



画像認識革命：CNNについて

このプレゼンテーションでは、画像認識技術の現状と課題をご紹介します。

私たちの身の回りで急速に進化している画像認識技術。その中心となるCNNについて詳しく解説します。



投稿者：Suwa

画像認識とは？なにに使われている？

基本概念

画像認識とは、コンピュータが画像内の物体や特徴を識別する技術です。

人間の視覚を模倣し、デジタル画像から意味のある情報を抽出します。

主要応用分野

- 医療診断支援
- 自動運転車
- 製造業の品質管理
- 顔認証システム



CNNとは？

畳み込みニューラルネットワーク

Convolutional Neural Networkの略称です。画像認識に特化した深層学習モデルです。

生物の視覚野にヒント

人間の視覚野の仕組みを模倣しています。局所的な特徴（たとえば、線・角・模様など）を検出し、徐々に抽象化します。

画像データの扱い方（後ほど詳しく説明）

画像を直接入力として扱います。従来のニューラルネットワークと異なり、空間的構造を保持します。

特徴抽出の自動化（後ほど説明）

特徴量を自動的に学習します。手作業による特徴設計が不要になりました。

CNNの登場前の手法の限界

画像認識技術は長い歴史がありますが、CNNが登場する以前の手法には大きな課題がありました。

手作業の特徴設計

エッジ検出やHOG特徴など、専門家が特徴量を設計する必要があります。汎用性に欠け、新しい対象に適応するのが困難でした。



複雑なパターン認識の限界

複雑な視覚パターンの抽象化が困難でした。人間には簡単な認識タスクでもコンピュータには難しかったのです。



スケール変化への弱さ

対象物の大きさや向きが変わると認識精度が大幅に低下します。実世界の変動に対応できませんでした。

階層的特徴学習の欠如

低レベルの特徴から高レベルの概念へと段階的に学習する仕組みがありませんでした。

(抽象化と具体化が苦手)

CNNの登場前の手法の限界：もう少し詳細に



手作業の特徴設計の困難さ

専門家がエッジやテクスチャなどの視覚的特徴を手動で定義する必要がありました。

新しい対象ごとに特徴を再設計するため、時間と専門知識が必要でした。



階層的特徴学習の欠如

低レベル特徴（線や色）から高レベル特徴（形や物体）への自然な抽象化ができませんでした。

人間の視覚系のような階層的処理能力がなく、複雑なパターン認識が困難でした。



環境変化への脆弱性

照明条件や背景の変化、物体の回転に対して極めて脆弱でした。

実世界の多様な状況に適応できず、実用的な応用が限られました。

これらの制約により、従来の画像認識手法は人間レベルの認識精度を達成できませんでした。CNNはこれらの問題を根本から解決する画期的なアプローチとなったのです。



CNNの登場：画像認識のブレイクスルー

従来技術の限界

手作業による特徴抽出に依存していました。

複雑な画像パターンの認識に課題がありました。

画像認識の革命

2012年のImageNetコンテストでAlexNetが圧勝しました。

以降、画像認識精度が急速に向上しています。

1

2

3

CNNの誕生

1998年、LeCunによって基本概念が提案されました。

生物の視覚野からインスピレーションを得ています。

CNNの転換点：2012年のImageNet革命

2012年、Alexei KrizhevskyらがAlexNetを用いてImageNetコンテストで圧勝した。従来の認識誤差率26%から16%へと一気に改善され、深層学習の可能性を世界に示した。

1

ImageNetチャレンジ
100万枚以上の画像、1000
カテゴリーを分類する世界
最大の画像認識コンテス
ト。

2

AlexNetの圧勝
次点の手法との差は10%以
上。他の参加者は従来技術
を使用していた。

3

技術的革新
5層の畳み込み層と3層の全
結合層からなる複雑なネッ
トワークを構築。

4

影響
この勝利を契機に、深層学
習が研究・産業界の主流と
なった。



Made with Gamma

CNNの仕組み

畠み込み層

画像から特徴を抽出します。フィルタを使って重要なパターンを検出します。

活性化関数層

ニューラルネットワークが複雑なパターンを学習できるように、非線形な変換を加えて情報の通り道に柔軟性を持たせる

プーリング層

特徴マップの次元を削減します。計算量を減らし、過学習を抑制します。

全結合層

最終的な分類を行います。畠み込みとプーリングで抽出された特徴を統合します。

ドロップアウト層

学習テクニック。ランダムに一部のニューロン（情報）を無効化して、過学習を防ぐ

CNNの学習：誤差逆伝播法



フィルタ適用

画像の一部を読み込む



積和演算

要素ごとに積と和を計算します。



活性化関数（代表例 ReLU）

非線形変換で表現力を向上させます。

⇒弱い信号はスルーして、強い信号だけ通す



誤差逆伝播

誤差を修正し精度を高めます。

① 画像の左上にフィルタを重ねる（ 3×3 の範囲を切り出す）

```
lua
```

コピーする 編集する

対象の範囲：

```
[[10, 10, 20],  
 [20, 20, 30],  
 [30, 30, 40]]
```

② 対応する値をかけて、全部足す（要素ごとの掛け算）

コピーする 編集する

```
(10x1) + (10x0) + (20x-1) +  
(20x1) + (20x0) + (30x-1) +  
(30x1) + (30x0) + (40x-1)  
  
= 10 + 0 + (-20) + 20 + 0 + (-30) + 30 + 0 + (-40)  
= -30
```

③ 結果を出力画像の最初の位置に記録

```
ruby
```

コピーする 編集する

出力画像（まだ途中）：
[[-30, ??, ??],
 [??, ??, ??],
 [??, ??, ??]]

④ フィルタを右にずらしてまたやる

次はこの範囲：

```
lua
```

コピーする 編集する

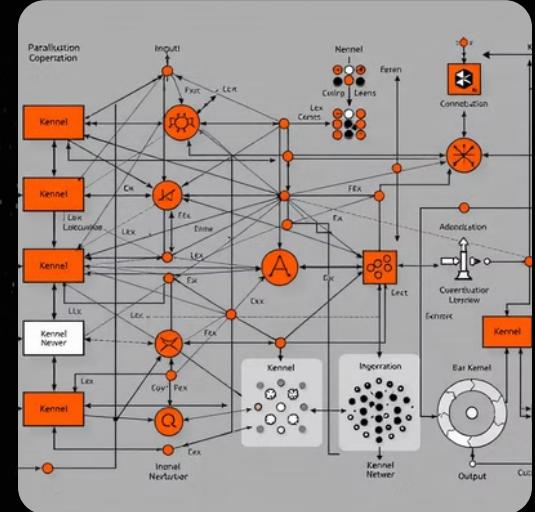
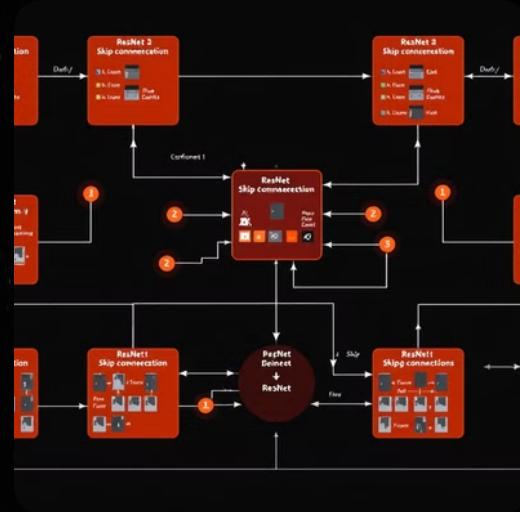
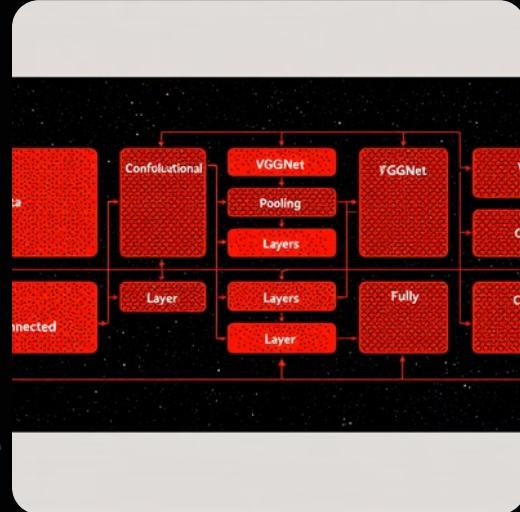
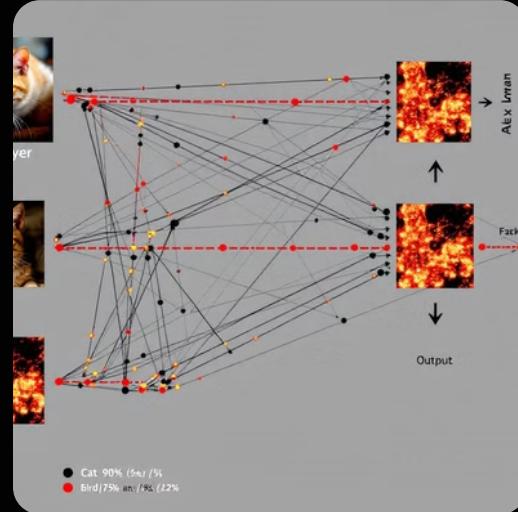
```
[[10, 20, 30],  
 [20, 30, 40],  
 [30, 40, 50]]
```

同じように掛け算 & 合計を計算して、出力画像に書き込む！

⑤ 最終的に出てくる「出力画像」が特徴マップ（feature map）

この特徴マップは、「この縦エッジフィルタにどの部分がどれだけ反応したか」を表して！
つまり、「この画像のどこに縦線っぽい模様があるの？」っていう**感度マップ**だと思ってOK！

CNNの進化：代表的なアーキテクチャ



1 LeNet (1989)

CNNの基本構造を世界に示した最初のモデル

2 AlexNet (2012)

深層学習ブームの火付け役。5つの畳み込み層と3つの全結合層を持ちます。画像認識コンテストで圧倒的な勝利を収めました。

3 ResNet (2015)

スキップ接続を導入し、非常に深いネットワークの学習を可能にしました。現在も広く使われています。

4 EfficientNet (2019~)

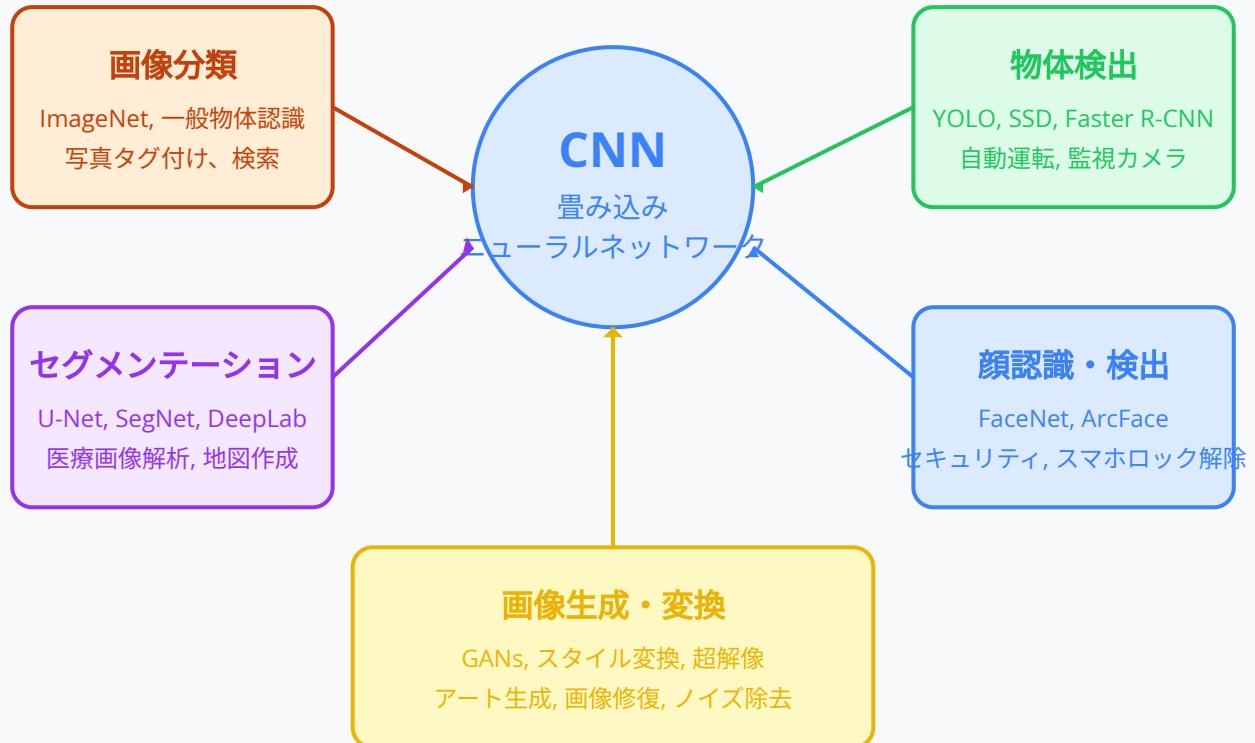
Googleが2019年に発表

層深さ、層の幅、画像サイズの3つをバランスよく調整して最適化します。

精度高い × 軽い × 速いなのでモバイルなど性能が低くても適用可能

CNN の応用分野

CNNの主な応用分野



CNN技術は画像認識から始まり、現在では様々なコンピュータビジョントスクの基盤となっています

CNNの課題と今後の展望

データ依存性

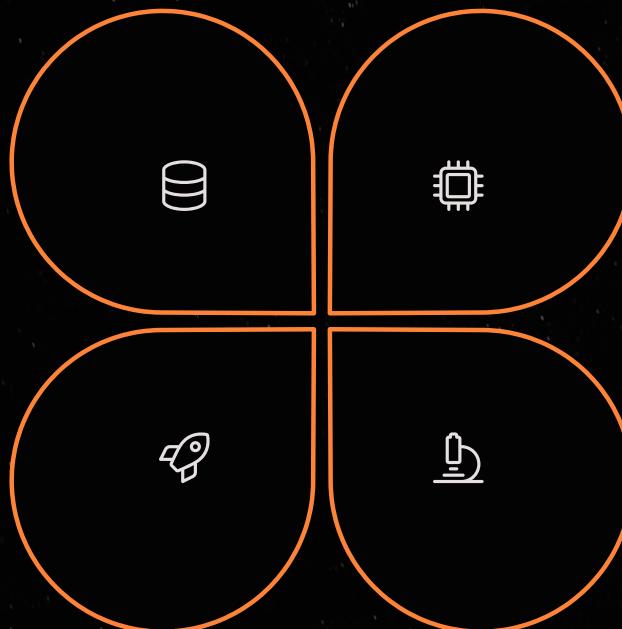
大量の学習データが必要です。

データ収集・データのラベリング
のコストが課題です。

学習の難しさ

手動でのデータのラベリングが必要。

自己教師あり学習の発展が期待されます。



計算コスト

高精度モデルは計算リソースを多く消費します。

スマートフォンや家電での効率的な実行に
課題があります。

説明可能性

深層 CNN はブラックボックス的な性質を持ち、判断根拠の説明が難しいです。

重要な意思決定への応用に制限があります。

CNNとTransformer の融合モデル

CNN（軽さ・効率性）×Transformer（柔軟性・全体把握力）

医療画像、ロボティクス、動画解析などで期待

項目	CNN（畳み込みニューラルネット）	Transformer
得意なこと	局所的な特徴抽出 (画像の縁・模様など)	グローバルな依存関係の理解（全体構造や文脈）
処理範囲	限られた範囲（カーネルでスライド）	全入力を一度に見る (Self-Attention)
主な用途	画像認識、物体検出など	言語処理（GPTなど）、画像認識（ViT）
欠点	全体を俯瞰する力が弱い	局所的な詳細をとらえるのがやや苦手、重い





まとめ

CNNの革命的影响

画像認識技術を飛躍的に進歩させました。医療、自動運転など多くの分野に変革をもたらしています。

技术の発展

畳み込み、プーリング、全結合層の連携により複雑な特徴抽出を可能にしました。各種アーキテクチャの進化により精度が向上しています。

未来への展望

データ効率や計算コストの課題を解決する新技术が登場しています。より軽量で高精度なモデルが今後も発展するでしょう。