

Implementation of Training Convolutional Neural Networks

備考

著者

Tianyi Liu, Shuangfang Fang, Yuehui Zhao, Peng Wang, Jun Zhang

掲載

arXiv preprint arXiv:1506.01195 [cs], 2015.

Abstract

深層学習とは、表現のレベルを学習することを基本とする機械学習の輝かしい一分野を指す。畳み込みニューラルネットワーク（CNN）は、ディープニューラルネットワークの一種である。並行して学習することができる。この記事では、CNNアルゴリズムのプロセスについて、前進プロセスと逆伝播の両方を詳しく分析しました。そして、典型的な顔認識問題をjavaで実装するために、特定の畳み込みニューラルネットワークを適用しました。そして、セクション4では、並列戦略を提案した。さらに、前進計算と後退計算の実時間を計測することで、最大のスピードアップと並列効率を理論的に分析した。

1. Intro

深層学習とは、機械学習の一分野であり、特徴や要素、概念の階層に対応する表現レベルの学習に基づいています。ここでは、下位レベルの概念から上位レベルの概念が定義され、同じ下位レベルの概念が多くの上位レベルの概念を定義するのに役立ちます。

深層学習とは、複数のレベルの表現や抽象化を学習することで、画像、音声、テキストなどのデータを理解するのに役立ちます。深層学習の概念は、人工ニューラルネットワークの研究から生まれたもので、隠れた層を多く含む多層パーセプトロンが深層学習の構造です。1980年代後半、人工ニューラルネットワークに用いられるバックプロパゲーションアルゴリズムの発明は、機械学習に希望をもたらし、統計モデルに基づいた機械学習の流れを作りました。

1990年代に入ると、SVM（Support Vector Machines）、Boosting、LR（Logistic Regression）など、さまざまなShallow Learningモデルが提案されました。これらのモデルの構造は、1つの隠れたノード（SVM、Boosting）、または隠れたノードのないモデル（LR）として見ることができます。これらのモデルは、理論的な分析と応用の両面で大きな成功を収めています。

2006年、カナダ・トロント大学の教授で機械学習の学部長であるジェフリー・ヒントんと、彼の学生であるルスラン・サラクチノフが「Science」誌に発表した論文が、学术界や産業界における機械学習の流れを作った。この記事のポイントは2つ。学習によって得られた特性は、データに対して本質的な説明を行い、視覚化や分類を容易にする。2) ディープニューラルネットワークの学習の難しさは、層ごとの事前学習によって克服できる。この記事では、教師なし学習によって、層別の事前学習を実現しています。*人工ニューラルネットワークの中でも、隠れた層を複数持つフィードフォワード・ニューラルネットワークや多層パーセプトロンは、通常、ディープ・ニューラル・ネットワーク（DNN）と呼ばれています。*畳み込みニューラルネットワーク（CNN）はフィードフォワード型ニューラルネットワークの一種である。1960年代、HubelとWieselは、猫の視覚系で局所的に敏感な方向選択に使われるニューロンを研究し、特殊なネットワーク構造が効果的に複雑さを減らすことができることを発見しました。

フィードバックニューラルネットワークの複雑さを効果的に軽減できることを発見し、Convolution Neural Networkを提案した。*CNNは、パターン認識や画像処理に広く用いられている効率的な認識アルゴリズムである。構造が単純で、学習パラメータが少なく、適応性が高いなどの特徴があります。*音声解析や画像認識の分野で注目されています。重みを共有したネットワーク構造により、生物学的なニューラルネットワークに近いものとなっています。ネットワークモデルの複雑さや重みの数を減らすことができます。

一般的にCNNの構造は2つの層で構成されています。1つは特徴抽出層で、各ニューロンの入力はその層の局所受容野に接続され、局所的な特徴を抽出します。局所的な特徴が抽出されると、その特徴と他の特徴との位置関係も決定されます。もう1つは特徴量マップ層で、ネットワークの各計算層は複数の特徴量マップで構成されています。すべての特徴マップは平面であり、平面内のニューロンの重みは等しくなっています。特徴量マップの構造は、畳み込みネットワークの活性化関数としてシグモイド関数を使用しており、これにより特徴量マップはシフト・インバリアンسを持つ。また、同じマッピング平面内のニューロンは重みを共有しているので、ネットワークの自由パラメータの数を減らすことができます。畳み込みニューラルネットワークの各畳み込み層の後には、局所平均と第2抽出

の計算を行う計算層があり、このユニークな2つの特徴抽出構造により、解像度を下げることができます。

CNNは主に、2次元グラフィックスの変位やズームなどの歪曲不変性を識別するために使用されます。CNNの特徴検出層は学習データによって学習するので、CNNを使うときは明示的な特徴抽出を避け、学習データから暗黙的に学習します。さらに、同じ特徴マッププレーンにあるニューロンは同一の重みを持っているので、ネットワークは並行して学習することができます。これは、互いに接続されたニューロンネットワークに対して、畳み込みネットワークの大きな利点です。CNNの局所的な共有重みの特殊な構造のため、**音声認識や画像処理において、CNNはユニークな利点を持っています。**そのレイアウトは、実際の生物学的なニューラルネットワークに近いものです。重みを共有することで、ネットワークの複雑さを軽減することができます。特に多次元の入力ベクトル画像を直接ネットワークに入力することができるので、特徴抽出や分類プロセスにおけるデータ再構築の複雑さを回避することができます。

顔認識は、人物の顔の特徴に基づいた生体認証技術です。顔認証システムの研究は1960年代に始まり、1980年代後半にはコンピュータ技術や光学的画像処理技術の発展に伴い改良が進み、1990年代後半には本格的な初期応用の段階に入っています。モニタリングシステムなどの実用的なアプリケーションでは、カメラで撮影された顔画像は低解像度であることが多く、ポーズのばらつきも大きい。姿勢の変化や低解像度の影響を受けると、顔認識の性能は急激に低下します。また、ポーズのばらつきは、顔認識に大きな課題をもたらします。顔認識には非線形な要素が含まれます。また、既存の機械学習法の中には、浅い構造を用いるものがあります。深層学習は、深い非線形ネットワーク構造によって、複雑な関数の近似を実現することができます。

この記事では、畳み込みニューラルネットワークを使って顔認識を解決します。畳み込みニューラルネットワークは、顔認識におけるポーズや解像度の影響を克服することができます。畳み込みニューラルネットワークは、学習時間が長く、認識計算量が大きいため、リアルタイムの要求を満たすことが難しく、また、遅延が許容範囲を超えてしまうことがあります。そこで、クラウド・プラットフォームを使用して、コンピューティング・プロセスを同時進行で高速化しています。

3. 驚異的なニューラルネットの原理

3.1. 方法論

****コンボリューション・ニューラル・ネットワーク・アルゴリズムは、2次元の画像情報を識別するために特別に設計された、多層パーセプトロンです。***入力層、畳み込み層、サンプル層、出力層と、常に複数の層を持っています。さらに、ディープネットワークのアーキテクチャでは、畳み込み層とサンプル層が複数になることもあります。CNNは、制限されたボルツマンマシンとしてではなく、す

すべての接続のために隣接する層のニューロンの層の前と後にする必要があります、畳み込みニューラルネットワークのアルゴリズムは、各ニューロンは、グローバルな画像を感じる行う必要はありませんが、ちょうど画像のローカルエリアを感じる。また、各ニューロンのパラメータは、同じに設定されている、すなわち、重みの共有、すなわち、同じ畳み込みカーネルを持つ各ニューロンは、画像をデコンボリューションする。

CNNアルゴリズムには、畳み込み処理とサンプリングという2つの主要なプロセスがある。

畳み込み処理: 学習可能なフィルタ F_x を使い、入力画像をデコンボリューションし（第一段階は入力画像、畳み込み後の入力はいずれも各層の特徴画像、すなわちフィーチャマップ）、バイアス b_x を加えて畳み込み層 C_x を得る。

サンプリング処理: 各近傍の n 個の画素がプーリングステップを経て1個の画素となり、スカラー重み付け W_{x+1} で重み付けされ、バイアス b_{x+1} を加えた後、活性化関数により、狭い n 回の特徴マップ S_{x+1} を生成します。