

傘による遮蔽に頑健なパーツ参照型歩行者検出

○ 新保祐人[†], 出口大輔[‡], 川西康友[†], 井手一郎[†], 村瀬洋[†]

○ Yuto SHIMBO[†], Daisuke DEGUCHI[‡], Yasutomo KAWANISHI[†],
Ichiro IDE[†], Hiroshi MURASE[†]

[†]: 名古屋大学 大学院情報科学研究科, shimboy@murase.m.is.nagoya-u.ac.jp,
{kawanishi, ide, murase}@is.nagoya-u.ac.jp

[‡]: 名古屋大学 情報連携統括本部, ddeguchi@nagoya-u.jp

<要約> 近年, 自動運転の実現に向けて自動車の周囲環境認識に関する研究が盛んに行なわれている。その中でも, 見落としが重大な事故につながる歩行者の検出は重要な課題である。しかし, 傘を差した歩行者は, 傘によって頭部が遮蔽されるため検出が困難な対象として知られている。そこで本発表では, 傘差し歩行者検出の高精度化を目的とし, パーツ参照型歩行者検出手法である Parts Driven Faster RCNN を提案する。具体的には, Faster RCNN の候補領域抽出部において, 傘と歩行者の関係を学習して傘の候補領域から歩行者の候補領域を推定することにより, 頭部が遮蔽された歩行者でも候補領域として抽出できるようにする。提案手法の有効性を確認するため, 雨天時に撮影した車載カメラ画像系列を用いて評価実験を行ない, 傘差し歩行者に対して検出精度が向上することを確認した。

<キーワード> 歩行者検出, 傘差し歩行者, ITS, パーツ参照型物体検出, Faster RCNN

1. はじめに

近年, 自動運転の実用化を目指して, 自動車の周囲環境認識に関する研究が盛んにおこなわれている。その中でも, 見落としが重大な事故につながる歩行者の検出は重要な課題である。

ここ数年のCNNに関する研究の進展に伴い, 画像中の物体認識の精度は劇的に向上している。なかでも, 画像から物体候補領域を抽出した後, その領域が対象物体であるか否かを識別するRCNN [1] と呼ばれる手法は, 歩行者検出においても高い精度を達成している。RCNN系の手法の中でも, Faster RCNN [2] は高精度かつ高速な検出が可能であり, 歩行者検出においても利用されている。

物体検出の検出精度を低下させる要因である遮蔽に対処する手法も数多く提案されており [3, 4], 特定状況下において高頻度に発生する遮蔽を考慮することが物体検出の精度向上につながる事が示されている。このことから, 検出対象の付随物体を明示的に検出モデルに組み込み, それらの関係を学習して検出に利用することが有効と考え

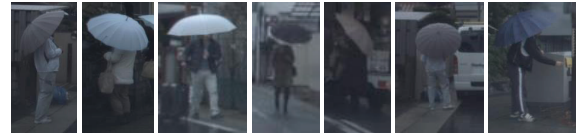


図 1 検出が困難な傘差し歩行者の例

られる。

雨天時における歩行者検出を考えた場合, 図 1 に示すように歩行者の頭部から胸部が傘によって遮蔽される状況が高頻度で発生する。歩行者検出において, 頭部は重要な形状であるため, 傘差し歩行者は検出が困難な対象とされている [5]。このような歩行者を検出しようとした場合, 従来のFaster RCNNによる歩行者検出手法では, 傘によって頭部が隠れることにより, そもそも歩行者候補領域抽出の段階で見落とすという問題が発生する。この問題を解決することにより, 傘差し歩行者検出の高精度化が望める。

ここで, 前述の付随物体を物体検出に利用するという考え方は, 雨天時の傘差し歩行者の検出に

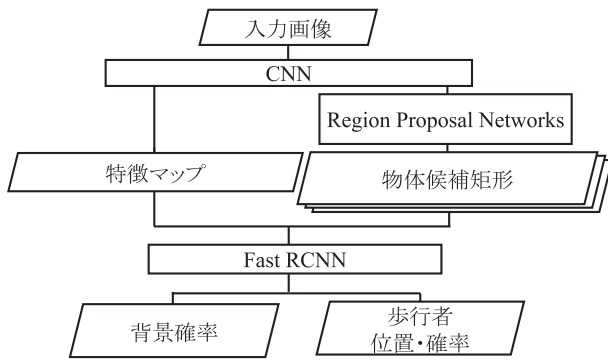


図 2 Faster RCNN の処理手順

においても利用できると思われる。そこで本発表では、車載カメラ画像を用いた傘差し歩行者検出の高精度化を目的として、パーツ参照型歩行者検出手法 Parts Driven Faster RCNN (PD-FRCNN) を提案する。歩行者を検出対象とした検出モデルに傘を付随物体として組み込み、それらの関係性を学習することができれば歩行者の検出精度の向上が期待できる。具体的には、Faster RCNN において物体候補領域抽出を行なう Region Proposal Networks (RPN) を拡張し、付随物体の位置から検出対象の位置を推定する Parts Driven Region Proposal Networks (PD-RPN) を導入することで傘差し歩行者検出の精度向上を図る。

2. Parts Driven Faster RCNN (PD-FRCNN)

2.1 Faster RCNN

Faster RCNN [2] は、一般物体認識のベンチマーク [6] で高い精度を記録したマルチクラス検出器である。歩行者クラスと背景クラスの 2 クラス検出器としても利用可能であり、歩行者検出においても広く応用されている。Faster RCNN の検出処理の流れを図 2 に示す。Faster RCNN は、Convolutional Neural Network (CNN) [7] による特徴抽出部、Region Proposal Networks (RPN) による物体候補領域抽出部、Fast RCNN [8] による識別部の 3 つの要素で構成される。以降では、Faster RCNN の候補領域抽出部である RPN を説明した後、提案手法である Parts Driven Region Proposal Networks (PD-RPN) について説明する。

2.2 Region Proposal Networks (RPN)

RPN は、CNN 特徴マップを入力として受け取り、物体候補となる領域の位置と物体らしさを入力する。CNN 特徴マップに対して 3×3 の枠を走査し、その枠内の局所特徴 \mathbf{x} をさらに畳み込むことによって走査枠毎に物体らしさを推定する。このとき、RPN は 1 つの走査枠からアンカと呼ばれる複数の領域に対する物体らしさを判定する。すなわち、RPN は CNN 特徴マップ上の 1 つの局所特徴 \mathbf{x} から、各アンカに設定された枠 $\mathbf{R}^i(\mathbf{x})$ ($i = 1, \dots, k$) に対するクラス確率 $\mathbf{p}^i(\mathbf{x}) = (p_0^i(\mathbf{x}), p_1^i(\mathbf{x}))$ および矩形回帰 $\mathbf{t}^i(\mathbf{x}) = (t_x^i(\mathbf{x}), t_y^i(\mathbf{x}), t_w^i(\mathbf{x}), t_h^i(\mathbf{x}))$ を出力する。ここで、 $p_0^i(\mathbf{x})$ は背景確率、 $p_1^i(\mathbf{x})$ は前景確率であり、 $\mathbf{t}^i(\mathbf{x})$ は $\mathbf{R}^i(\mathbf{x})$ から推定された物体候補領域への相対位置および相対サイズを表す。

RPN は数層の畳み込み層によって構成されているため、そのネットワークパラメータは誤差逆伝播法による学習が可能である。学習時には、局所特徴 \mathbf{x} に対して各アンカ i に設定されている枠 $\mathbf{R}^i(\mathbf{x})$ と検出対象の学習サンプルの重なり率を求めて、 $\mathbf{R}^i(\mathbf{x})$ へラベル $y^i(\mathbf{x}) \in \{0, 1\}$ を付与する。 $\mathbf{R}^i(\mathbf{x})$ と重なり率が 0.7 以上の検出対象クラスの学習サンプルがあればポジティブ ($y^i = 1$) とし、重なり率が 0.3 未満の学習サンプルがあればネガティブ ($y^i = 0$) とする。また、検出対象クラスに属する学習サンプルに対して、重なり率が 0.7 以上となる $\mathbf{R}^i(\mathbf{x})$ が存在しなければ最も重なり率が高い $\mathbf{R}^i(\mathbf{x})$ のラベルをポジティブとする。また、 $\hat{\mathbf{t}}^i$ は事前に学習データとして与えられた $\mathbf{R}^i(\mathbf{x})$ からポジティブサンプルへの相対位置および相対サイズであり、矩形回帰の目標値を表す。以降では、簡略化のため \mathbf{p}^i や \mathbf{t}^i のように引数 (\mathbf{x}) を省略して表記する。

学習は、以下の損失関数 $L(\mathbf{p}^i, y^i, \mathbf{t}^i, \hat{\mathbf{t}}^i)$ を最小化するように確率的勾配降下法と誤差逆伝播法を用いて行なう。

$$L(\mathbf{p}^i, y^i, \mathbf{t}^i, \hat{\mathbf{t}}^i) = \frac{1}{N_\alpha} \sum_i L_\alpha(\mathbf{p}^i, y^i) + \lambda \frac{1}{N_\beta} \sum_i y^i L_\beta(\mathbf{t}^i, \hat{\mathbf{t}}^i), \quad (1)$$

$$L_\alpha(\mathbf{p}^i, c) = -\log p_c^i \quad (2)$$

$$L_\beta(\mathbf{t}^i, \hat{\mathbf{t}}^i) = \sum_{j \in \{x, y, w, h\}} S_{L_1}(t_j^i - \hat{t}_j^i), \quad (3)$$

$$S_{L_1}(x) = \begin{cases} 0.5x^2 & \text{if } |x| < 1 \\ |x| - 0.5 & \text{otherwise,} \end{cases} \quad (4)$$

ただし、 N_α は学習におけるミニバッチサイズ、 N_β は CNN 特徴マップに対する走査枠の個数である。 $L_\alpha(\mathbf{p}^i, c)$ は正解ラベル $c \in \{0, 1\}$ のクラス確率 p_c^i に対する対数損失であり、クラス分類に関する損失である。また、 $L_\beta(\mathbf{t}^i, \hat{\mathbf{t}}^i)$ はアンカ毎の出力 \mathbf{t}^i と $\hat{\mathbf{t}}^i$ の差に対する矩形回帰損失である。 λ は $L_\alpha(\mathbf{p}^i, c)$ と $L_\beta(\mathbf{t}^i, \hat{\mathbf{t}}^i)$ の重みを調節するパラメータである。これらの損失関数を最小化することによって、RPN は入力された CNN 特徴マップ上の全ての $\mathbf{R}^i(\mathbf{x})$ に対する物体確率と矩形回帰を出力するようになる。

2.3 Parts Driven Region Proposal Networks (PD-PRN)

RPN による物体候補抽出では、検出対象の一部に遮蔽が発生すると見えの情報が欠落してしまうため、検出対象の前景確率 p_1^i が低くなったり、矩形回帰出力 \mathbf{t}^i が誤った値になるという問題がある。しかし、検出対象に付随する物体の位置や姿勢などから検出対象の候補領域を推定できれば、検出対象自体の遮蔽の影響に関係なく候補領域を抽出することが可能である。そこで、検出対象自体の見えの情報に反応して矩形回帰を行なう「直接的な物体候補領域抽出」と、付随物体の位置から検出対象の位置を推定する「間接的な物体候補領域抽出」を同時に行なうように RPN を改良する。このように検出対象に付随するパーツを参照しながら物体候補領域抽出を行なう手法を Parts Driven Region Proposal Networks (PD-RPN) と呼ぶ。PD-RPN の概念図を図 3 に示す。

従来の RPN の学習では、学習サンプルの真のクラスを考慮せず、全ての学習サンプルを前景クラス ($y^i = 1$) か背景クラス ($y^i = 0$) として扱っていた。しかし、提案する PD-RPN では直接的な物体候補領域抽出と間接的な物体候補領域抽出を同時に行なうため、学習サンプルの真のクラスに応じて、矩形回帰損失の値を変化させる必要がある。そこで、サンプルに付与するラベルを $y^i \in \{0, 1, 2\}$ に拡張する。ここで、背景クラスを $y^i = 0$ 、検出対象のクラスを $y^i = 1$ とし、検出対象の物体領

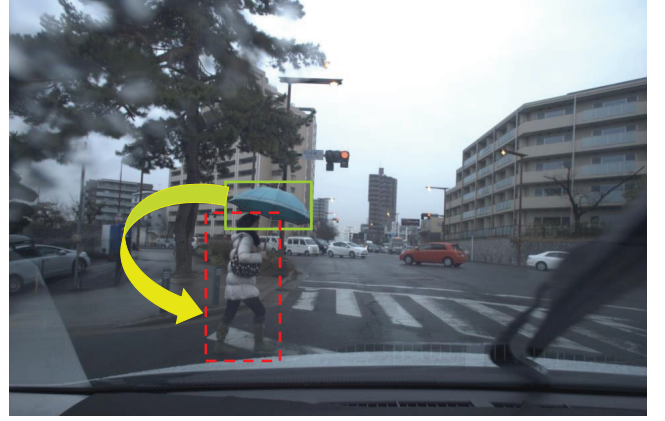


図 3 Parts Driven Region Proposal Networks によるパーツ参照型物体候補領域抽出。直接的な歩行者候補領域の抽出と、傘候補領域（緑実線）からの歩行者候補領域（赤破線）の推定を同時に行なう。

域を推定するために参照する付随物体のクラスを $y^i = 2$ とする。これにより、学習において検出対象クラスと付随物体クラスを区別できるようにする。また、学習データセットにおいて付随物体のサンプルは同一画像内に存在する検出対象のサンプルと 1:1 で対応付けられているものとする。ここで、付随物体サンプルへの矩形回帰目標 $\hat{\mathbf{t}}^i$ を対応付けられた検出対象サンプルへの矩形回帰目標へ変換する関数を $f(\hat{\mathbf{t}}^i)$ とする。以上の条件の下、RPN の損失関数 $L(\mathbf{p}^i, y^i, \mathbf{t}^i, \hat{\mathbf{t}}^i)$ を学習サンプルのクラス y^i に応じて値が変化するように以下の損失関数 $L'(\mathbf{p}^i, y^i, \mathbf{t}^i, \hat{\mathbf{t}}^i)$ へ拡張する。

$$L'(\mathbf{p}^i, y^i, \mathbf{t}^i, \hat{\mathbf{t}}^i) = \frac{1}{N_\alpha} \sum_i L'_\alpha(\mathbf{p}^i, y^i) + \lambda \frac{1}{N_\beta} \sum_i [y^i \geq 1] L'_\beta(\mathbf{t}^i, \hat{\mathbf{t}}^i, y^i), \quad (5)$$

$$L'_\alpha(\mathbf{p}^i, c) = \begin{cases} -\log p_0^i & \text{if } c = 0 \\ -\log p_1^i & \text{otherwise} \end{cases} \quad (6)$$

$$L'_\beta(\mathbf{t}^i, \hat{\mathbf{t}}^i, c) = \begin{cases} L_\beta(\mathbf{t}^i, \hat{\mathbf{t}}^i) & \text{if } c \leq 1 \\ L_\beta(\mathbf{t}^i, f(\hat{\mathbf{t}}^i)) & \text{otherwise} \end{cases} \quad (7)$$

ここで、 $[y^i \geq 1]$ は $y^i \geq 1$ のとき 1 となり、 $y^i = 0$ のときは 0 となる。すなわち、アンカ領域 \mathbf{R}^i がク

ラス $y^i = 2$ の付随物体サンプルと重なったとき、矩形回帰損失 $L'_\beta(t^i, \hat{t}^i, c)$ は出力値 t^i と \hat{t}^i の差に対してではなく、 t^i と $f(\hat{t}^i)$ の差に対して損失を与える。この損失関数を用いて学習することにより、RPN は付随物体から抽出された CNN 特徴に反応して、その付随物体と対応付けられた物体の候補領域を推定するようになる。なお、この領域の前景確率には付随物体の前景確率 p^i がそのまま用いられる。これにより、検出対象の遮蔽を要因とする前景確率の低下を抑制する。また、式 (7) 内の条件分岐により、検出対象物体に対しては従来通りその物体の候補領域が出力される。

Parts Driven Faster RCNN (PD-FRCNN) では、この PD-RPN により出力される物体候補矩形のうち、前景確率の上位 N 個を Fast RCNN の入力として与え、詳細な識別を行なう。

3. 傘差し歩行者検出の実装方法

3.1 学習

傘差し歩行者検出に対する PD-FRCNN の学習について説明する。PD-FRCNN の学習では、Faster RCNN における RPN を 2.3 節で提案した PD-RPN に置き換えて学習を行なう。学習には以下の 4-Step Alternating Training アルゴリズム [2] を使用する。

STEP 1 : PD-RPN の学習

識別を行なう Fast RCNN の学習には、画像からの特徴抽出を行なう CNN と物体候補領域が必要である。そこで、初めに物体候補領域を抽出する PD-RPN を学習する。傘差し歩行者検出においては、歩行者を検出対象 ($y^i = 1$) とし、傘を付随物体 ($y^i = 2$) として扱う。式 (5) の損失関数を用いて、確率的勾配降下法と誤差逆伝播法によって特徴抽出部の CNN と PD-RPN のネットワークパラメータを学習する。学習においては各画像における PD-RPN の出力のうち上位 2,000 個を Fast RCNN の学習に使用する。

STEP 2 : Fast RCNN の学習

STEP 1 で学習された PD-RPN により抽出された候補領域のうち、各画像毎に前景確率の上位 2,000 個を用いて Fast RCNN を学習す

る。Fast RCNN は、歩行者 ($y = 1$) と背景 ($y = 0$) の 2 クラス分類器として学習される。学習において入力される物体候補領域には対応付けられた学習サンプルの真のクラスラベル $y \in \{0, 1\}$ と矩形回帰の目標値 t^y が付与されているものとする。Fast RCNN は、入力された CNN 特徴マップ上に物体候補領域を射影し、その領域に対するクラス確率 $\mathbf{p} = (p_0, p_1)$ および物体候補領域からの矩形回帰値 t^y を出力とする。損失関数には以下のマルチタスク損失が用いられる。

$$L(\mathbf{p}, y, t^y, \hat{t}^y) = L_\alpha(\mathbf{p}, y) + \lambda[y \geq 1]L_\beta(t^y, \hat{t}^y), \quad (8)$$

ただし、 $L_\alpha(\mathbf{p}, y) = -\log p_y$ は正解クラス y に対する対数損失であり、クラス分類に対する損失である。また、 L_β は正解クラス y に対する矩形回帰の目標値 t^y および出力値 t^y に対して定義される。すなわち、背景クラスの場合は矩形回帰の目標が存在しないため、 L_β は無視される。この損失を用いて、確率的勾配降下法と誤差逆伝播法によって特徴抽出部の CNN と Fast RCNN のネットワークパラメータを学習する。

STEP 3 : PD-RPN の再学習

STEP 2 で学習された特徴抽出部の CNN のパラメータを固定して、STEP 1 と同様に PD-RPN のネットワークパラメータのみを再学習する。

STEP 4 : Fast RCNN の再学習

STEP 2 で学習された特徴抽出部の CNN のパラメータを固定し、STEP 3 で学習された PD-RPN による物体候補領域を用いて Fast RCNN のネットワークパラメータのみを再学習する。

この繰り返し処理によって、PD-RPN と Fast RCNN において特徴抽出を行なう CNN のネットワークパラメータを共有して学習することができる。以上が提案する PD-FRCNN の学習である。以降の検出処理では、この学習方法によって学習されたネットワークにテスト画像を入力し、画像内の傘差し歩行者を検出する。

3.2 検出処理

Faster RCNN の検出処理は、前節で述べた学習によって得られたネットワークにテスト画像を入力することで行なわれる。初めに、入力画像は一定サイズに拡張されてから CNN に入力され、数層の畳み込み層を通ることによって CNN 特徴マップが抽出される。次に、特徴マップが PD-RPN に入力され、物体候補領域が抽出される。このとき、抽出される物体候補領域には、歩行者を直接抽出した領域と傘の位置から歩行者の位置を間接的に推定された領域も含まれる。ここで、遮蔽の影響によって歩行者候補領域の物体らしさが低下している可能性があるため、傘候補領域から推定された歩行者候補領域の物体らしさは元の傘候補領域の物体らしさが用いられる。PD-RPN によって抽出された物体候補領域は前景確率の値で整列され、上位 300 個が Fast RCNN による識別の対象となる。そして、CNN 特徴マップと PD-RPN で抽出された物体候補領域が Fast RCNN に入力され、各物体候補領域の歩行者確率 $\mathbf{p} = (p_0, p_1)$ と矩形回帰 \mathbf{t}_i が出力される。ここで、 p_0 は背景確率、 p_1 は歩行者確率である。最後に、得られる検出結果には歩行者の周辺でいくつかの検出枠が重なって出力されることが多い。そこで、Non Maximum Suppression によって重なり率が 0.7 以上の検出枠のうち歩行者確率の小さな枠を除去する。以上が傘差し歩行者に対する PD-FRCNN の検出処理である。

4. 評価実験

4.1 パラメータ

Faster RCNN の実装には、Girschick が提供している Python のプログラム [10] を用いた。特徴抽出を行なう CNN の構造には ZF モデル [11] を用いた。学習における各種パラメータや RPN および Fast RCNN の検出パラメータには、実装 [10] の初期設定を使用した。その設定における RPN のアンカ数は 9 個である。各アンカ枠の大きさを表 1 に示す。

4.2 データセット

本実験では、雨天時車載カメラ画像系列を 2 本用意し、1 つの画像系列を学習用データセット、も

う一方の画像系列を評価用データセットとした。各データセットの詳細を表 2 に示す。学習および評価のために、各画像中の歩行者および傘の外接矩形を手によって付与した。また、学習サンプルの拡充のために、水平方向に反転させたポジティブサンプルも学習に用いた。

4.3 評価方法

各手法で出力された矩形 \mathbf{b}_d は、正解矩形 \mathbf{b}_g に対して以下の条件式を満たすときに正解として扱った。

$$\frac{|\mathbf{b}_d \cap \mathbf{b}_g|}{|\mathbf{b}_d \cup \mathbf{b}_g|} \geq 0.5, \quad (9)$$

ここで、 $|\cdot|$ は領域のピクセル数である。

本実験の評価指標には Free-response Receiver Operating Characteristic (FROC) 曲線と F 値を用いた。FROC 曲線は縦軸に検出率、横軸に画像 1 枚当たりの誤検出数 (False Positives Per Image) をとる曲線であり、曲線がプロット領域の左上に位置するほど精度が良いことを表す指標である。FROC 曲線は、評価に用いる検出結果のスコアのしきい値を変化させることによって描画した。F 値は再現率と適合率を総合的に評価する指標である。これら 2 つの指標を用いて、従来の Faster RCNN と提案する PD-FRCNN の検出精度を比較した。

4.4 実験結果

検出結果である FROC 曲線を図 4 に示す。本論文で提案した PD-FRCNN の曲線が従来手法である Faster RCNN の曲線よりも左上に位置していることが分かる。すなわち、提案手法は従来手法と比べて検出精度が向上したといえる。しかし、曲線右側における最終的な再現率には大きな影響はなかった。

また、各手法の最高 F 値およびその時の再現率と適合率を表 3 に示す。この表より、提案手法は再現率をほとんど低下させることなく、適合率が 2.7% だけ上昇していることが確認できた。また、検出結果を評価用画像に描写した例を図 5 に示す。それぞれの手法で最高 F 値をとるときの検出結果を赤枠で表示した。

表 1 アンカの大きさ

アンカ番号	1	2	3	4	5	6	7	8	9
サイズ (pixels)	128 ²	128 ²	128 ²	256 ²	256 ²	256 ²	512 ²	512 ²	512 ²
縦横比 (w,h)	(1,1)	(1,2)	(2,1)	(1,1)	(1,2)	(2,1)	(1,1)	(1,2)	(2,1)

表 2 データセットの内容

データセット	画像の枚数	画像の大きさ	歩行者の総数	傘差し歩行者の総数
学習用データセット	1,673	960 × 720	1,857	1,763 (95%)
評価用データセット	1,669	960 × 720	1,397	1,276 (91%)

表 3 各手法の性能比較

手法	再現率	適合率	F 値
提案手法 : PD-FRCNN	0.862	0.704	0.775
従来手法 : Faster RCNN	0.866	0.677	0.760

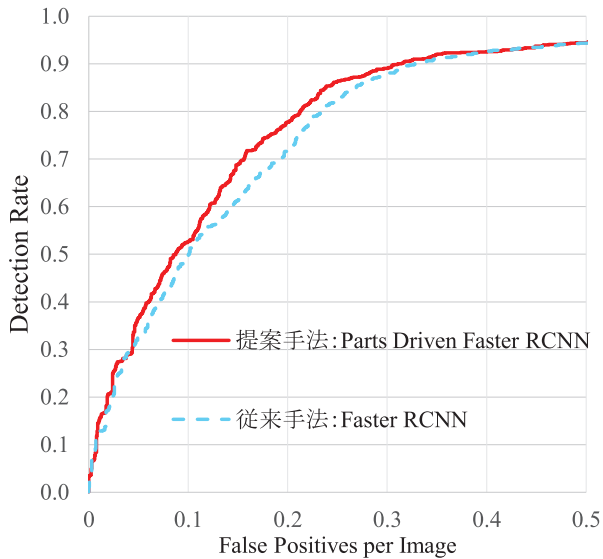


図 4 検出精度の比較

4.5 考察

図 4 に示した通り，最終的な再現率に変化はなく，提案手法は従来手法と比べて適合率が上昇するという結果が得られた．その理由として，PD-RPN により抽出される歩行者候補領域の質が改善

したことが考えられる．各手法が最高 F 値をとるときの検出スコアのしきい値は，提案手法で 0.762，従来手法で 0.689 であった．最高 F 値をとるときの再現率は各手法共に約 86% であったため，提案手法では正検出に対する検出スコアが上昇したことが分かる．これは傘の情報を利用して候補矩形抽出を行なうことによって，候補領域の質が改善されたためだと考えられる．実際に提案手法と従来手法の結果を比較した際，図 6 のように検出枠が改善されている例が数多く見られた．

また，どちらの手法でも検出できなかった歩行者の例を図 7 に示す．この図のように，遠方の小さな歩行者や近方の大きな歩行者の未検出が多く見られた．このような歩行者を検出するには，RPN におけるアンカの数を増やすことにより，多様な大きさの物体候補領域を抽出できるようにすることが有効だと考えられる．また，識別部の Fast RCNN を 2 つ用意し，入力される物体候補領域の高さに応じて各 Fast RCNN の出力するクラス確率を重みづけする Scale-aware Fast RCNN [12] などの手法を導入することも有効だと考えられる．

5. むすび

本発表では，付随物体の情報を積極的に利用し歩行者を検出するパーツ参照型歩行者検出手法 Parts



(a) 提案手法 (b) 従来手法

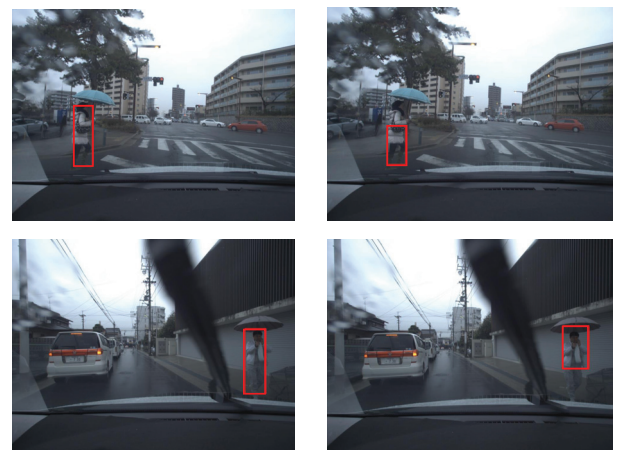
図 5 検出結果例

Driven Faster RCNN (PD-FRCNN) を提案した。具体的には、Faster RCNN において物体候補領域抽出を行なう RPN を、付随物体から検出対象の位置を推定する PD-RPN に拡張することにより、検出対象の遮蔽に頑健な物体候補抽出を可能とした。雨天時車載カメラ画像を用いた評価実験によって、提案する PD-FRCNN は従来の Faster RCNN と比べて傘差し歩行者の検出精度が向上することを確認した。

今後は、物体候補領域抽出部だけでなく識別部に対してもパーツ参照の概念を導入した手法を考案する予定である。

謝辞

本研究の一部は、科学研究費補助金による。



(a) 提案手法 (b) 従来手法

図 6 検出枠の改善例



(a) 遠方の歩行者 (b) 近方の歩行者

図 7 検出できなかった歩行者の例

参考文献

- [1] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," In Proc. 2014 IEEE Conf. on Computer Vision and Pattern Recognition, pp.580–587, June 2014.
- [2] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," In Proc. 29th Annual Conf. on Neural Information Processing Systems, Advances in Neural Information Processing Systems, vol.28, pp. 91–99, Curran Associates, Inc., Dec. 2015.
- [3] B. Pepik, M. Stark, P. Gehler, and B. Schiele, "Occlusion patterns for object class detection," In Proc. 2013 IEEE Conf. on Computer Vision and Pattern Recognition, pp.3286–3293, June 2013.
- [4] S. Tang, M. Andriluka, and B. Schiele, "Detection and tracking of occluded peo-

- ple,” *Int. J. of Computer Vision*, vol.110, no.1, pp.58–69, Oct. 2014.
- [5] 吉田 英史, 出口 大輔, 井手 一郎, 村瀬 洋, 町田 貴史, 小島 祥子, “生成型学習法を用いた傘をさした歩行者の高精度な検出に関する検討,” 電子情報通信学会技術研究報告, PRMU2012-50, Sept. 2012.
- [6] M. Everingham, L. V. Gool, C. K. Williams, J. Winn, and A. Zisserman, “The Pascal Visual Object Classes (VOC) challenge,” *Int. J. of Computer Vision*, vol.88, no.2, pp.303–338, June 2010.
- [7] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” In *Proc. 26th Annual Conf. on Neural Information Processing Systems, Advances in Neural Information Processing Systems*, vol.25, pp. 1097–1105, Curran Associates, Inc., Dec. 2012.
- [8] R. Girshick, “Fast R-CNN,” In *Proc. 2015 IEEE Int. Conf. on Computer Vision*. pp. 1440–1448, Dec. 2015.
- [9] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei, “ImageNet Large Scale Visual Recognition Challenge,” *Int. J. of Computer Vision*, vol.115, no. 3, pp. 211–252, Dec. 2015.
- [10] R. B. Girshick, “Faster R-CNN (Python implementation),” <<https://github.com/rbgirshick/py-faster-rcnn>>, Mar. 2016. (2017/1/6 参照)
- [11] M. D. Zeiler and R. Fergus, “Visualizing and understanding convolutional networks,” In *Proc. 2014 European Conf. on Computer Vision*, pp. 818–833, Springer International Publishing, Sept. 2014.
- [12] J. Li, X. Liang, S. Shen, T. Xu, and S. Yan, “Scale-aware fast R-CNN for pedestrian detection,” *arXiv preprint arXiv:1510.08160*, Oct. 2015.