

深層学習による車両検出ツールを用いた tricycleの分布把握に関する研究

荒井 勇輝¹・屋井 鉄雄²

¹非会員 東日本電信電話株式会社
(〒163-8019 東京都新宿区西新宿三丁目19番2号)
E-mail: arai.y.0925@gmail.com

²正会員 東京工業大学 環境・社会理工学院 土木・環境工学系 都市・環境学コース
(〒226-8502 横浜市緑区長津田町4259)
E-mail: tyai@enveng.titech.ac.jp

東南アジア途上国ではLAMATと呼ばれる交通モードが、現地の人々の主たる交通手段や、就労先等としての役割を担っている。しかし、政府などによる管理の曖昧さや個人による運営が種である点より、運行台数や分布等の市場特性に関するデータについては存在せず、その実態については不明瞭な点が多く存在する。そこで、本研究ではフィリピンのLAMATのtricycleを対象にその分布を把握する為のツールの開発を行い、その分布を明らかにした。衛星写真にはGoogleEarthを用い、検出手法には近年画像分野で飛躍を遂げている深層学習の理論に基づく方式を適用したツールを構築し、tricycleの分布を明らかにした。

Key Words : LAMAT, Developing countries, Image recognition, Vehicle detection, Deep learning

1. 序論

近年、東南アジア途上国においてはモータリゼーションの進行に伴い、慢性的渋滞や大気汚染の深刻化等が問題となり、それに伴う経済的損失が甚大なものとなっている。これらの諸問題に対し、公共交通の導入等による問題の解消が図られるが、その計画の段階においては現地の交通状態を定量的に把握する必要があると考えられる。

東南アジアにて走行が確認される交通モードとして自家用車やバス等の公共交通の存在が挙げられる。またその他の交通モードとして、LAMAT¹⁾ (Locally Adapted, Modified, and Advances Transport) の存在が挙げられる。LAMATの定義は、自家用車と典型的な大量旅客輸送公共交通の間に位置する、“自転車タクシーから定員25名のミニバスまでわたる”交通モードである。LAMATの存在は現地住民の主たる交通手段や就労先等としての役割を担っており、その運行実態についても定量的な把握を行う必要がある。しかしながらLAMATには既定の路線形態を持たないものや、個人による運行形態による運営が成されているものが数多く存在し、分布等をはじめとした市場特性

については不明瞭な点が数多く存在する。

そこで本研究では交通計画策定の一助とすべく、LAMAT の都市内での分布を明らかにすることを目的としたツールの開発を行い、対象とした LAMAT の分布を明らかにする。方法論として、画像認識分野にて発展を遂げている深層学習による手法を適用する。具体的には、衛星写真上より検出対象とする物体を検出し、検出された物体に対し CNN (深層学習の一手法である畳み込みニューラルネットワーク) によるカテゴリ判別を行う、R-CNN²⁾ (Region with CNN)の方法論を基にした車両検出ツール(以下、ツール)を構築する。その後、構築ツールを用い対象地域内での LAMAT の分布を推定し、その市場特性を明らかにする。

2. 既往研究と本研究の位置づけ

LAMAT の運行形態に迫る研究として Wang³⁾ らは、メトロマニラにおけるジープニー利用者に対してアンケートを行い、利用者の選択要因をはじめとした意識に関する分析を行った。また、Joewono⁴⁾ らは、インドネシアの

ミニバスや NMT (non-motorized vehicles) を対象に、運賃等に着目の上、利用者視点でのサービス水準をアンケート調査より明らかにした。

これらの研究は、LAMAT の利用者の意識や社会的存在価値に関し、言及したものであるが、LAMAT そのものの情報に関して言及した研究は見当たらない。

衛星写真上より深層学習をはじめとした機械学習手法を用い、車両の検出を行う研究は幾つか存在する。車両検出の新たなアルゴリズムの開発を行った研究として、Ram⁵⁾らは、自動車の対称性に基づき検出を行うアルゴリズムの開発を行った。既存の手法を改良した上で車両の検出を行う研究として、Shenquan⁶⁾らは、車両の特徴量を新たにBING (Binarized Normed Gradients) より取得することで検出の精度を向上させた。また、Sincha⁷⁾らは5段階カスケード分類器による検出の検討を行った。これは従来のカスケード分類器を改良し、車両形状を基にした特徴量等を追加した上で検出を行った。

この様に機械学習を用い、車両検出を行う研究では、精度向上や、即時性の向上を目的とした研究が多く存在する。しかし、車両の色や形が独特なLAMATを対象に検出をしたものは見当たらない。本研究では検出結果を基に分布等の市場特性を明らかにすることを目的とした点で異なる。

3. R-CNNを基にしたツールの理論・概要

a) 構築ツールの概要

本研究での構築ツールはR-CNNの理論(以下②～④)にて適用)を基とする。分布取得までのフローを示す。

① 分布把握対象地域の衛星写真画像の収集

GoogleEarth上より分布把握地域の衛星写真画像を収集。本研究では目視にて物体識別が可能、縦1080pixel×横1920pixel(約70m×110m)1枚を衛星写真画像と定義する。

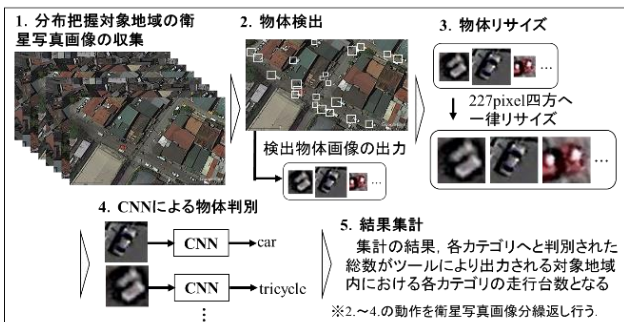


図-1 ツールによる分布取得の概要

② 物体検出

既存の物体検出手法を用い、衛星写真上より物体として認識されたものを全て検出物体画像として出力する。

③ 物体リサイズ

物体判別を行う際、CNNの規格上入力は227pixel四方とする為、各検出物体画像を一律でリサイズを行う。

④ 物体判別

CNNによりリサイズ済みの各検出物体画像全てに対して、カテゴリの判別を行う。

⑤ 結果の集計

判別結果の集計を行い、対象地域内における対象車両の走行台数、推定分布を出力。

b) カスケード分類器による物体検出

ツールにおける、物体検出ではカスケード分類器を採用した上で物体検出を行う。カスケード分類器による検出は以下のフローで行われ、物体検出の概念を図2へ示す。

- ① 衛星写真上へ検出窓を大きさを変えながら複数回滑らせ、学習した特徴量と重なる部分を持つ領域を物体候補領域として出力。
- ② 物体候補領域の内、指定枚数よりも多く候補領域が重なった部分を検出物体として正方形で出力。

物体検出に際し、数千枚の検出対象物(以下,正例)と非検出対象物(以下,負例)の画像より特徴の学習を予め行う。また物体検出実行時、検出される物体の最小、最大サイズのピクセル数(以下,minSize,maxSize)及び物体候補領域が何枚重なった時に物体領域とするか定める minNeighbors, 物体候補領域領域の矩形がどの程度の割合で重なると重なりと定義するか定める eps, 検出窓を minSize から maxSize まで変化させ候補領域を探索する際、変化させる検出窓の大きさ変化の幅を示す scaleFactor 値の設定をそれぞれ予め設定した上で検出へと使用する必要がある。

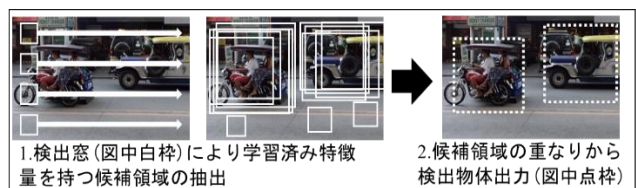


図-2 カスケード分類器による物体検出の概念

c) 物体リサイズ手法について

検出物体の画像の判別を次の段階で行うに当たり、判別部の規格上、227pixel 四方へとリサイズを行う必要がある。検出物体は minSize と maxSize で定めた値の範囲で出力がなされる。それらを一律で単純拡大した場合、リサイズ後の画像としてリサイズ前の pixel 数が小さい検出物体画像である程リサイズ後に空間周波数へ影響し、GoogleEarth の様な低解像度の場合には特に、その後の判別へと影響することが考えられる。そこで、本研究ではカスケード分類器にて定めた minSize と maxSize の指標を基に空間周波数成分を一定範囲に保つリサイズ手法を適用する。具体的には maxSize の大きさにて検出された検出物体画像に対しては一律で 227pixel 四方へ単純拡大によるリサイズを行い、その他のサイズの検出物体画像に対しては、maxSize の pixel との比で同様に単純拡大を行い、残りの空白部分へは、単純拡大後の検出物体画像の最外枠部の pixel の平均で補間し、空間周波数成分を一定範囲に保つ。図-3 では、minSize=30pixel 四方、maxSize=50pixel 四方と定め、40pixel 及び 50pixel 四方の検出物体画像をリサイズした例である。また、本リサイズでは単純拡大の手法として超解像を用いた。

d) CNNによる検出物体の判別

本研究ではニューラルネットワーク中の中間層で畳み込み、プーリング処理を交互に行い、より抽象的な情報を基に物体判別を行うCNNを用いた。実装に当たりディープラーニング用のライブラリであるCaffeを用いた。検出物体画像の判別前に、判別器に対し判別先の各カテゴリの特徴学習を行う。本研究では判別器の学習にて、構

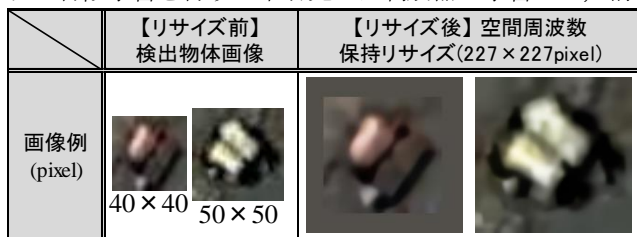


図-3 適用リサイズ手法によるリサイズ例

表-1 構築ツールの精度検証時の使用用語、指標名

No.	精度検証時使用用語名	概要	指標名	概要	式
i	tricycle実運行台数(真値)	衛星写真中で目視にて確認できたtricycleの台数	物体検出に関する指標	検出率	衛星写真真画像中で目視で確認することが出来たtricycleの内、検出物体として検出した割合
ii	検出物体画像中のtricycle数	全検出物体画像中で欠損が無く、目視にてtricycleとして確認できた枚数		物体判別に関する指標	accuracy
iii	検出物体中判別先tricycle数: 出力値(ツールによるtricycle走行台数)	全ての検出物体画像に対しツールにより物体判別を行いtricycleと判別された枚数	recall rate		検出物体中で目視にて確認できた全てのtricycle内でtricycleと判別出来た割合
iv	正判別tricycle台数	検出物体画像中でtricycleと目視で確認できたもの(ii)中でtricycleと判別された数			

築済みのネットワークに対し追加学習を行う転移学習(Fine-tuning)を用いる。転移学習とは、汎用的な教師データを基に学習済みネットワーク構造の出力層を判別先のカテゴリ数へ変更し、再度学習する手法を指す。ネットワークアーキテクチャをはじめから構築するよりも学習データ数が少なく、高精度な判別が可能である。本研究にて転移学習の適用先のネットワークアーキテクチャとして、Caffenet (以下、図-4) を使用する。

4. R-CNNを基にした車両検出ツールの構築

a) 検出対象LAMATについて

本研究では、フィリピン国内のLAMATであるtricycleを分布把握対象LAMATとした。tricycleは運営の大部分が個人により行われ、路肩の利用者を拾い目的地へ向かう。(以下、図-5)

b) 物体検出部、判別部の構築

本節では、ツールの物体検出部、物体判別部の構築の検討を行う。精度検証時の使用用語、指標を表-1 へ示す。

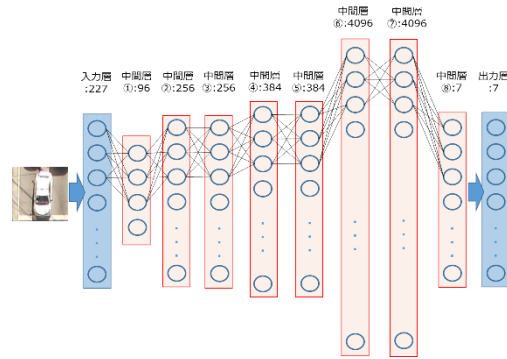


図-4 CaffeNet モデル構造



図-5 本研究対象の tricycle (左：地上撮影 右：衛星撮影)

また、構築した物体検出の精度及び物体判別の精度検証の為に、本節では精度検証用の衛星写真画像として、図-6へ示す衛星写真画像(目視にて確認した i. tricycle 実運行台数は18台)を対象に検出、判別を行う。

c) 物体検出部の構築

物体検出部の学習にて正例にはtricycleの画像を、負例には非検出対象物体である路面や、屋根模様等を含むデータを収集し学習を行う。以下の図-7へ例を示す。

検出時のパラメータを表-2へ示した値へ設定の上、図-6へ示した検証用の衛星写真上よりtricycleの検出を行った。ii. 検出物体中のtricycle数を目視にて数えた結果18台であり、検出率は18/18=1となった。

また、tricycleの他、68枚の路面や屋根模様等も検出がされた(図-8)。これら全86枚の検出物体に対し、物体判別を行いtricycleの抽出を行う。物体判別の段階においてtricycleと判別された総数が、ツールにより推計されるtricycle台数となる(表-1中iii値)。

d) 物体判別部の構築

物体検出部にて検出された検出物体画像を空間周波数成分を保つリサイズを行った後に判別を行う。判別先のカテゴリは検出される可能性を有するカテゴリで構成した。具体的にはtricycleの他、同程度のpixel数で示される



図-6 精度検証時の使用衛星写真画像

データ名	正例	負例
概要	GoogleEarth上のtricycle画像	tricycle以外(路面,屋根等)
例		
枚数	4000枚	4000枚

図-7 カスケード分類器の学習データセット

表-2 カスケード分類器の物体検出時パラメータ

変数名	minSize	maxSize	scaleFactor	minNeighbors	eps
値	30×30	50×50	1.01	2	0.3

bike, 車両欠損や路面、屋根模様等のnegの全3カテゴリとした(図-9)。

本研究では転移学習を用い、汎用性を示すテストデータへの判別精度と汎化誤差が収束した(図-10)各カテゴリ450枚、学習回数10000回、学習率0.0005にて学習した。

数百枚の学習で誤差が収束したのは転移学習を行った為である。次に、検出物体全86枚(内tricycle18枚)をリサイズし、構築した判別器で判別を行った。表-3へ結果を示す。

表-3へ示した結果より、ツールにより出力されるtricycleの台数値は、誤判別も含め19台となり、目視にて認識した真値である18台に近い数字を本構築ツールでは出力出来た。本研究ではここで構築した検出器、判別器をツール構築時に採用する。検出、判別の精度を向上させ、各評価値を常に1とすることで正確な分布を取得することが課題だが、negカテゴリ内での特徴の多様さや分布把握対象の解像度や自然環境による要因より検出漏れや誤判別の完全な解消は難しい。そこで、それらが発生することを踏まえ、台数の補正を行い真値の推定を行う。



図-8 検出物体画像例

表-3 各検出画像の物体判別の結果,判別に関する評価

		判別先カテゴリ				各評価値	
		tricycle	bike	neg	計	accuracy	14/19 ≒ 0.737
本来カテゴリ	tricycle	14	0	4	18	recall rate	14/18 ≒ 0.778
	tricycle以外	5	1	62	68	検出物体中判別先tricycle数	14+5 = 19(台)
	計	19	1	66	86		

カテゴリ	bike	tricycle	neg
画像例			
学習データ(枚)	450	450	450

図-9 物体判別部の学習データセット

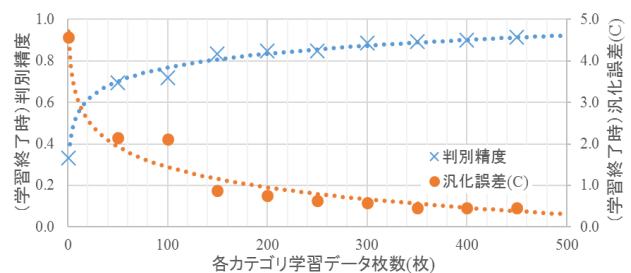


図-10 学習データ枚数別の判別精度, 汎化誤差

e) 出力台数補正に関する検討

出力台数の補正に際し、真値であるtricycle実運行台数とツールによるtricycleの出力台数である検出物体内判別先tricycle数の関係は各指標を用い以下の式(1)で示される。

$$i. \text{tricycle実運行台数} = \text{iii. 検出物体内判別先tricycle数} \times \frac{\text{accuracy}}{\text{recall rate} \times \text{検出率}} \quad (1)$$

本研究では、出力台数の検出物体内判別先tricycle数へ、係数を乗じ真値の推定を行う。式(1)より $\frac{\text{accuracy}}{\text{recall rate} \times \text{検出率}}$ 部を補正係数と定義し、複数の衛星写真画像より補正係数の傾向を分析する。本研究では45枚のtricycleが走行する衛星写真画像を対象に補正係数を算出しヒストグラムの作成と値による信頼区間を検討した(図-11、表-4)。

本研究にて構築したツールで出力されるtricycleの台数へ表-4の値を乗じ、対象地域内の運行台数推定を行う。

5. 構築ツールによる衛星写真画像内の分布推定

本章では、都市内でのtricycleの分布推定を行う。本研究では、マニラ都市圏のラスピニャス2km四方(縦32枚×横18枚の全576枚の衛星写真画像により構成)を対象に推定を行う。ツールの適用可能性検討の為、分布把握対象地域の一部(図-12点線内)の推定値と真値との比較を行う。

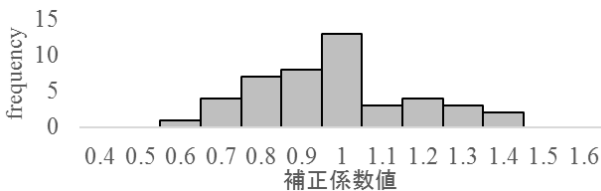


図-11 補正係数ヒストグラム

表-4 補正係数信頼区間

信頼区間	95%		99%	
	下限	上限	下限	上限
補正係数値	0.551	1.345	0.417	1.479

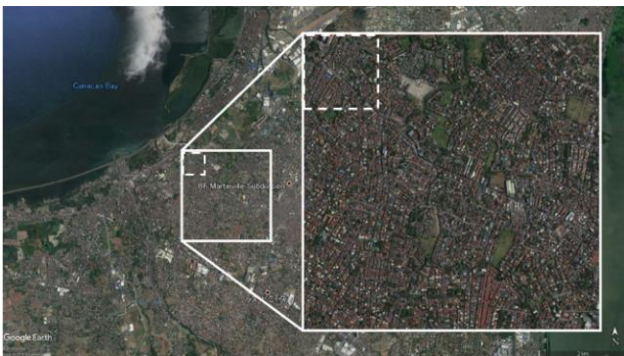


図-12 分布把握対象範囲

a) ツールの適用可能性検討

図-12中、点線内500m四方(縦8枚×横5枚の計40枚)では、目視で46台のtricycleが確認された。40枚全てよりツールにより台数出力と集計を行った結果、tricycle数は計66台となり、補正係数を乗じた推定結果を表-5へ示す。結果より、真値である46台を推定区間に収まる結果を導出出来、ツールとしての適用可能性が示された。

b) 広域を対象としたtricycle台数の推定

本節では広域を対象としたtricycleの台数推定を行った。対象地域の衛星写真全576枚よりツールを用い各衛星写真上のtricycleの台数を出力し、その結果を補正及び、集計した上で分布傾向を示した図を対象地域の衛星写真と共に図-13へ示す。推定値算出には補正係数の中間値である、0.948を乗じた。図-13中では1格子あたり縦4枚×横3枚の全12枚の衛星写真画像中での推定値の合算で示している。

図-13より、tricycleの分布傾向を視覚的に表現できた。得られたtricycleの分布の傾向を分析した上で、その市場特性を検討の上、計画への活用を期待する。

6. 本研究のまとめと今後の課題

本研究では、GoogleEarth上を走行するtricycleを対象に、深層学習手法を基にした車両検出ツールの適用性を明らかにし、ツールを用い分布傾向を明らかにした。今後の課題として、他物体検出手法との比較や、他衛星写真画像や他LAMATに対する本研究手法の適用可能性等について検討を行う必要がある。

表-5 tricycle推定台数結果

信頼区間	95%		99%	
	下限	上限	下限	上限
推定台数(台)	36.366	88.770	27.522	97.614



図-13 ツールにて推定されたtricycle推定分布

参考文献

- 1). Veng Kheang PHUN, Tetuso YAI (2015), State of the Art of Paratransit Literatures in Asian Developing Countries, EASTS 11th Conference, pp. 1169
- 2). Girshick, Ross, et al. "Rich feature hierarchies for accurate object detection and semantic segmentation." Proceedings of the IEEE conference on computer vision and pattern recognition. 2014
- 3). Rui Wang, Toshiyuki Okamura, Yutaro Kaneko, Fumihiko Nakamura(2013), "Passengers' Attitudes to the Service items of Jeepneys in Metro Manila by Different Lifestyles", Journal of the Eastern Asia Society for Transportation Studies, Vol10,23
- 4). Tri Basuki Joewono, Hisashi Kubota(2005), "The Characteristics of paratransit and Non-Motorized transport in Baudung, Indonesia", Journal of the Eastern Asia Society for Transportation Studies, Vol6,pp.262-277,2005
- 5). Sundaresh Ram,Jeffery.J.Rodriguez(2016),"Vehicle Detection in Aerial Images Using Multiscale Structure Enhancement And Symmetry",Icip(IEEE) pp.3817-3821
- 6). Shenquan Qu,Ying Wang,Gaofeng Meng, and Chunhong Pan(2016),"Vehicle Detection in Satellite Images by Incorporating Objectness and Convolutional Neural Network",Journal of Industrial and Intelligent Information Vol.4,No.2,March 2016
- 7). DmitrySincha,MikhailChervonekis,Pavel Skribtsov(2016),"Vehicle Detection and Classification in Aerial Images",Indian Journal of Science and Technology,Vol 9