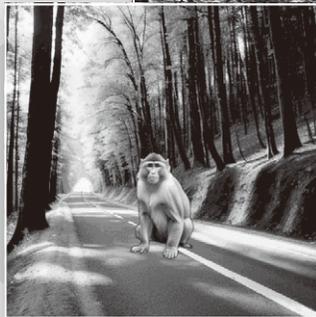
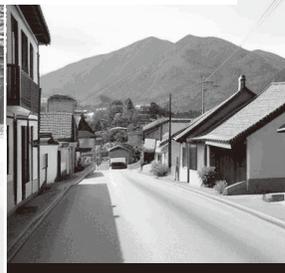


# 生成AI× エッジ・デバイスで AI画像認識

ご購入はこちら



ダウンロード・データあります



岩田 利王

第3回

シンプルな6層CNNの仕組み…  
カラー・ボール認識を例に

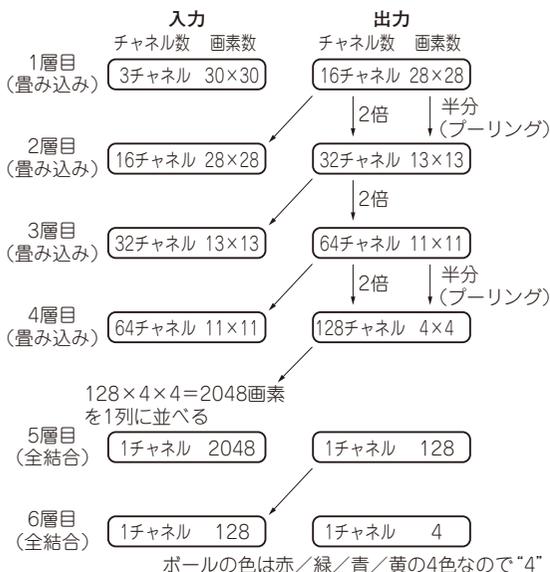


図1 画像認識の流れ～各層のチャンネル数と画素数

前回までの連載ではGoogle Colab<sup>注1</sup>上でPythonプログラムを走らせて次を行いました。

- ①生成AIでカラー・ボールの画像セットを自作
- ②PCで学習。学習済みモデルを生成
- ③PCで推論。認識率や実用性を確認

今回は②③のアルゴリズムとPythonプログラムについて説明します。なお、プログラムは以下のURLから入手できます。

<https://www.cqpub.co.jp/interface/download/contents2025.htm>

また、次号の生成AI特集では次のように進めます。

- ④ラズベリー・パイで推論。USBカメラからの動画

注1: Google Colab (Google Colaboratory) とはブラウザ (Google Chrome 推奨) 上で動く Python 実行環境であり、GPU も無料で使用できます (使用量制限あり)。高価なグラフィックス・ボードのないPCでも生成AIが使えるようになります。

注2: Pythonプログラム内で30×30画素にリサイズします (実際の画像より小さくする)。

のリアルタイム認識

- ⑤FPGAで推論。CMOSイメージセンサからの動画のリアルタイム認識

## ● CNNは畳み込みが4層、全結合が2層からなる

本連載では、画像認識に適したアルゴリズムであるCNN (Convolutional Neural Network, 畳み込みニューラル・ネットワーク) を使用します。図1のように6層からなり、最初の4層は畳み込み層、その後の2層は全結合層といいます。

▶前提条件: 本CNNの入力は3チャンネル、30×30画素

1層目の入力はカラー画像なのでRGB、つまり3チャンネルになります。画素数は30×30です<sup>注2</sup>。畳み込みの場合、層が進むたびにチャンネル数が徐々に増え、画素数は徐々に減っていきます。

▶1～4層目までは畳み込み層

4層目の出力ではチャンネル数は128チャンネルまで増え、画素数は4×4まで減ります。そして全チャンネルの全画素 (128×4×4=2048) が1列に並べられ、5層目の入力となります。

▶5～6層目は全結合層

5層目からは全結合層になります。入力は2048画素です (チャンネル数は1チャンネル)。全結合の場合、層が進むたびに画素数は減っていき、6層目の出力は4画素となります。このCNNはカラー・ボールを認識するもので、クラス数 (ボールの色) が4つ (赤/緑/青/黄) だからです。

▶最終層の出力は4クラスの特徴量となる

4画素のうち最大の値を持つものが推論結果です。

## ● 1層目は畳み込み層

図1のように1層目では3チャンネル→16チャンネルと増え、画素数は30×30→28×28と小さくなります。

▶畳み込みという一種の2次元フィルタ演算

前回生成した学習済みモデルには、3×3の要素を持つ重み係数 (カーネル) が含まれます。カーネルを図2のように横に1個ずらしながら積和演算を行います。このような作業を畳み込みといいます。右端まで