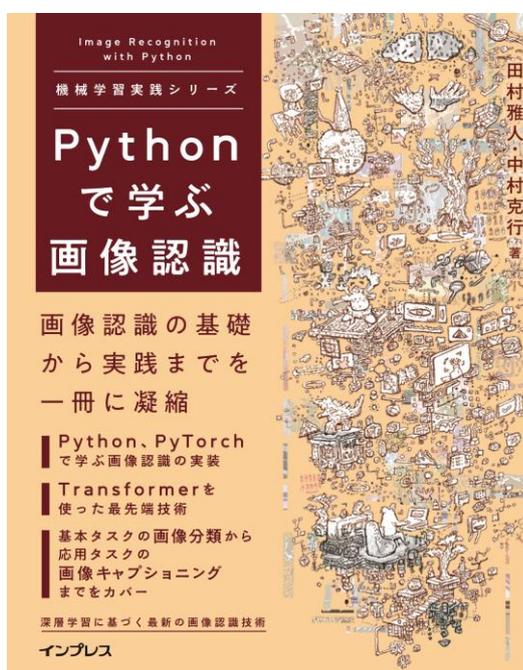


各 位

2023年3月22日
株式会社インプレス**画像認識の基礎から実装までを解説した、中級者以上向けの実践的な技術書
『Pythonで学ぶ画像認識』を3月22日（水）に発売**

インプレスグループでIT関連メディア事業を展開する株式会社インプレス（本社：東京都千代田区、代表取締役社長：小川 亨）は、画像認識技術の基礎から実装までを解説した書籍『Pythonで学ぶ画像認識 機械学習実践シリーズ』を2023年3月22日（水）に発売いたします。

**■画像認識の基礎から実践までを一冊に凝縮**

近年の技術の発展により、画像認識技術はますます身近になっています。スマートフォンの顔認証によるロック解除や、オンライン会議での人物と背景を認識して背景をぼかす技術など、さまざまな場面において画像認識技術は人々の生活をサポートしています。本書では発展著しい画像認識について基礎から実践までを具体的に学ぶことを目的とし、Pythonでの実装を示しながらその手法を丁寧に解説しています。また、最新の画像認識手法について紙幅を割き、多くの部分で深層学習の実装について学ぶことができる構成になっています。

■基本から深層学習を用いた最新技術までカバー

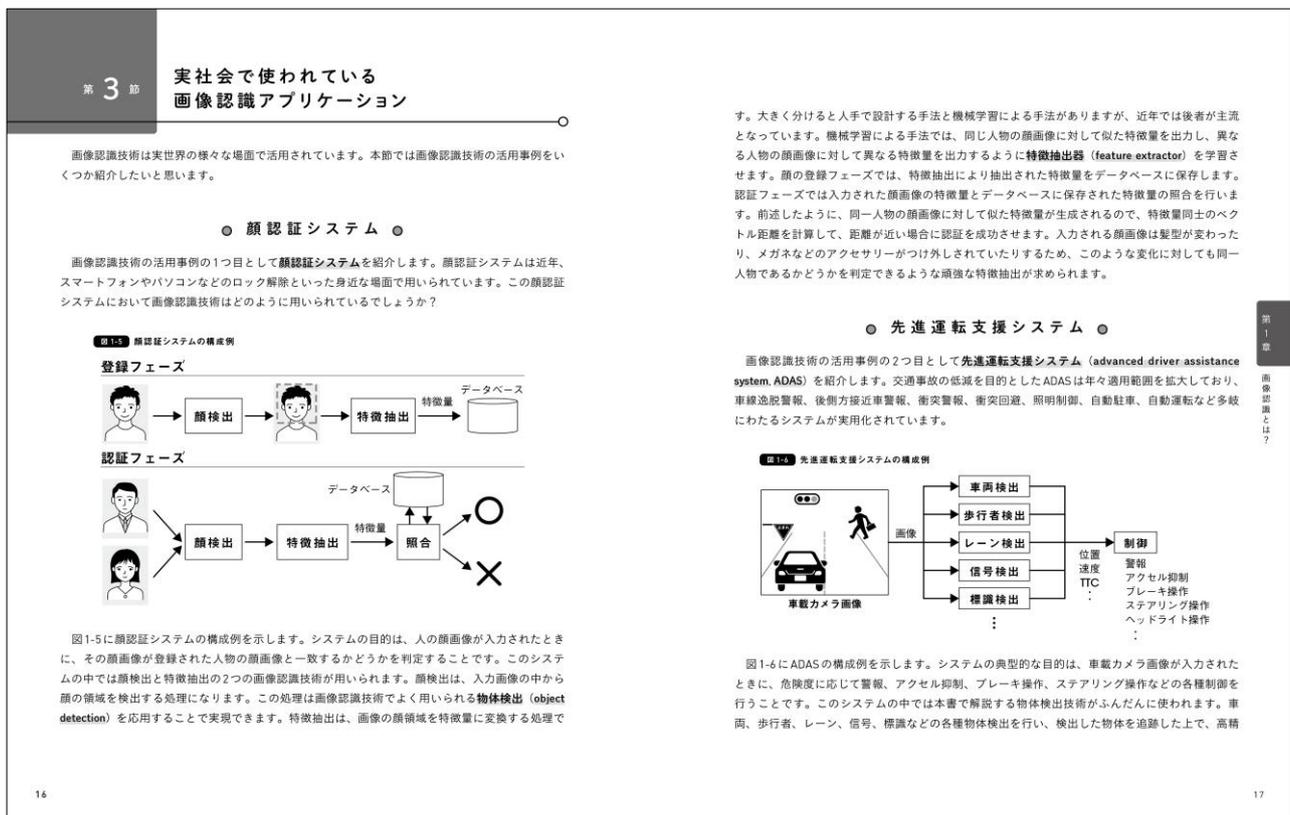
本書は6章構成です。第1章では、画像認識の概要やアプリケーション、本書の前提となる開発環境について解説します。第2章では、画像読み込みやフィルタ処理などの基本的な操作から始め、畳み込み演算やアテンションといった深層学習における特徴抽出の基礎についても解説します。第3章では、単純な画像分類モデルを使って画像認識の学習から評価までの流れを確認します。また、深層学習の基礎を解説し、深層学

習を本格的に使い始める前の準備を行います。第4章では、画像認識の最も基本的なタスクである画像分類について解説します。順伝播型ニューラルネットワーク、畳み込みニューラルネットワーク（CNN）、Transformer という異なる構造を持つ3種類のネットワークを使った手法を解説し、それぞれのネットワークの特徴を理解できるようにします。第5章では、画像認識アプリケーションで広く用いられている物体検出について解説します。この章ではCNNを使った手法とTransformerを使った手法を1つずつ紹介し、それぞれの構造が物体検出でどのような利点を持つのかを解説します。第6章では、画像認識と自然言語処理の融合技術である画像キャプションを解説します。CNNとLong Short-Term Memory（LSTM）の組み合わせによる代表的な手法、アテンション機構を利用した手法、Transformerを応用した手法を解説します。

■本書は以下のような方におすすめです

- ・ 画像処理のエンジニア・研究者
- ・ 理系の大学生
- ・ Pythonで画像処理を実装したい人

■紙面イメージ

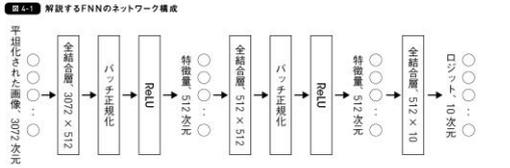


画像認識の基礎から丁寧に解説しています。

まずは DNN の中でも最も基本的なネットワークである**順伝播型ニューラルネットワーク** (feedforward neural network, **FNN**) を使った CIFAR-10 の画像分類器の解説をします。FNN は線形関数を使って前後の層の全ノードを接続する**全結合層** (fully-connected layer) を使うネットワークです。全結合層と ReLU 関数を複数回適用することで表現力の高い非線形な関数を実現します。データ関係の処理は多クラスロジック回帰のものと同様で、まずはモデルの解説から始めます。

● FNN モデルの構造 ●

図 4-1 に本節で解説する FNN のネットワーク構成を示します。今回の FNN では多クラスロジック回帰のときと同様に $32 \times 32 \times 3$ の画像を平坦化した 3072 次元のベクトルを入力とします。FNN において隠れ層の数と隠れ層の特徴量次元はハイパーパラメータとなります。今回は隠れ層の数を 2 とし、隠れ層の特徴量次元を 512 に設定します。活性化関数には第 3.2 節で解説した ReLU 関数を使用します。隠れ層の最後に全結合層を加えて、出力層の 10 クラス分のロジットを生成します。



● バッチ正規化 ●

今回のネットワークでは、全結合層と ReLU 関数の間に**バッチ正規化** (batch normalization) [Ioffe

and Szegedy, 2015] と呼ばれる処理を入れています。バッチ正規化は全結合層や次節で出てくる畳み込み層と活性化関数の間で使われることが多く、学習を安定化させ、収束の速度を上げる効果があります。Ioffe と Szegedy によってバッチ正規化が提案されて以来、ほとんどのモデルでバッチ正規化、あるいはそれに類する正規化が行われており、ニューラルネットワークを学習させるための重要な要素となっています。

バッチ正規化の具体的な計算方法を示します。バッチ正規化は学習時と推論時に異なる計算をします。学習時のミニバッチ入力 $X = \{x_1, x_2, \dots, x_B\}$ に対するデータ x_i の正規化結果 $f_{BN}(x_i)$ は以下のように計算されます。

$$\mu_i = \frac{1}{B} \sum_{j=1}^B x_{ij} \tag{4.1}$$

$$\sigma_i^2 = \frac{1}{B} \sum_{j=1}^B (x_{ij} - \mu_i)^2 \tag{4.2}$$

$$f_{BN}(x_i) = \gamma \odot \left\{ (x_i - \mu_i) \odot \left(\sqrt{\sigma_i^2 + \epsilon} \right) \right\} + \beta \tag{4.3}$$

$$\mu_{avg} = \alpha \mu_{avg} + (1 - \alpha) \mu_i \tag{4.4}$$

$$\sigma_{avg}^2 = \alpha \sigma_{avg}^2 + (1 - \alpha) \sigma_i^2 \tag{4.5}$$

ここで、 ϵ は分母が小さくなりすぎるのを防ぐための定数、 γ および β は出力の平均と分散を調整するための学習により値が決まるパラメータです。通常は初期状態での出力の平均が 0 に、分散が 1 になるように、 γ と β はそれぞれ 1 と 0 で初期化されます。バッチ正規化では、まず式 (4.1) および式 (4.2) により入力データの各次元のバッチ平均と分散を計算します。その後、式 (4.3) で各次元の値がバッチ全体で平均が 0 に、分散が 1 になるように標準化し、 γ および β で出力の分散と平均を調整します。また、式 (4.4) および式 (4.5) により、バッチ平均と分散の**指数平滑移動平均** (exponential moving average, EMA) を計算します。 α は更新の度合いを調整するハイパーパラメータになります。移動平均はミニバッチが入力されるたびに更新され、最終的な移動平均の値は推論時に使われます。

推論時には式 (4.1) および式 (4.2) の計算は行いません。式 (4.1) および式 (4.2) で計算されるバッチ平均と分散はミニバッチを構成するデータによって変化するため、入力されるミニバッチのサンプル構成によってバッチ正規化の結果が変化します。そのような正規化を推論時に行うとミニバッチの構成によって推論結果が変わってしまいます。これを防ぐために、推論時には学習時に計算した移動平均を使ってバッチ正規化の計算をします。計算は以下ようになります。

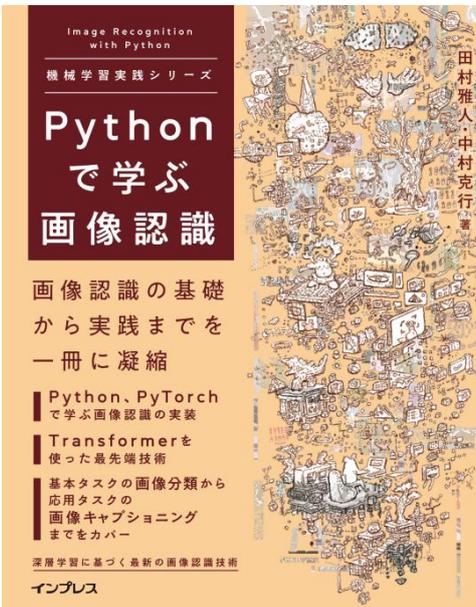
1: 平均が 0、分散が 1 となるようにスケールリングすることは標準化と呼ばれることが多いですが、技術的にはバッチ正規化と呼ばれます。

最新の画像認識技術を実装レベルで学ぶことができます。

■本書の構成

- 第 1 章 画像認識とは？
- 第 2 章 画像処理の基礎知識
- 第 3 章 深層学習を使う準備
- 第 4 章 画像分類
- 第 5 章 物体検出
- 第 6 章 画像キャプションング

■書誌情報



書名：Python で学ぶ画像認識 機械学習実践シリーズ

著者：田村雅人・中村克行

発売日：2023 年 3 月 22 日（水）

ページ数：352 ページ

サイズ：B5 変形判

定価：3,850 円（本体 3,500 円＋税 10%）

電子版価格：3,850 円（本体 3,500 円＋税 10%）※インプレス直

販価格

ISBN：978-4-295-01599-4

◇Amazon の書籍情報ページ：

<https://www.amazon.co.jp/dp/4295015997/>

◇インプレスの書籍情報ページ：

<https://book.impress.co.jp/books/1122101074/>

■著者プロフィール

田村雅人（たむら・まさと）

2016年に東京大学大学院工学系研究科修士課程を修了し、(株)日立製作所に入社。パブリックセーフティ向け映像認識技術の研究開発に従事し、人物検出や追跡、人と物体の関係性検出など、映像監視に使われるAI技術を開発。2021年よりアメリカ・シリコンバレーにオフィスをおく Hitachi America, Ltd.に出向し、世界トップレベルの研究者たちと映像認識技術の開発に従事。

中村克行（なかむら・かつゆき）

2007年 東京大学大学院 新領域創成科学研究科 博士課程修了。同年、(株)日立製作所 中央研究所入社。物体追跡、一人称視点映像解析、マルチモーダル認識などの研究開発に従事し、2022年より同社研究開発グループ 先端 AI イノベーションセンタ 知能ビジョン研究部長。2015-2016年 スタンフォード大学コンピュータサイエンス学科 客員研究員。2005年 FIT ヤングリサーチャー賞、2011年 IEEE Consumer Electronics Society Best Paper Award in Television Technology など受賞。

■機械学習実践シリーズについて

中級者以上に向けた、特定の技術分野のアルゴリズムの紹介と、そのアルゴリズムを実装したコードを解説する、より技術的・実践的なシリーズです。

以上

【株式会社インプレス】 <https://www.impress.co.jp/>

シリーズ累計7,500万部突破のパソコン解説書「できる」シリーズ、「デジタルカメラマガジン」等の定期雑誌、IT関連の専門メディアとして国内最大級のアクセスを誇るデジタル総合ニュースサービス「Impress Watch シリーズ」等のコンシューマ向けメディア、「IT Leaders」、「DIGITAL X」、「Web 担当者 Forum」等の企業向け IT 関連メディアブランドを総合的に展開、運営する事業会社です。IT 関連出版メディア事業、およびデジタルメディア&サービス事業を幅広く展開しています。

【インプレスグループ】 <https://www.impressholdings.com/>

株式会社インプレスホールディングス（本社：東京都千代田区、代表取締役：松本大輔、証券コード：東証スタンダード市場 9479）を持株会社とするメディアグループ。「IT」「音楽」「デザイン」「山岳・自然」「航空・鉄道」「モバイルサービス」「学術・理工学」を主要テーマに専門性の高いメディア&サービスおよびソリューション事業を展開しています。さらに、コンテンツビジネスのプラットフォーム開発・運営も手がけています。

【本件に関するお問合せ先】

株式会社インプレス 広報担当：丸山

E-mail: pr-info@impress.co.jp URL: <https://www.impress.co.jp/>

※弊社はテレワーク推奨中のため電話でのお問い合わせを停止しております。メールまたは Web サイトからお問い合わせください。