

### 1.はじめに

歩行者検出は、画像内から歩行者の位置と大きさを推定する技術であり、自動運転において重要な機能となる。歩行者検出の代表的な手法である、Center and Scale Prediction(CSP)[1]は、Convolutional Neural Network(CNN)をベースとし、高精度に歩行者の中心点とスケールを推定する。一方で、自動運転では、歩行者の一部にオクルージョンが生じているケースでも正しく検出する必要がある。CSPはオクルージョンが発生すると、歩行者の確信度が低下し、未検出となる。そこで本研究では、CSPを用いて歩行者候補領域を検出した後、Compositional Nets[2]によるオクルージョンを考慮した最終判定を行う2段階の歩行者検出手法を提案する。

### 2.関連研究

本研究で用いる、CSPとCompositional Netsの構造について述べる。

#### 2.1 Center and Scale Prediction

CSPのネットワーク構造を図1に示す。CSPは歩行者の中心点と縦横のスケールをヒートマップで推定するkeypoint型の検出手法である。ヒートマップ内の極大値を歩行者の中心点とし、極大値が歩行者の確信度となる。CSPはオクルージョンが発生していても可視部分の歩行者の特徴から中心位置を特定し、検出領域の推定が可能である。しかしながら、オクルージョンが発生している歩行者に対する確信度は低く、検出できないケースもある。

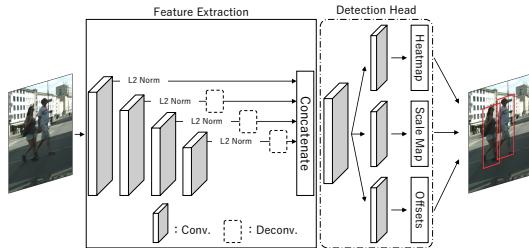


図1: Center and Scale Prediction

#### 2.2 Compositional Nets

Compositional Netsはオクルージョンが発生している度合をヒートマップとして出力する。また、物体らしさも合せて出力する。そのため、物体の判定とオクルージョン発生の度合を同時に捉えることができる。Compositional NetsはDCNNで畠み込み処理を行い、画像の特徴を抽出する。そして、抽出した特徴マップをvMF kernelsに入力し特徴ベクトルに確率を付与する。Occluder kernelsでは、その確率を基にオクルージョン部分とそれ以外の部分に分類したオクルージョンマップを出力する。Class mixturesでは、コンテキスト情報とオブジェクト情報を合わせた物体マップを出力する。オクルージョンマップと物体マップを統合し、最終的なクラススコアとオクルージョンマップを出力する。

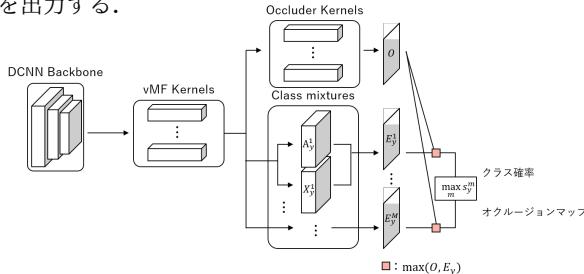


図2: Compositional Nets

### 3.提案手法

本研究では、CSPを用いて歩行者候補領域を検出した後、Compositional Netsによるオクルージョンを考慮した

最終判定を行う2段階の歩行者検出手法を提案する。提案手法の流れを図3に示す。1段階目では、入力画像をCSPに入力し、歩行者の確信度マップを求める。このとき、確信度のマップに対して、閾値を低く設定して多くの候補領域を出力する。次の、2段階目では、歩行者候補領域を1つずつCompositional Netsに入力し、候補領域のクラス確率とオクルージョンマップを出力する。クラス確率を基に候補領域を歩行者か否かを判定する。

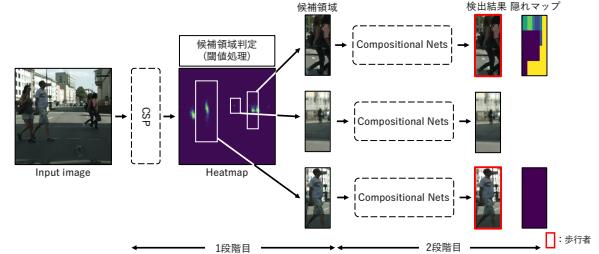


図3: 提案手法の2段階判定

### 4.評価実験

本実験では、従来のCSPと提案手法を比較し、その有効性を評価する。

#### 4.1 実験概要

学習、評価にはCityPersonsデータセットを用いる。評価指標にはMiss Rate(未検出率)を用いる。評価は、Bare, Partial, Heavy, Reasonableのオクルージョン度合ごとにMiss Rateを算出し、従来手法(CSP単体)との比較を行う。

#### 4.2 実験結果

従来手法と提案手法を比較した結果を表1に示す。表1より、提案手法を従来手法と比較して、Heavyにおいてしきい値0.25とした時、Miss Rateの精度を6.76ポイント向上させることができた。提案手法の検出結果とオクルージョンマップを図4に示す。図4から提案手法は、歩行者のオクルージョン領域の推定に成功し、正しい判定結果を出力することができ、検出できていることが分かる。

表1: 実験結果 (Miss Rate)

	閾値	Reasonable	Bare	Partial	Heavy
従来法	0.5	20.28	15.49	21.03	60.90
	0.1	17.93	13.30	18.03	55.55
提案手法	0.25	<b>16.82</b>	<b>12.37</b>	<b>16.48</b>	<b>54.14</b>
	0.5	18.65	13.96	19.10	58.65

ジョーンマップを図4に示す。図4から提案手法は、歩行者のオクルージョン領域の推定に成功し、正しい判定結果を出力することができ、検出できていることが分かる。



図4: 定性的結果

### 5.おわりに

本研究では評価実験からオクルージョンの考慮をした2段階の判断により精度が向上することを確認した。今後は、本提案手法の更なる高精度化を目指す。

### 参考文献

- [1] W. Liu *et al.*, “Center and Scale Prediction:A Box -free Approach for Pedestrian and Face Detection”, CVPR 2019.
- [2] A. Kortylewsk, *et al.*, “Compositional Convolutional Neural Networks:A Deep Architecture with Innate Robustness to Partial Occlusion”, CVPR 2020.