

AI 特許紹介(28)  
AI 特許を学ぶ！究める！  
～MASK R-CNN 特許～

2021年5月10日  
河野特許事務所  
所長 弁理士 河野英仁

「AI 特許紹介」シリーズは、注目すべき AI 特許のポイントを紹介します。熾烈な競争となっている第4次産業革命下では AI 技術がキーとなり、この AI 技術・ソリューションを特許として適切に権利化しておくことが重要であることは言うまでもありません。

AI 技術は Google, Microsoft, Amazon を始めとした IT プラットフォーマ、研究機関及び大学から毎週のように新たな手法が提案されており、また AI 技術を活用した新たなソリューションも次々とリリースされています。

本稿では米国先進 IT 企業を中心に、これらの企業から出願された AI 特許に記載された AI テクノロジー・ソリューションのポイントをわかりやすく解説致します。

## 1.概要

特許権者 Facebook

出願日 2018年3月15日

登録日 2020年7月14日

登録番号 US10713794

発明の名称 オブジェクトインスタンスのセグメンテーションに機械学習を使用するための方法とシステム

794 特許は、オブジェクトの検出と分類をピクセル単位で行う MASK R-CNN に関する。

## 2.特許内容の説明

図1は、Mask R-CNN のフレームワークを示す。

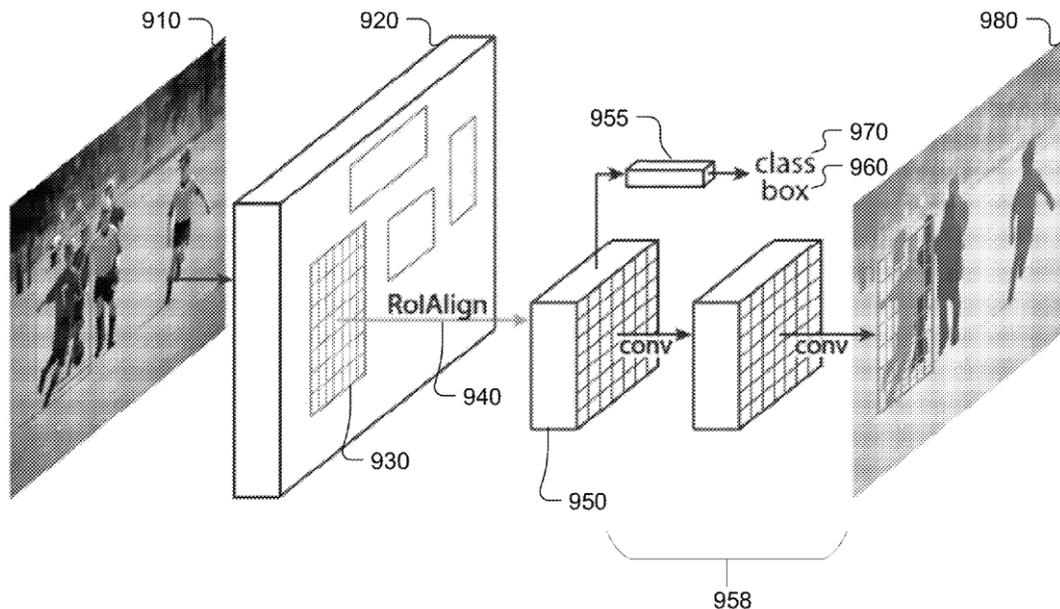


図 1

図 1 では、サッカーをしている数人の人間を示す入力画像 910 の例を示している。ResNet、または、特徴ピラミッドネットワーク (FPN) 等の畳み込みバックボーン深層ネットワークを用いて、入力画像 910 を処理し、特徴マップ 920 を生成する。

バックボーンの出力はまた、領域提案ネットワーク (RPN) によって使用され、特徴マップ 920 内の領域にマッピングされる任意の数の RoI930 を識別する。次に、RoI930 などの個々の RoI は、RoIAlign940 と呼ばれるプロセスを使用して処理され、より小さな RoI または領域特徴マップ 950 (例えば、 $7 \times 7$ ) を抽出する。より小さな領域特徴マップは、固定された所定のサイズ (例えば、 $n \times n$  または  $n \times m$ ) を有する。

次に、RoI 特徴マップ 950 は、(1)オブジェクト 960 を (バウンディングボックスを使用して) 検出し、(2)オブジェクト 970 を (例えば、人、犬などとして) 分類し、(3)セグメンテーションマスクを介してオブジェクト 980 をセグメント化する。

これらの 3 つのタスクのそれぞれは、畳み込みニューラルネットワークなどの 1 つまたは複数のニューラルネットワーク層を使用して実行することができる。例えば、全結合層 955 は、物体検出 960 および物体分類 970 に使用される。セグメンテーションマスクブランチのためのネットワーク 958 は、各 RoI に適用され、ピクセル毎にセグメンテーションマスクを予測する完全畳み込みネットワーク (FCN) である。

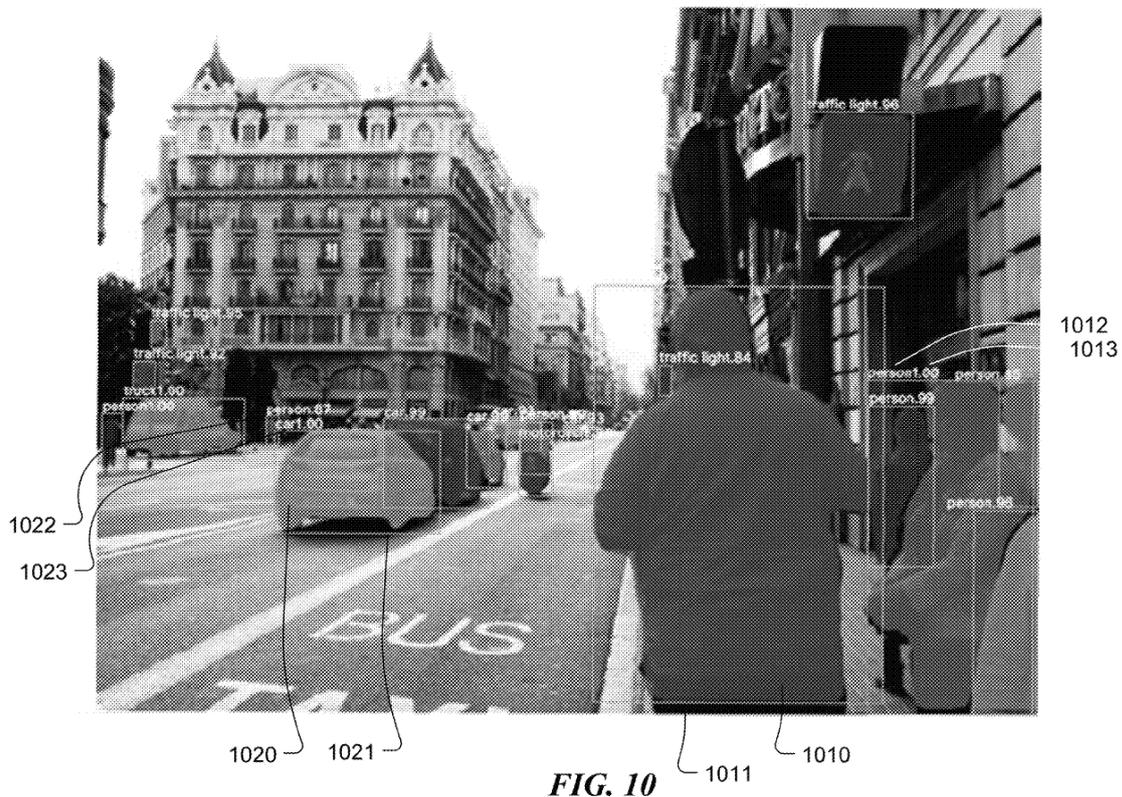


図 2

図 2 は、Common Objects in Context (COCO) テストセット (特定のオブジェクトの様々な画像を含む) に対して Mask R-CNN を使用して画像処理を実行した結果の例を示している。

図 2 は、分類のためのグラウンドトゥールースラベル (例えば、「車」または「人」のラベル)、バウンディングボックスオブジェクト検出 (例えば、車などのオブジェクトの周りのバウンディングボックス)、および、インスタンスセグメンテーション (たとえば、異なる色で表され、各オブジェクトインスタンスに対応する特定のピクセルを識別するマスク) を備えたトレーニング画像を示している。

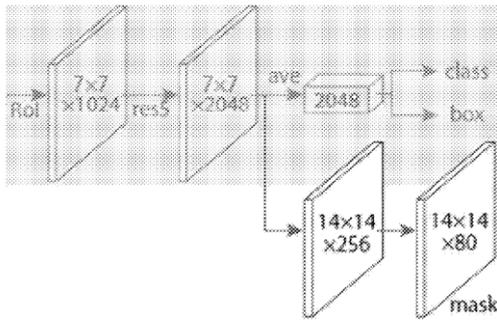


FIG. 13A

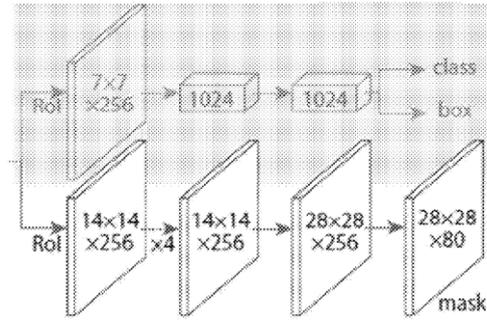


FIG. 13B

図 13A ResNet C4 ネットワークヘッド

図 13B FPN ネットワークヘッド

図 13A は ResNet C4 ネットワークヘッド、図 13B は FPN ネットワークヘッドを示す。両方の図で、それぞれのヘッドは RoI を入力として受け取り、3 つの予測を出力する。RoI で検出されたオブジェクトの分類 (class)、オブジェクトのバウンディングボックス (box)、および、オブジェクトのセグメンテーションマスク (mask) である。

図の数字は空間分解能とチャネルを示し、矢印はコンテキストから推測できる畳み込み、デコンボリューション、または完全結合レイヤーを示す。例えば、図 13B において、「 $\times 4$ 」は、4 つの連続する畳み込みのスタックを示し、 $28 \times 28 \times 256$  マップの生成は、デコンボリューションによるものである。

図 13A において、「res5」は、ResNet の第 5 ステージを示す。これは、簡単にするために、最初の畳み込みがストライド 1 の  $7 \times 7$  RoI で動作するように変更されている。FPN の場合、バックボーンにはすでに res5 が含まれているため、より少ないフィルターを使用するより効率的なヘッドが可能となる。示されている例では、ResNet-C4 を使用して生成されたマスクは  $14 \times 14 \times 80$  であり、FPN を使用して生成されたマスクは  $28 \times 28 \times 80$  である。

### 3. クレーム

794 特許のクレーム 1 は以下の通りである。

1. コンピュータシステムにおける方法において、  
 トレーニング画像にアクセスし、  
 第 1 のニューラルネットワークを使用してトレーニング画像の特徴マップを生成し、  
 特徴マップにおける関心領域を特定し、

サンプリング領域によって定義されたサンプリング位置に基づいて、関心領域の領域特徴マップを生成し、前記サンプリング領域と関心領域は、特徴マップ内の同じ領域に整列しており、

第2のニューラルネットワークを使用して領域特徴マップを処理することにより、対象領域に関連付けられたインスタンスセグメンテーションマスクを生成し、インスタンスセグメンテーションマスクを使用して第2のニューラルネットワークをトレーニングし、

トレーニングされた第2のニューラルネットワークは、画像に描かれたオブジェクトインスタンスのインスタンスセグメンテーションマスクを生成するように構成される。

#### 4. MASK R-CNN に関する論文

本特許に関連する論文<sup>1</sup>「MASK R-CNN」が Kaiming He 氏らにより発表されている。

下記図は、COCO テストセットを用いた MASK R-CNN の結果を示している。各オブジェクトがピクセル単位での確に分類されていることが理解できる。



	backbone	AP	AP <sub>50</sub>	AP <sub>75</sub>	AP <sub>S</sub>	AP <sub>M</sub>	AP <sub>L</sub>
MNC [10]	ResNet-101-C4	24.6	44.3	24.8	4.7	25.9	43.6
FCIS [26] +OHEM	ResNet-101-C5-dilated	29.2	49.5	-	7.1	31.3	50.0
FCIS+++ [26] +OHEM	ResNet-101-C5-dilated	33.6	54.5	-	-	-	-
<b>Mask R-CNN</b>	ResNet-101-C4	33.1	54.9	34.8	12.1	35.6	51.1
<b>Mask R-CNN</b>	ResNet-101-FPN	35.7	58.0	37.8	15.5	38.1	52.4
<b>Mask R-CNN</b>	ResNeXt-101-FPN	<b>37.1</b>	<b>60.0</b>	<b>39.4</b>	<b>16.9</b>	<b>39.9</b>	<b>53.5</b>

上記テーブルは、COCO テストにおける、インスタンスセグメンテーションマスクの AP (Average Precision) を示している。MNC と FCIS は、それぞれ COCO の 2015

<sup>1</sup> Kaiming He Georgia Gkioxari Piotr Doll'ar Ross Girshick, "MASK R-CNN", arXiv:1703.06870v3 [cs.CV] 24 Jan 2018

と 2016 のセグメンテーションチャレンジの勝者である。Mask R-CNN は、マルチスケールトレーニングテスト、水平フリップテスト、および OHEM を含むより複雑な FCIS+++よりも優れた結果を達成している。

以上

#### 著者紹介

河野英仁

河野特許事務所、所長弁理士。立命館大学情報システム学博士前期課程修了、米国フランクリンピアースローセンター知的財産権法修士修了、中国清華大学法学院知的財産夏季セミナー修了、MIT(マサチューセッツ工科大学)コンピュータ科学・AI 研究所 AI コース修了。

[AI 特許コンサルティング](#)、[医療 AI 特許コンサルティング](#)の他、米国・中国特許の権利化・侵害訴訟を専門としている。著書に「世界のソフトウェア特許(共著)」、「FinTech 特許入門」、「[AI/IoT 特許入門 2.0](#)」、「[ブロックチェーン 3.0\(共著\)](#)」がある。