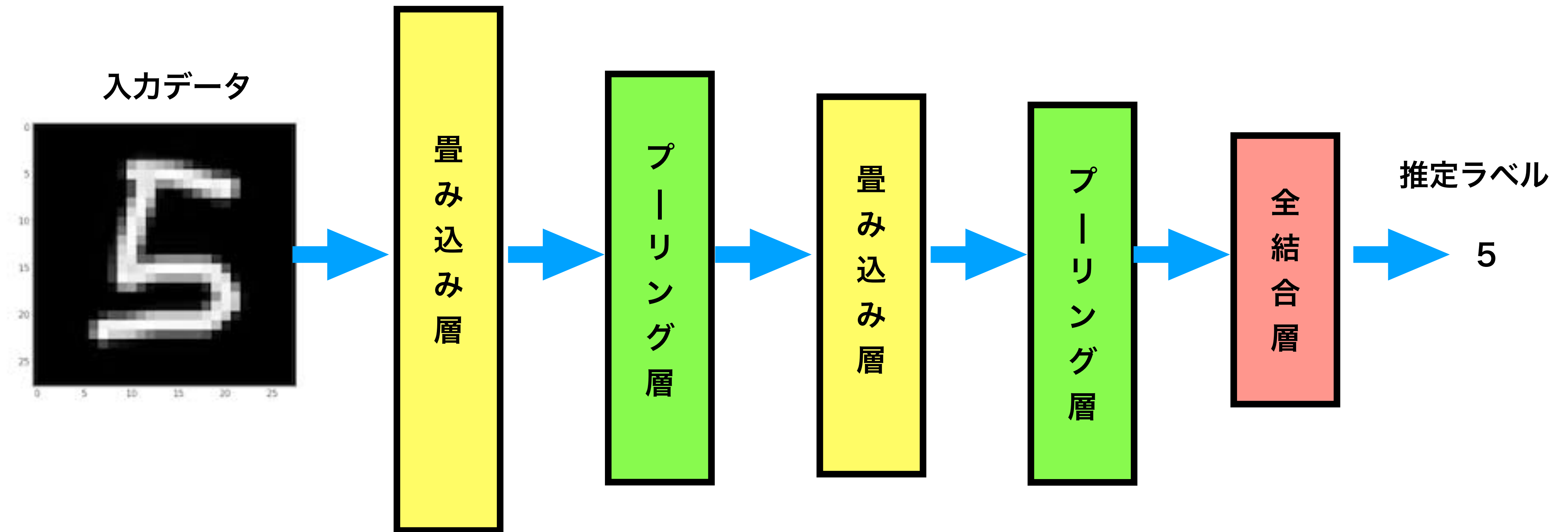
畳み込みニューラル ネットワーク

VGG-16



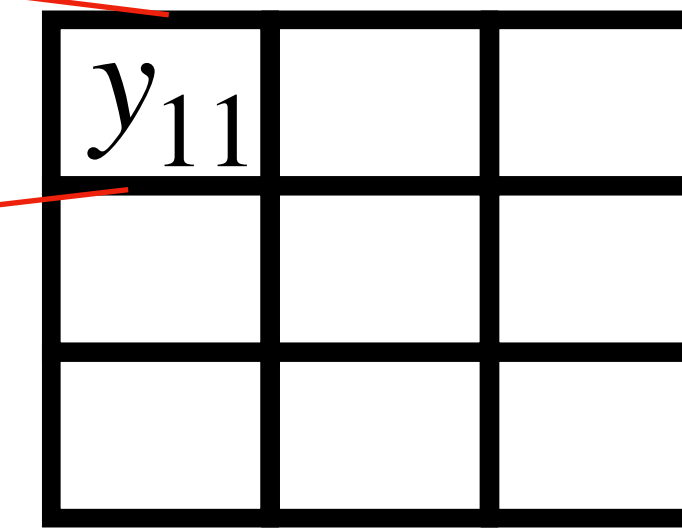
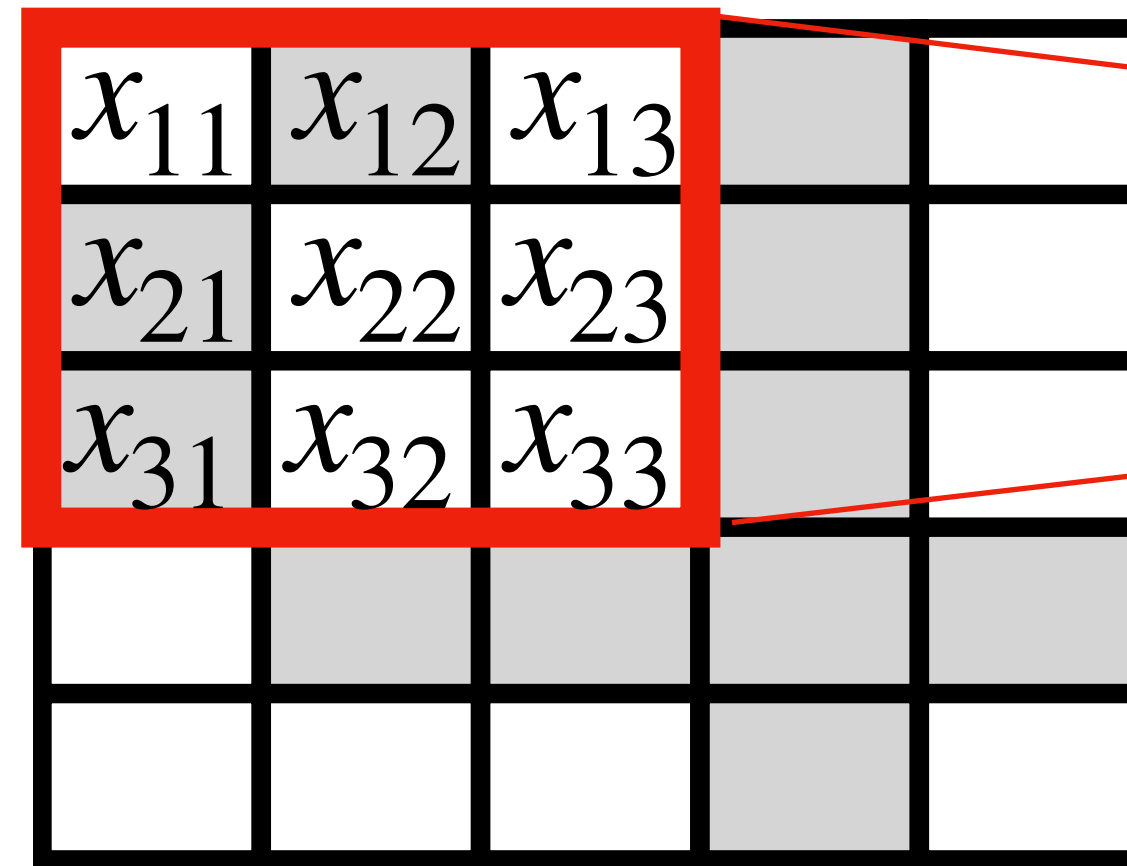
畳み込みニューラルネットワーク

- 2次元畳み込み演算に基づく層 = 畳み込み層(学習パラメータを含む)
- 2次元プーリング演算に基づく層 = プーリング層
- 全結合層

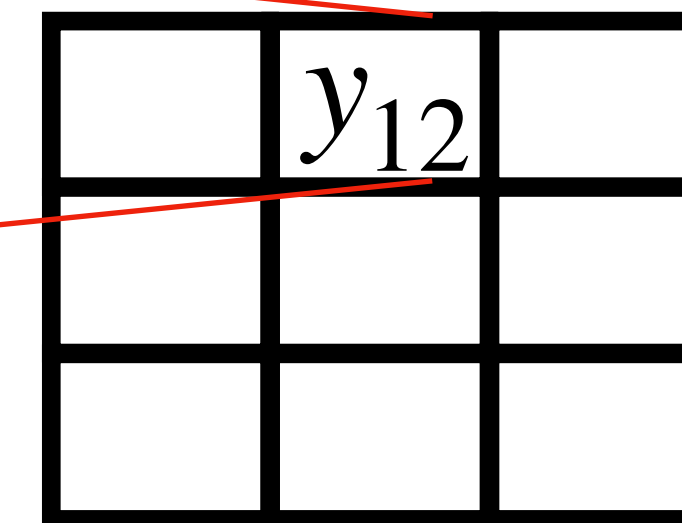
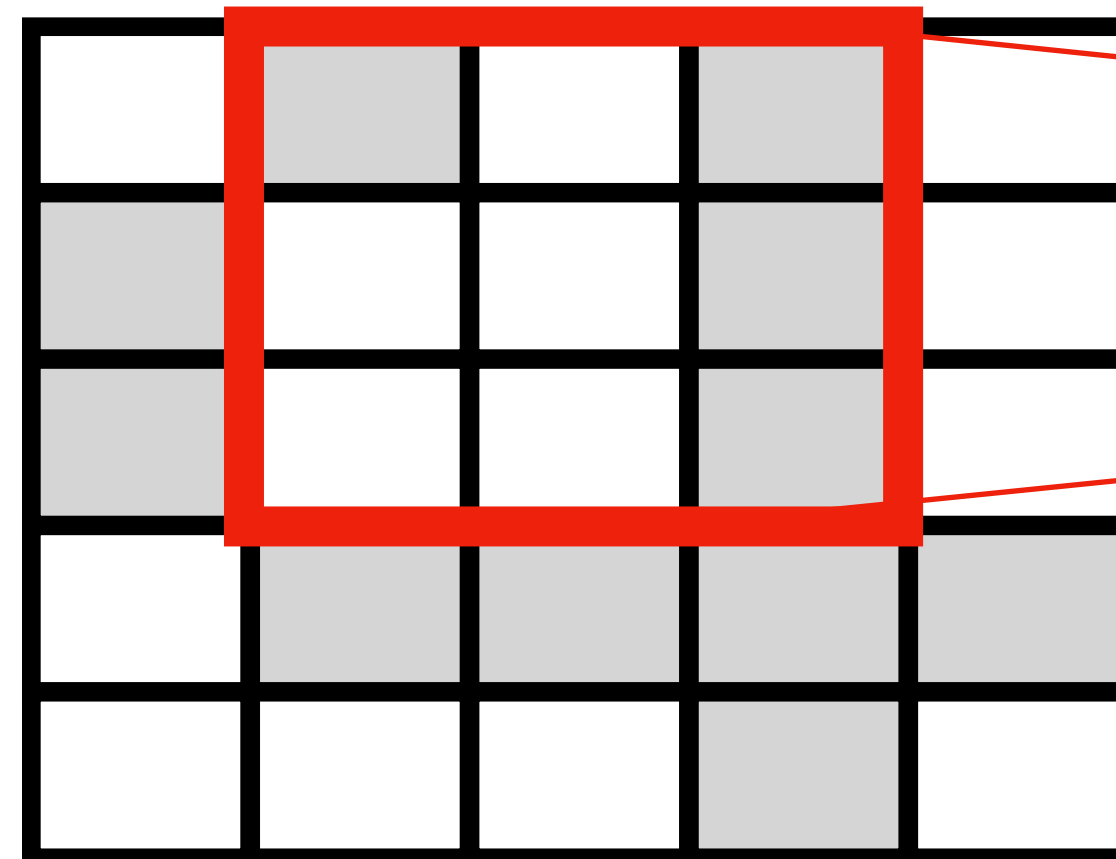


畳み込み層(1)

窓内のデータに学習可能パラメータを乗じて和を取る



$$y_{11} = \sum_{1 \leq i, j \leq 3} w_{ij} x_{ij}$$

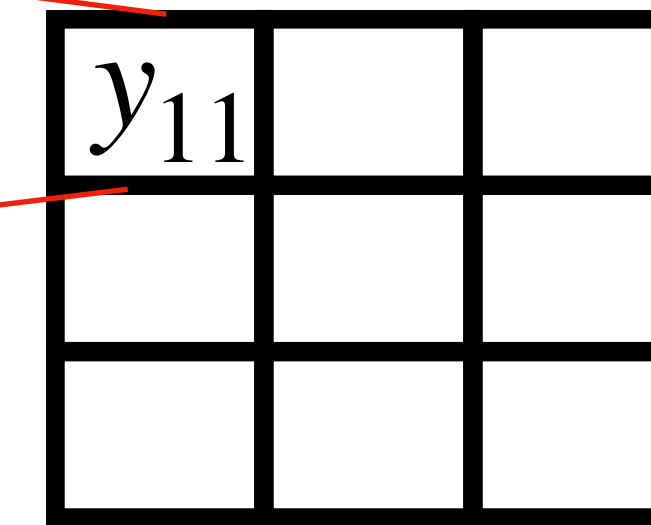
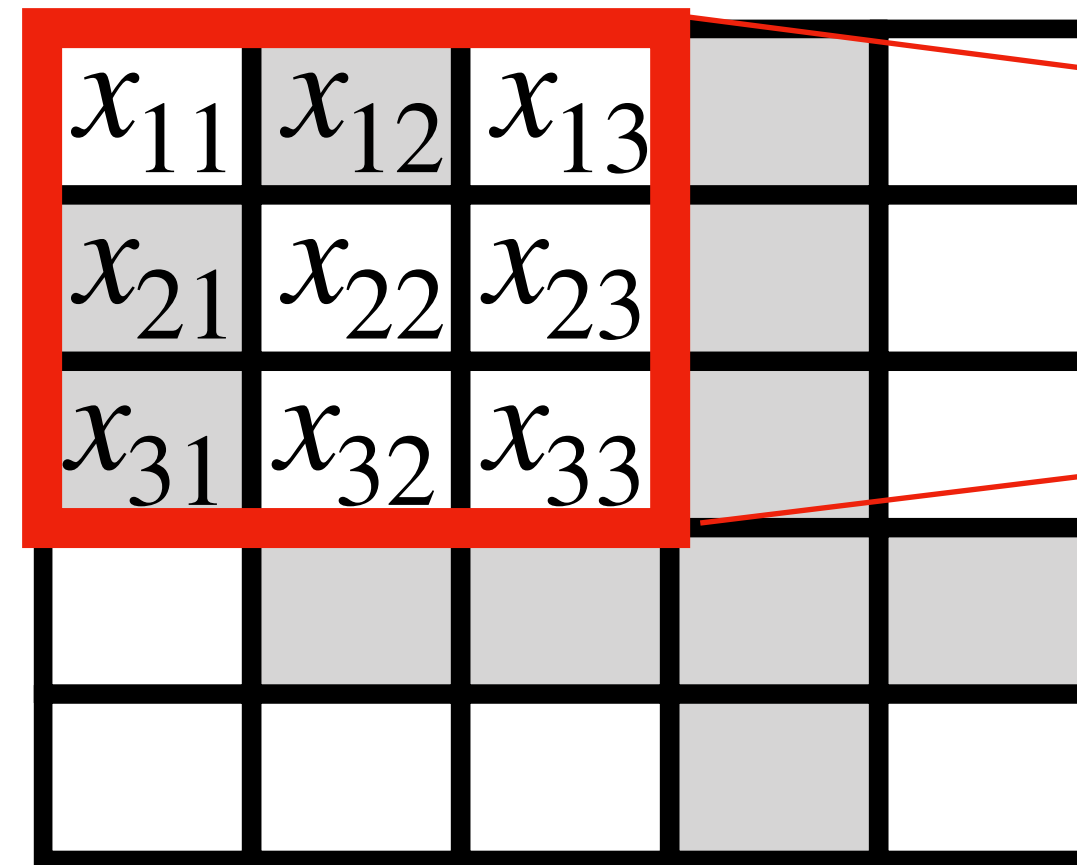


畳み込み層(2)

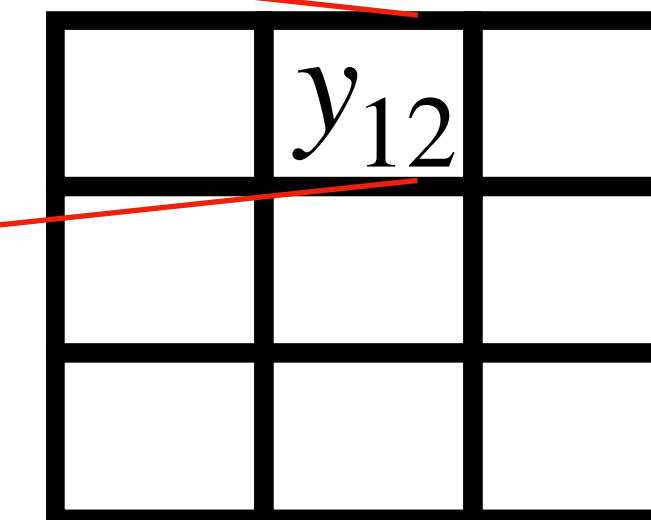
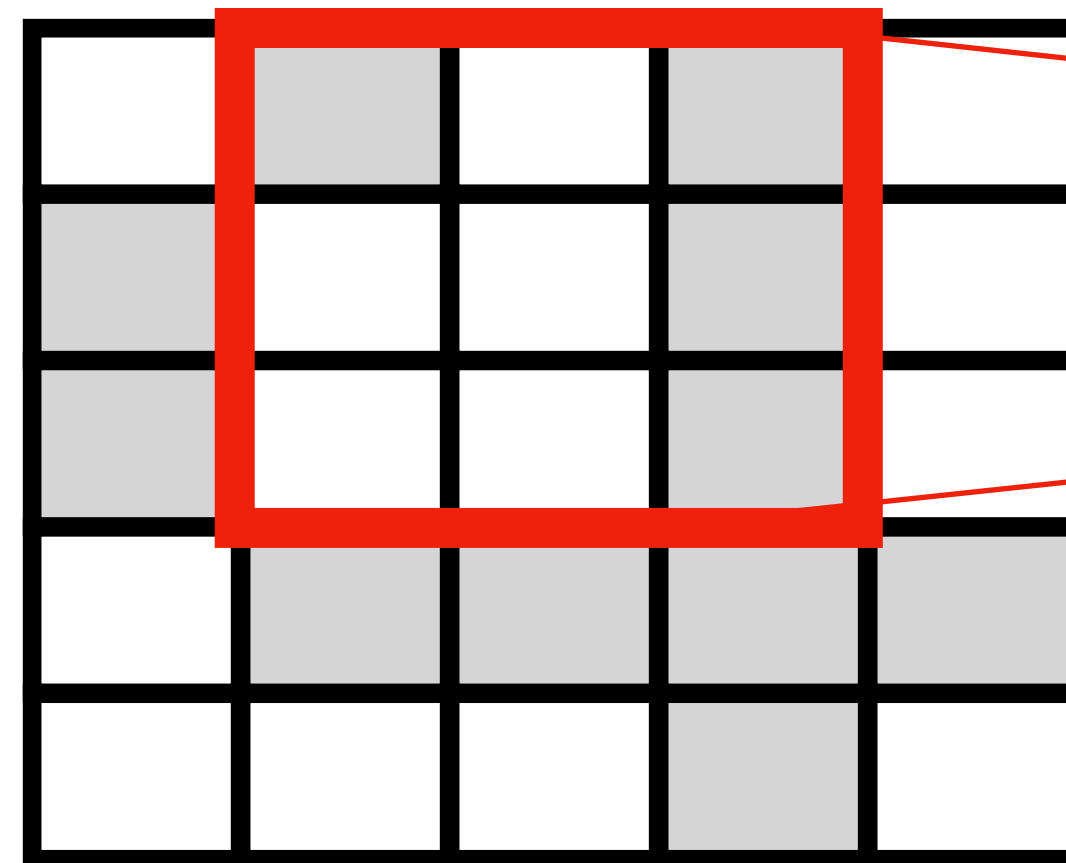
- w_{ij} は学習パラメータ(畳み込みカーネルと呼ばれることもある)であり、これを学習することで画像認識に適切な2次元フィルタが学習される(特徴学習・表現学習)
- 画像の局所的性質を抽出している
- 上の説明は少し簡単化しており、本当は複数枚のチャネルから複数枚チャネルを生成する

プーリング層

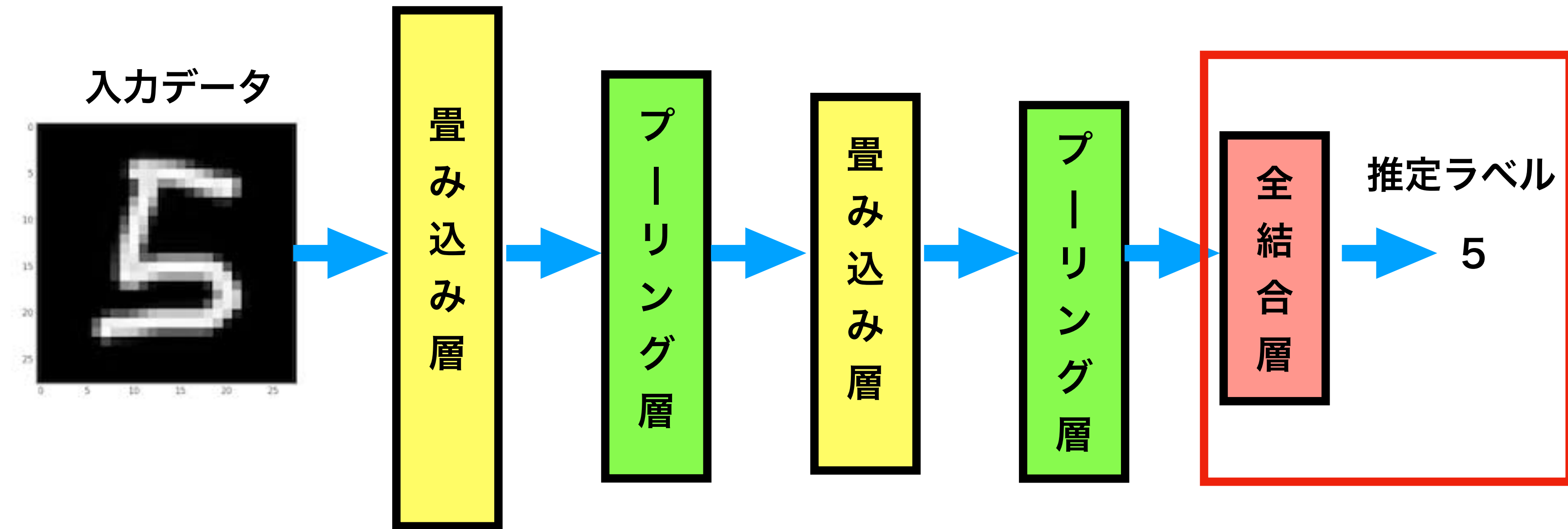
情報を縮約し、層のサイズを縮小させる



$$y_{11} = \max_{1 \leq i, j \leq 3} x_{ij}$$

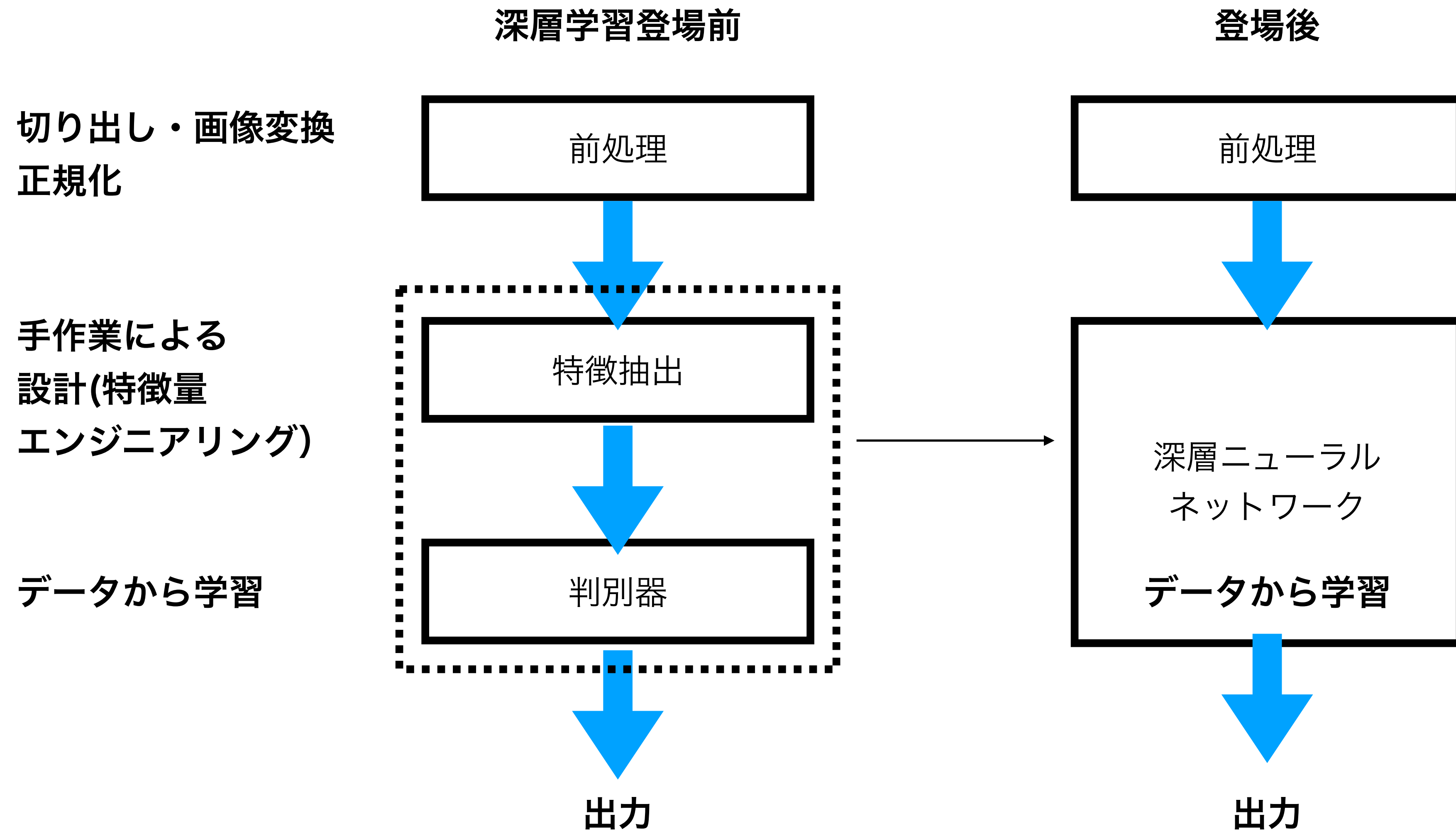


畳み込みニューラルネットワークの構造



- 全結合層は $f(Wx + b)$ 型の通常のNN
- 最終層はソフトマックス層
- 現在においても優れた画像クラス判別のアルゴリズムは畳み込みNN構造

パターン認識構成のためのアプローチ



(過去)特徴量エンジニアリング→(現在) DNNによる表現学習

PyTorchで畳み込みNN

```
class Net(nn.Module):
    def __init__(self):
        super(Net, self).__init__()
        self.conv1 = nn.Conv2d(1, 2, 3) # 28x28x1 -> 26x26x2
        self.conv2 = nn.Conv2d(2, 3, 3) # 26x26x3 -> 24x24x3
        self.pool = nn.MaxPool2d(2, 2) # 24x24x3 -> 12x12x3
        self.fc1 = nn.Linear(12 * 12 * 3, 16)
        self.fc2 = nn.Linear(16, 10)

    def forward(self, x):
        x = F.relu(self.conv1(x)) # 第1畳み込み層
        x = F.relu(self.conv2(x)) # 第2畳み込み層
        x = self.pool(x) # プーリング層
        x = x.view(-1, 12 * 12 * 3)
        x = F.relu(self.fc1(x)) # 全結合層 1
        x = self.fc2(x) # 全結合層 2
        return F.log_softmax(x, dim=1)
```

入力チャンネル数

出力チャンネル数

カーネルサイズ

全結合層

本講義のまとめ

- **畳み込みニューラルネットワーク**