

ディープラーニングを活用した OD 交通量と 経路調査方法の基礎研究

安藤 正幸¹・形屋 陽一郎²

¹正会員 株式会社日本海コンサルタント AI 技術室 (〒921-8042 石川県金沢市泉本町 2 丁目 126 番地)
E-mail: m-andou@nihonkai.co.jp

²正会員 株式会社日本海コンサルタント AI 技術室 (〒921-8042 石川県金沢市泉本町 2 丁目 126 番地)
E-mail: y-kataya@nihonkai.co.jp

交通計画や道路整備に当たり、道路利用者の起終点交通量 (OD) 交通量及び利用経路を把握することは重要事項である。しかし、現在使用されている OD 交通量は、サンプル数やリアルタイム性に問題がある。また、OD 交通量調査や経路調査において、車両照合のためナンバープレート調査が多用されるが、ドライバーのプライバシーへの過剰な反応が危惧される。

このため、人工知能 (AI) におけるディープラーニング (CNN) を活用し、ナンバープレートではなく車両の形と色で車両照合を行い、高精度でリアルタイムな OD 交通量と利用経路を把握する手法を研究した。研究の結果、教師写真及びマッチング対象車両写真をマスキングし、適切な小片写真に細分化することで照合精度の向上が図られた。

Key Words: AI, Deep Learning, CNN, Origin and Destination survey, Route survey

1. 研究の背景と目的

交通計画において、道路利用者の起点と終点が明瞭な交通量 (OD 交通量) の把握は、交通計画・マネジメントを行う上で大変重要である。しかし、現状の交通計画に用いる OD 交通量は、全国道路・街路交通情勢調査の起終点調査によるものであり、表-1 に示す問題点が挙げられる。このため、これまで多くの OD 交通量取得方法が研究されている²⁾。

また、整備対象路線の選定や整備優先度決定に当たり、道路利用者の経路把握も重要である。経路は ETC2.0 データなどプローブデータにより把握できるが、プローブ機器を搭載した特定の車両の経路であるため、より多くの車両の経路情報を把握する必要がある。

表-1 現況の OD 交通量の問題点

①サンプル率は 2.3% ¹⁾ と低いため、より現状を反映する (サンプル率が高い) データが必要である。
②5 年毎の調査結果であるため、調査年以外の道路整備や大規模店舗の開業などによるリアルタイムな変化が反映されない。
③起終点間の一日の合計交通量であるため、渋滞対策に必要な朝夕ピーク時間帯の動向が分からない。

さらに、これまでナンバープレート調査により OD 交通量や経路を調査してきたが、近年ではプライバシーに対する意識が高いため、ナンバープレート調査に対する過剰な反応が危惧される。

このような現状を踏まえ、本研究の目的は、プライバシーに配慮しつつ、現状よりサンプル率が高い OD 交通量及び経路情報を把握する調査手法の開発とする。

2. 予備研究の概要

(1) OD 交通量及び経路の調査方法

OD 交通量及び利用経路の把握は、次の手順で行う。

【OD 交通量及び利用経路の把握方法 (図-1 参照)】

- ①計画地の出入部及び主要交差点にビデオカメラ (A ~D) を設置し、撮影する。
- ②流入地点のカメラ A により計画地への流入車両を撮影する。
- ③流出部カメラ D で撮影された車両と流入車両を照合し、一致した場合 OD 交通量として集計する。
- ④主要交差点 (カメラ B, C, D) で撮影された車両と流入車両を照合し、経路を推定する。

(2) 車両の照合方法

これまでの車両照合は、車両固有のナンバープレートを用いて同一車両を判定していた。具体的には、調査時に通行車両のナンバープレートを調査員が読み上げ、録音（記録）し、調査後、調査結果の整理しデータを入力する。調査結果から、ナンバープレートが一致する車両の確認調査地点や時系列を整理し、起終点及び経路を特定する。このようなナンバープレート調査では、以下の問題が生じることがある。

【ナンバープレート調査における課題】

- ①片側 2 車線以上の道路で歩道側の第 1 車線は問題ないが、より道路中央側の車線におけるナンバープレートが読みづらい。
- ②ピーク時など交通量が多いと記録が間に合わない。
- ③誤認・読み間違い・記録ミスの可能性はある。

本研究における車両照合は、プライバシーに配慮し、ナンバープレートによる照合ではなく、AI（人工知能）による画像解析を用いて、車両の「形」と「色」で照合する。照合方法を図-2に示す。

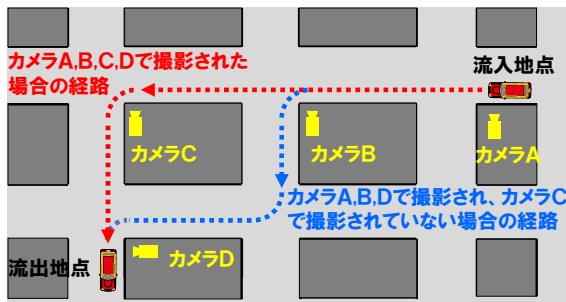


図-1 OD 交通量・経路把握説明図

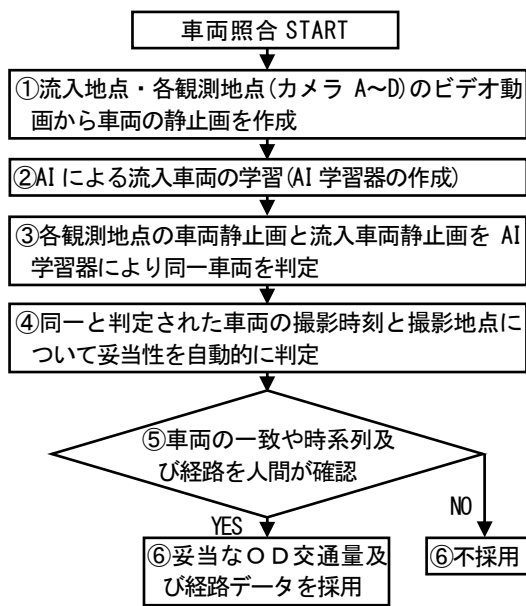


図-2 車両照合フロー図

(3) 学習方法

車両照合における「②AI による流入車両の学習」は、機械学習「教師あり学習」とし、流入車両静止画を「教師写真」とする。また、照合する写真（カメラ B～D）を「マッチング写真」とする。なお、カメラの「傾き」や「設置地点の明るさ」及び「設置高さ」さらには「車両との離れ」が各カメラで異なるため、流入車両静止画の回転・明るさ変更・角度補正・拡大縮小を組み合わせた画像を作成し、教師写真として学習する。

学習方法は、ディープラーニング（深層学習）により行い、アルゴリズムは画像認識が得意とされている Convolution Neural Network（CNN）によりモデルを構築した。プログラミング言語は Python を使用し、フレームワークは Chainer を採用した。

(4) 車両照合実験の実施とその結果

AI による車両照合の精度や課題を確認するため、車両照合実験を行った。実験は、P1～P3 に Web カメラを設置し、8 台の車両を各カメラで撮影した(図-3 参照)。

カメラの設置位置による影響を確認するため、P1 はやや日影の地上約 7m の位置に設置し、P2, P3 は撮影条件が類似した明るい場所の地上約 3m の位置に設置した。

車両照合は、①P1 の画像を教師写真として学習し、P2, P3 を照合した場合、②P2 の画像を教師写真として学習し、P1, P3 を照合した場合、③P3 の画像を教師写真として学習し、P1, P2 を照合した場合の 3 ケースを行った。車両照合結果を図-4に示す。

実験結果より、P1 の画像を学習し、撮影条件が異なる P2, P3 の画像と照合した場合の正答率は約 3～4 割と低く、同様に P2 もしくは P3 の画像を学習し、撮影条件が異なる P1 の画像と照合した場合の正答率も約 1～2 割と低いものであった。これに対し、撮影条件が類似した P2 と P3 による照合（P2 を学習し P3 を照合、P3 を学習し P2 を照合）の正答率は約 7 割であり、比較的高い精度であった。以上より、撮影条件が類似していれば高精度の照合が可能であることが判明した。

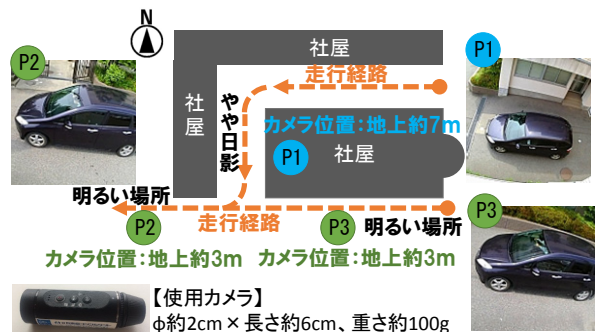


図-3 実験概要の説明図

3. 教師写真の加工による照合精度の向上

(1) 予備実験における課題と改善点

予備実験において、車両照合の正答率は、カメラ設置位置が類似している場合でも約 7割であり、カメラ設置位置が類似しない場合は約 1~4 割程度であった。正答率の低い要因は、カメラ設置位置に加え、画像には車両と背景（建物や植栽及びフェンスなど）が写っており、背景も車両の情報として学習するため、背景が異なった場合、同一車両と判定できないことが考えられる。

この改善点として、下記事項を実施する。

a) 改善点 1：マスキング

教師写真とマッチング写真の車両以外の部分を同一色でマスキングして、不要な情報を削除する。

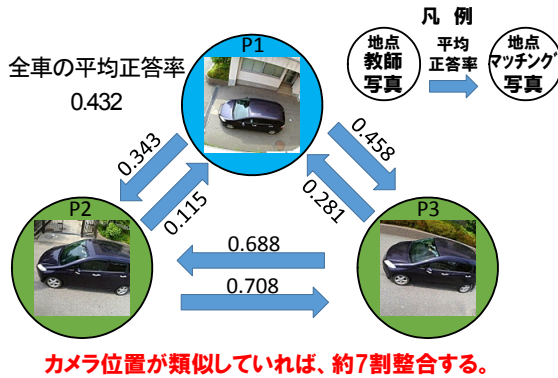


図-4 実験結果の説明図

表-2 今回の研究内容及び検討事項

手順	今回の研究内容
1	3 観測地点 (P1, P2, P3) のビデオ動画から実験車両 (8 台) の静止画を作成 (3 箇所×8 台×5 枚/台)
2	作成した静止画をマスキングした後、拡大縮小・回転・彩度変更を行い、16pixel 毎に細分化し、小片写真を作成
3	各観測地点の小片写真 (写真中のマスキング面積が 20%以下) を用いて、地点ごとの学習器を作成
4	車両照合の方法：多数決による判定 マッチング対象車両の静止画から 50 枚の小片写真を作成した時、学習器が 50 枚中の 25 枚以上を「1 号車」と判定した場合、マッチング対象車両は 1 号車と判定
5	【今回の検討事項】 ①細分化サイズ検証：4 サイズを比較検討 ②学習回数の検証：10 ケースを比較検討 ③学習プログラムの検証：3 タイプを比較検討

b) 改善点 2：写真の細分化

ディープラーニングに当たり、漫然と車両の全体的な色や形を学習するのではなく、フロントグリルやテールランプ及びアルミホイールなど特徴的な部分を意図的に学習・判定するために、教師写真及びマッチング写真を細分化し、小片写真を作成する。

以上の改善を踏まえ、今回の研究内容を表-2 に示す。また、具体的な内容を以下に詳述する。

(2) マスキング方法

教師写真及びマッチング写真は、各調査地点のビデオ動画 (60fps で撮影) からナンバープレートが写っていない静止画 5 枚を切り出した。マスキングは、切り出した静止画の車両以外の部分を、車両装飾色としても利用頻度が少ない緑色で塗りつぶした。

(3) 静止画の細分化と細分化サイズの検討

マスキングされた教師写真及びマッチング写真を細分化し、小片写真を作成した。この時、小片のサイズを小さくすると、車両の小さな部分に着目して学習するため、形状の情報が乏しくなる。また、小片のサイズが大きくなると車両全体を捉えることになり、特徴的な部分の占める割合が減少する。細分化サイズの差異による正答率の差異を分析するため、マスキングされた元写真 (640×360 pixel) を 4 種類の大きさ (32×32 pixel, 64×64 pixel, 128×128 pixel, 256×256 pixel) に細分化した。また、細分化は、16pixel ずつ左方向及び下方向にずれた画像を切り出した (図-5 参照)。

なお、マスキング面積が大部分を占める写真を学習す

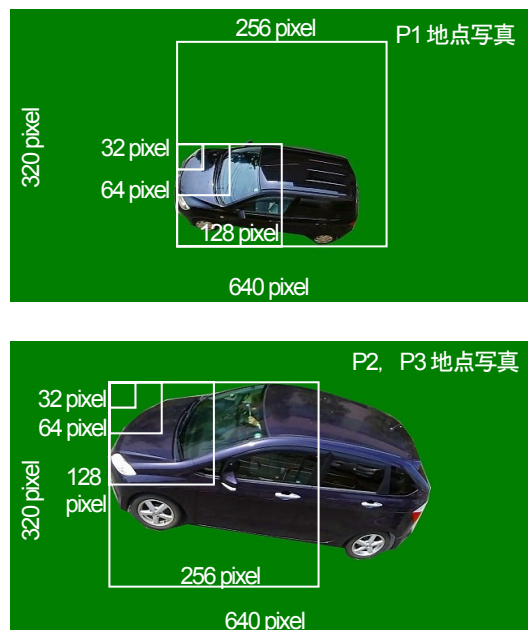


図-5 マスキング写真と写真の細分化 (1号車の例)

ると、別車のマスキング面積が大部分を占める写真と区別できなくなるため、マスキング部分の面積が全体の 20%以下となる写真を採用した。ただし、256×256 pixel では車両以外の部分が多く、20%以上を破棄した場合教師写真が乏しくなるため、60%のマスキング部分を許容した。

(4) 教師写真及びマッチング写真の整理

8 台の実験車両に対して、細分化した教師写真及びマッチング写真の枚数を表-3 に示す。

なお、教師写真の細分化においては、カメラの設置角度やカメラの傾き及びカメラ設置位置の明るさの差異があるため、マスキング写真に対して細分化の前に、拡大縮小・回転・彩度変更の加工を施した。

(5) 車両判定方法（多数決による判定）

車両照合において、マッチングする車両が、学習した実験車両 8 台のうち、どの車両かを判定する。マッチング写真は 1 台の写真を細分化した数十枚の小片写真であり、この小片写真を学習機に投入し、学習車両の内いずれの車両かを判定する。この判定により過半数（閾値 50%以上）が指摘する車両を、当該車両の判定とする。

【判定例】

不明車両（実は 1 号車）の静止画から 50 枚の小片写真を作成した時、

- ①50 枚の内 26 枚以上が「1 号車」と判定されれば、不明車両は 1 号車と判定する（正答）。
- ②50 枚の内 26 枚以上が「2 号車」と判定されれば、不明車両は 2 号車と判定する（誤答）。
- ③判定が分散し、過半数を超えて判定された車両が無い場合は、「車両不明」とする（誤答）。例えば、「1 号車」と判定された枚数が 20 枚で、最も多い場合でも、過半数を超えていないため、「車両不明」とする（誤答）。

表-3 細分化した教師写真・マッチング写真枚数

サイズ	P1	P2	P3
32×32 pixel	3,527,920 3,779	7,588,088 8,310	7,667,000 8,424
64×64 pixel	2,605,955 2,595	6,310,370 6,685	6,402,561 6,809
128×128 pixel	1,179,134 1,036	4,079,152 4,171	4,129,917 4,210
256×256 pixel	1,276,275 972	4,375,624 4,893	4,320,591 4,833

上段：教師写真枚数，下段：マッチング写真枚数
※1 台当たり 5 枚の静止画を抽出し、8 台分の合計値

(6) 学習回数の検討

ディープラーニングにおいて、学習回数が重要な要素となるため、学習回数は 1,000~10,000 回の 1,000 回毎に 10 個の学習器を作成し、正答率を比較した。

(7) 学習プログラムの検討

画像解析のアルゴリズム（CNN）において、畳み込み層とプーリング層の層数により照合精度が異なると考えられるため、畳み込み層とプーリング層を 2 層、3 層、4 層とした場合の 3 タイプのプログラムを作成し、正答率を比較した。

4. 車両照合結果

(1) 学習器精度の検証

車両照合に当たり、学習器の精度を確認するため、学習した写真を作成した学習器で照合し、正答率について検証した。検証の結果、32×32 pixel では、層数に依存せず、学習回数が多くなると正答率は 1 になるが、64×64 pixel 以上では、層数や学習回数に関係なく、正答率は 1 であった。なお正答率は、次式(1)により算出した。

$$\text{正答率} = \text{正答台数} / \text{全実験台数 (8 台)} \quad (1)$$

(2) 車両照合結果

車両照合は、3 調査地点（P1, P2, P3）において、各々学習器を作成し、他の調査地点の車両を照合した。

照合結果を表-5~7 に示す。なお、学習回数が異なっても正答率が等しくなる場合があった。

a) P1 地点の車両を学習し、P2, P3 地点の車両を照合

表-4 より、正答率は写真サイズの増大に伴い向上するものの、256×256 pixel では 60%のマスキング面積を許容した（画像の中に車両の占める割合が低くなる）ため、

表-4 P1 地点の車両を学習し P2・P3 地点の車両を照合した時の全車の平均正答率

学習地点 照合地点	P1					
	P2			P3		
	2層	3層	4層	2層	3層	4層
32×32 pixel	0.625 2,000	0.600 1,000	0.750 7,000	0.600 1,000	0.600 9,000	0.675 7,000
64×64 pixel	0.750 8,000	0.750 3,000	0.600 8,000	0.750 1,000	0.750 3,000	0.625 1,000
128×128 pixel	0.700 9,000	0.750 8,000	0.750 1,000	0.800 6,000	0.750 2,000	0.750 8,000
256×256 pixel	0.425 4,000	0.500 1,000	0.375 2,000	0.450 4,000	0.375 2,000	0.300 6,000

上段は正答率，下段は学習回数

表-5 P2 地点の車両を学習し P1・P3 地点の車両を照合した時の全車の平均正答率

学習地点	P2					
	P1			P3		
照合地点	2層	3層	4層	2層	3層	4層
32×32 pixel	0.625	0.450	0.500	0.850	0.875	0.875
	4,000	3,000	5,000	8,000	8,000	3,000
64×64 pixel	0.750	0.750	0.850	0.875	0.975	1.000
	1,000	2,000	7,000	1,000	9,000	4,000
128×128 pixel	0.725	0.875	0.750	0.925	0.975	1.000
	9,000	1,000	9,000	5,000	6,000	4,000
256×256 pixel	0.625	0.500	0.525	0.900	0.900	0.900
	9,000	1,000	10,000	5,000	7,000	6,000

上段は正答率, 下段は学習回数

表-6 P3 地点の車両を学習し P1・P2 地点の車両を照合した時の全車の平均正答率

学習地点	P3					
	P1			P2		
照合地点	2層	3層	4層	2層	3層	4層
サイズ	2層	3層	4層	2層	3層	4層
	32×32 pixel	0.650	0.525	0.500	0.875	0.875
64×64 pixel	0.875	0.750	0.725	1.000	1.000	1.000
	9,000	1,000	6,000	4,000	5,000	3,000
128×128 pixel	0.700	0.850	0.875	1.000	1.000	1.000
	2,000	4,000	8,000	4,000	4,000	4,000
256×256 pixel	0.700	0.850	0.875	0.975	1.000	1.000
	2,000	4,000	8,000	6,000	6,000	1,000

上段は正答率, 下段は学習回数

正答率が低下したものと考えられる。また、P1 地点は高さ約 7m 地点にカメラを設置しており、P2、P3 地点は高さ約 3m 地点にカメラを設置しているため、撮影された車両の大きさが大きく異なっている (図-6 参照)。教師写真とマッチング写真における車両の大きさの差異により、照合精度が悪いものと思われる。

b) P2 地点の車両を学習し、P1、P3 地点の車両を照合

表-5 より、正答率は写真サイズの増大に伴い向上するが、256×256 pixel では低下する傾向は a) と同様である。被写体車両の大きさが異なる P2 と P1 では正答率は a) とほぼ同様であるが、被写体車両サイズがほぼ等しい P2 と P3 では、4 層構成の CNN プログラムで正答率が 1.0 となった。

c) P3 地点の車両を学習し、P1、P2 地点の車両を照合

表-6 より、正答率は写真サイズの増大に伴い向上するが、256×256 pixel で低下する傾向は a)、b) と同様である。被写体車両の大きさが異なる P2 と P1 では正答率は a) とほぼ同様であるが、被写体車両サイズがほぼ等しい P2 と P3 では、CNN プログラムの層数にかかわらず、64×



教師写真 (1号車)
98,835~1,223,926
枚/台/サイズ

マッチング写真 (1号車)
8~274
枚/台/サイズ

図-6 教師・マッチング写真の小片集合写真



正答写真 64 枚

誤答写真 72 枚

正答率=64枚/全照査数 136枚=0.471

図-7 正答及び誤答写真の小片集合写真

64 pixel 及び 128×128 pixel で正答率 1.0 となった。

以上より、今回の研究では、全体の整合率が最も高かったのでは、3 層構成の CNN プログラムで 128×128 pixel の写真を学習・照合する場合であった (全車の平均正答率 0.821)。この時のディープラーニングにおける学習回数は 4,000 回であった。なお、学習器はランダム抽出した写真を学習しているため、偶然性があり、4,000 回が最適回数とは言い難い。

(3) 車両照合の考察

車両照合の考察にあたり、例として、P1 地点の教師写真及びマッチング写真の集合写真を図-6 に示す。

照合結果の正答写真及び誤答写真の小片集合写真を図-7 に示す。これより、学習車両と照合車両の被写体の大きさが大きく異なり、撮影された角度が異なるものの、1号車を 1号車と判定した正答枚数は 64 枚と最も多く、正答率は 0.471 であった。しかし、閾値 50%未満であるため、「車両不明」と判定する。

ここで、誤答写真の小片集合写真と正答の小片集合写真を比較すると大きな差はない。そこで、誤答箇所を分類し、誤答車両の教師写真を対比した (図-8 参照)。

対比の結果、色が近似している 3号車や 5号車との誤答が多く確認された (61 枚 44.9%)。しかし、色が全く異なる 2号車や 6号車との誤答もわずかではあるが、確認された (11 枚 8.1%)。

以上より、1号車に対し最も多く誤認された 5号車は、



図-8 誤答箇所と教師写真との対比 (128×128 pixel)

フロントランプ形状や後部座席のドアノブ位置が異なるものの、同程度の大きさのワンボックス車両であり、全体的に色が近似しているため、誤認が大きかったものと考えられる。また、3号車は大きさは異なるものの、色が似ているため、小片化し、部分的に拡大して判定したため、誤認されたものと考えられる。なお、枚数は少ないものの色が全く異なる車両との誤認は、写真の明暗の補正を図るため、教師写真に彩度を変更したが、この影響が大きいものと考えられる。よって、明度の変更は軽微に行う、もしくは変更しない方が良く考えられる。

(4) 閾値の感度分析

車両照合に当たり、対象者車両の特定において、マッチング写真の過半数 (閾値 50%以上) が指摘する車両を、当該車両の判定した。より厳密に対象車両を特定するために閾値をより高くすることが考えられる。

そこで、閾値を 50~80%とした場合の正答率について整理した (図-9 参照)。

図-9 より、閾値を上げることで対象車の特定が厳密となるが、逆に先述した例に示す通り、マッチング写真の中で最も多く正答車と判定されているのに、閾値により

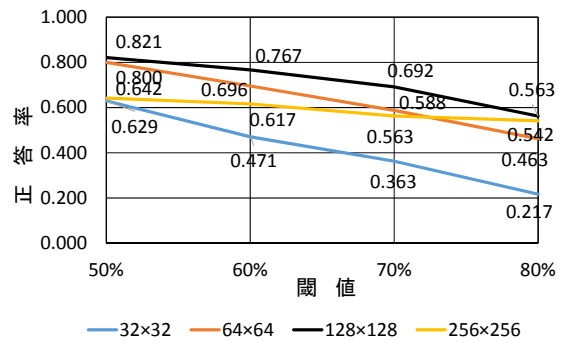


図-9 閾値と正答率の関係

表-7 検討項目とその内容

検討項目	検討内容
細分化サイズ	4 サイズ (32×32 pixel, 64×64 pixel, 128×128 pixel, 256×256 pixel)
学習回数	10 ケース (1,000~10,000 回の 1000 回毎)
学習プログラム	3 タイプ (CNN の層構成を 2 層, 3 層, 4 層)

「車両不明」と判定されるため、正答率が減少する。今回は、実験車両が 8 台と少ないため、閾値を過半数 (50%を超える) としたが、より多くの車両を学習した場合、より低い閾値であっても正答である場合が想定されるため、実験台数を増やした検証が必要である。

4. 研究結果のまとめと今後の課題

(1) 研究結果のまとめ

OD 交通量や経路調査のため、車両の色・形による車両照合手法の研究を行った。車両照合は、ディープラーニング (CNN) の教師あり学習とした。学習に当たり、車両以外の風景が写った写真をそのまま学習した場合、精度が低い結果であった。これを踏まえ、車両以外の風景をマスキングし、細部の情報を学習するため、マスキング写真を細分化した小片写真を学習した。この時、小片写真の大きさや学習回数及びプログラムの構成により照合精度も変わるため、表-7 に示す検討項目を組み合わせさせて照合精度を比較した。

この結果、128×128 pixel の教師写真を 3 層構成の CNN プログラムで 4,000 回学習した場合が最も正答率が高かった (表-8 参照)。

なお、本研究に使用した PC のスペック及び各計算時間を表-9 に示す。

表-8 学習条件と全車の平均正答率

学習条件	全車の平均正答率
車両以外をそのまま学習	0.432
車両以外をマスクし、 細分化した小片写真を学習	0.821 (類似撮影条件では 1.000)

(2) 今後の課題とフィールド実験

今回の研究により教師写真の加工により車両照合の精度向上が確認されたが、以下の課題があると考ええる。

課題 1：実験車台数が少ない。

課題 2：明るさの変化が少ない

課題 3：片側 2 車線以上の道路での精度

これらの課題を踏まえ、今後は、実際の道路において、フィールド実験を行う。具体的には、以下の内容を予定する。

【フィールド実験内容】

- ①比較的小規模のエリアにカメラを複数台設置し、車両を撮影する。
 - ②一般車の中に起終点及び経路が既知であるモニター車を複数台走行させる。
 - ③一般車及びモニター車写真のマッチングにより起終点及び経路を推定する。
 - ④学習台数と照合結果の関係や、既知の経路と推定経路の差異及び片側 2 車線以上の道路における照合精度を確認し、実用化に向けた課題を明確化する。
- また、AI によるナンバープレートの認識精度が高くなっていることを踏まえ、ナンバープレートを暗号化するブラックボックスを作成して車両を識別することが考えられる。具体的には以下の流れで車両を識別する ID を作成する。

表-9 PC スペックと所要時間

項目	スペック・所要時間
CPU	Intel®Xeon®CPU E5-1650 v2@3.5GHz 3.5GHz
RAM	16 GB
GPU	ASUS® GEFORCE GTX 1080 8GB
OS	Windows8.1 Pro 64bit
プログラム言語	Python
フレームワーク	Chainer
教師データ 作成時間	約 90 分/台/サイズ
学習時間	約 22~35 分/1000 回/8 台/サイズ
照合時間	26~32 秒/台/サイズ

【ナンバープレート暗号化ブラックボックス】

- ①ビデオカメラで撮影されたナンバープレートから登録地、種別番号、ひらかな及び一連指定番号を認識する。
 - ②認識された内容を入力値として、暗号化プログラムにより車両を識別する ID 番号を自動発生させる。
 - ③ID 作成の後、ビデオカメラ映像を自動削除する。
 - ④一連の作業をブラックボックスが自動で行う。
- 作成した ID で車両照合を行い、OD 交通量と経路を調査することが可能と考える。

参考文献

- 1) 交通センサスの概要, <http://www.mlit.go.jp/road/ir/ir-council/census2/1st/3.pdf>
- 2) 例えば、鈴木健太郎、萩原武司、林竜太郎、斎藤多恵子：スマートフォンアプリを活用した起終点調査の可能性に向けた検討, 土木計画学研究・講演集, Vol49, pp217 1-8, 2014.

(2018.7.31 受付)

BASIC STUDY ON OD TRAFFIC VOLUME AND ROUTE SURVEY METHOD UTILIZING DEEP LEARNING

Masayuki ANDOU, Yoichiro KATAYA

In traffic planning and road improvement, it is important to grasp the origin and destination (OD) traffic volume and usage route of road users'. However, the OD traffic volume in use has problems with the amount of sample and the adjustment recent demand. Also, in OD traffic survey and route survey, license plate survey is frequently used for vehicle matching, however it is concerned excessive reaction to driver's privacy.

For this reason, the method which grasp highly accurate and adjusted the real OD traffic volume and usage path with using Deep Learning (CNN) in Artificial Intelligence (AI) to match up vehicles by its shape and color rather than license plate was studied. As a result of this study, matching accuracy was improved by masking and subdividing the teacher pictures and matching pictures.