

データ利活用における車載映像収集基盤と個人情報保護

Platform of Collecting On-vehicle Video Images and Protection of Personal Information for Data Utilization

岡田 康貴
Yasutaka OKADA

関 竜介
Ryusuke SEKI

村下 君孝
Kimitaka MURASHITA

山田 浩
Hiroshi YAMADA

要旨

自動車のIoT化が進み、車載機とクラウドとの常時接続によって新たなサービスやビジネスを創出するコネクティッドカー社会が到来している。CPSの世界では、フィジカル空間の情報をコネクティッドカーが情報収集、センターで知識化し、新たな付加価値を創出する。この付加価値創出の源泉となる車載機で収集したデータの活用が期待されている。

その一つとして、車両周囲の映像があるが、車載カメラ映像には歩行者の顔やナンバープレートなど個人情報が含まれている。EUではGDPRが施行されており、社会的にも個人情報保護が強化されている。

映像は非構造データであるため、そのままでは活用が難しく、映像の分析には映像が持つ情報を抽出するタグ付けが必要である。また、さまざまなサービスへの適用に向けては、タグ付けの高度化と個人情報保護の両立が重要である。そのためには、車載機上で物体検出技術を実行し、収集した画像を特定するタグ付けと、タグ情報を活用した個人情報保護領域の特定とを実現する必要がある。

本稿では、物体検出処理の車載機動作の実現可能性と個人情報保護への適用可能性を考察した結果について報告する。

Abstract

Recently, there have been many developments related to the Internet of Things (IoT) in the automotive sector. Cars are continuously connected to the cloud through on-vehicle devices, which in turn has led to the connected car society which creates a new service and business. In the world of Cyber Physical System (CPS), huge data in the physical space is collected by connected cars, and converted into knowledge in the center, which creates new added values. The utilization of data collected by on-vehicle devices which are the source of this value creation is expected.

A video image around the vehicle, as one of the data, is captured by the on-vehicle cameras and may include personal information such as the face of pedestrians and number plates of vehicles. The General Data Protection Regulation (GDPR) is in force in the EU for the protection of such personal information.

As the video image is non-structured data, it is difficult to be utilized in its original form. Therefore video analysis requires tagging to extract information contained in the video. Furthermore, it is important to implement both of sophistication of the tagging process and the protection of personal information so that the video image can be utilized in a variety of services. In order to achieve this objective, it is necessary to equip the on-vehicle devices with object detection technology, which can tag the collected images for identifying them and define the area of personal information protection by utilizing the tagged information.

This study reports the feasibility of object detection processing by on-vehicle devices and the application possibility of personal information protection.

1. はじめに

自動車のIoT化が進み、車載機とクラウドとの常時接続によって新たなサービスやビジネスを創出するコネクティッドカー社会が到来している。コネクティッドカーにおけるCPS (Cyber Physical System) の世界では、フィジカル空間の情報をコネクティッドカー(車載機)が情報収集し、サイバー空間(センター)で大規模データ処理技術などを用いて知識化する。そこで新たに見出した価値ある情報・知見によって新しいサービスを創出し、さらにサービスを車載機経由でユーザーへ提供する(図1)。新たなサービスを創出するため、付加価値を生む源泉となる車載機で収集したデータの活用が期待されている。

一方で、車両周囲の映像の活用が期待されているが、車載カメラ映像には歩行者の顔やほかの車のナンバープレートなど個人情報が含まれている。EUにおいては2018年にGDPR (General Data Protection Regulation: 一般データ保護規則)が施行され、また、日本においても個人情報保護法にGDPRの要求事項を反映する改正が議論されており、社会的にも、個人情報保護が強化されている。

映像は非構造データであるため、そのままでは活用が難しく、映像の分析には映像が持つ情報を抽出するタグ付けが必要である。さまざまなサービスへの適用に向けては、タグ付けの高度化と個人情報保護の両立が重要である。そのためには、車載機上で物体検出技術を実行し、収集した画像を特定するタグ付けと、タグ情報を活用した個人情報保護領域の特定、とを実現する必要がある。

本稿では、物体検出処理の車載機動作の実現可能性と個人情報保護への適用可能性を考察した結果について報告する。

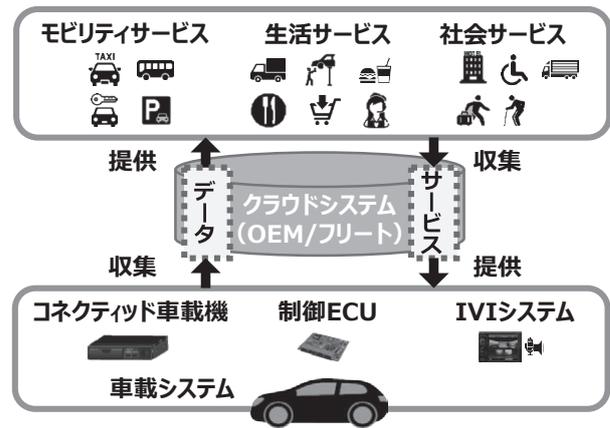


図1 Cyber Physical System (CPS)

2. コネクティッドカーにおけるデータ収集システム

図1のシステムを実現するためには、数百から数千万台のコネクティッドカーが収集した巨大なサイズの車載データを扱うため、

- ・センターに送信する通信コストが膨大になる
- ・センター側でのデータ処理に大量の時間とコストがかかる

という問題が生じる。また、収集した車載データには、データを利用するサービス提供者(サービサー)にとって必要でないデータも含まれている。そこで、当社では車載機のデータを効率よく収集するシステムとして、On-demand Data Collectionの開発を進めている(図2)。

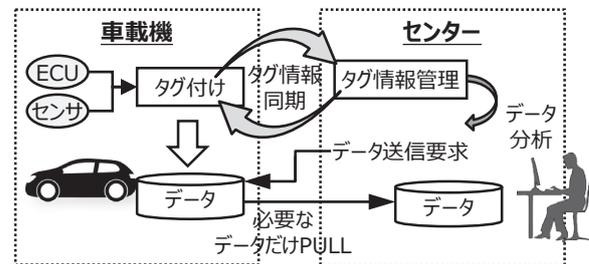


図2 On-demand Data Collection

On-demand Data Collectionでは、車速、操舵角、車両の位置・時刻情報などの車両データからイベントを検出しタグ付け(例えば"急ブレーキ","急

ハンドル"など)を行う。タグ情報だけをセンターに送り、サービサーは必要なデータだけをセンターに集められたタグ情報から検索し、車載機から受け取る。この仕組みにより、センター側へのデータ送信や、センター側でのデータ処理を最小化し、サービサーが必要とする情報だけを効率良く収集することができる。

現在、通信機能を持つ業務用ドライブレコーダー¹⁾や車両運行管理を支援するテレマティクスサービス向けデータロガー²⁾などへの適用を目指し、製品に向けた開発を進めている。これらの活用サービスとして、

- ・急ブレーキや急ハンドルなど、事故につながる可能性が高い場所の可視化
- ・タクシーやバス運転手の運転行動履歴収集などが期待されているが、さらに次のサービス事例にも活用したい、というニーズがある。

<サービス事例>

- ① Level3 自動走行車両開発
- ② 後続車両に向けた情報配信サービス
- ③ ダイナミックマップ作成やマーケティング

このようなサービス向けのデータ収集を実現するためには、フィジカル空間の物標（物体）を車載カメラで検出し、タグ付けする必要がある。図3に検出する物標の例を示す。

また昨今の個人情報保護強化の流れから、収集したデータに含まれる個人情報の保護も求められている。個人情報保護への対応としては、映像内の個人情報に該当する領域（歩行者の顔やナンバープレートなど）を個人情報のタグを利用してマスクする。

前記サービス事例①②③への対応と個人情報保護を両立するためには、車載機上でマルチクラスオブジェクト認識（以下、画像AI）を動作させ、タグ付けのための物標や個人情報保護のための領域を高精度に検出する必要がある。車載エッジへ適用可能な軽量、かつ、高性能な画像AIを開発す

る（図4）。



a) 群衆、子供、等



b) バリケード、三角コーン、等



c) 標識、看板、等

図3 サービス実現に必要な物標の例

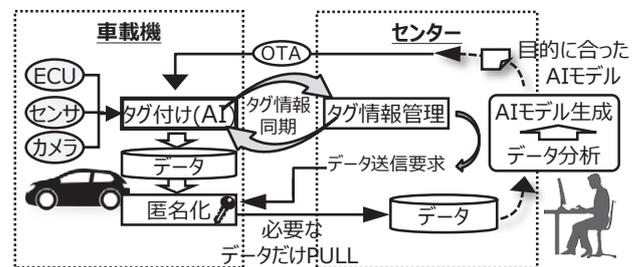


図4 画像AIの適用

3. DNNによる物体検出手法

収集した画像の意味づけや個人情報保護領域を特定するためのタグ付け手法として、Deep Neural Network (DNN) を用いた画像AIが、有力な候補と考える。

DNNは物体検知の一手法であり、これまでの方式から一線を画す高い検知性能をもつ手法であることから、近年、特に注目されている。DNNを用いたEnd-to-Endで物体検知可能な代表的なモデルに、Faster R-CNN³⁾、SSD⁴⁾、YOLOv3⁵⁾がある。

MS-COCO⁶⁾ データセットでのそれぞれのモデルの性能を表1に示す。

表1 主要な物体検知モデルの性能

物体検知モデル	DNN入力サイズ [pixel]	推論速度 [FPS]	精度 (mAP ^{*(1)})
Faster R-CNN	224	5.8	59.1
SSD	321	16.4	45.4
YOLOv3	320	45.5	51.5

表1より、Faster R-CNNは三つのモデルの中で最も高い精度をもつ一方で、推論速度はSSDやYOLOv3と比較して遅いことが分かる。すなわち、推論時間と精度にはトレードオフの関係があるといえる。今回、われわれが想定する用途で検証して、推論速度が速く精度も高いYOLOv3モデルを車載機で動かすDNNの候補として選択した。

まず、選択したYOLOv3モデルの性能を評価した。表2、表3に顔とナンバープレートのデータを学習し、推論速度と適合率^{*(2)}、モデルサイズを評価した結果を示す。顔はWIDER FACE⁷⁾で公開されている190,000枚のデータを使用し、ナンバープレートは今回の検証用に12,000枚の当社作成データを使用した。

表2 顔 (YOLOv3)

DNN入力サイズ	推論速度 [FPS]	適合率	モデルサイズ [MB]
224×224	30.8	0.87	246.3
416×416	14.3	0.94	246.3
608×608	7.8	0.95	246.3

表3 ナンバープレート (YOLOv3)

DNN入力サイズ	推論速度 [FPS]	適合率	モデルサイズ [MB]
224×224	23.7	0.85	246.3
416×416	13.6	0.94	246.3
608×608	8.5	0.94	246.3

次に、DNN入力サイズ224×224の適合率が低い原因を調査した。画像上の物体サイズと検知性能の関係を把握するため、ドライブレコーダー映像(1920×1080 [pixel])中の物体を物体サイズに基づいて三種類 (Small, Medium, Large) に分類した。三種類の物体サイズに関しては、MS-COCOで用いられている分類を踏襲した(表4)。

表4 画像中の物体のサイズ分類

サイズ分類	物体領域サイズS [pixel]
Small	$S \leq 32 \times 32$
Medium	$32 \times 32 < S \leq 96 \times 96$
Large	$96 \times 96 < S$

表5および表6に、三種類のDNN入力サイズ(224×224, 416×416, 608×608)における、表4の物体サイズに分類される顔およびナンバープレートの適合率と再現率^{*(3)}を示す。表5および表6より、224×224ではSmallサイズの適合率、再現率が有意に低いことが分かった。Smallサイズに分類される顔、ナンバープレートともに目視で識別可能であることから、Smallサイズでの高い適合率が必要である。416×416以上では物体サイズによる有意

*⁽¹⁾ mAP (mean Average Precision)

物体検知精度を評価する指標であり、ある物体の情報が与えられた時点までの適合率 (Precision) の平均であるAP (Average Precision) の平均で計算され、数値が高いほど物体検知の精度が高いことを示している。

*⁽²⁾ 適合率

正と予測したデータのうち、実際に正であるものの割合で表され、正確性を評価する指標である。適合率が高いほど、正確性が高いことを示している。

*⁽³⁾ 再現率

実際に正であるもののうち、正であると予測されたものの割合で表され、取りこぼしを評価する指標である。再現率が高いほど、取りこぼしが少ないことを示している。

な性能差はない。そのため、個人情報保護への適用には DNN 入力サイズは 416 × 416 以上とする必要がある。

表 5 顔の検知精度

DNN入力サイズ	物体サイズ	適合率	再現率
224×224	Small	0.77	0.46
	Medium	0.94	0.78
	Large	0.98	0.92
416×416	Small	0.92	0.70
	Medium	0.96	0.87
	Large	0.98	0.94
608×608	Small	0.93	0.74
	Medium	0.97	0.88
	Large	0.97	0.93

表 6 ナンバープレートの検知精度

DNN入力サイズ	物体サイズ	適合率	再現率
224×224	Small	0.74	0.68
	Medium	0.95	0.90
	Large	0.89	1.00
416×416	Small	0.90	0.92
	Medium	0.98	0.95
	Large	0.89	0.98
608×608	Small	0.92	0.95
	Medium	0.98	0.96
	Large	0.89	0.98

以上から、既存の YOLOv3 のモデルは入力サイズ 416 × 416 以上であれば高い検知精度を確保できる。しかし、表 2、表 3 より、DNN 入力サイズに関わらず、われわれが想定する車載機へ適用するにはモデルサイズが大きく、サイズ削減が必要なこと、また、DNN 入力サイズ 416 × 416 以上での推論速度の高速化が必要なが分かった。

4. 車載機への適用に向けた課題

DNN を用いた画像 AI を車載機へ適用するため、物体検出を高速かつ高精度に行える軽量なモデルが必要である。

<課題>

- (1) 車載機上でのリアルタイム推論処理
- (2) 車載プラットフォームに搭載可能な物体検出モデルのサイズ
- (3) 上記 (1) (2) と高い検知精度の両立

5. 課題へのアプローチ

高い検知精度を維持しつつ、推論の高速化、モデルサイズを削減する次の二つの軽量化技術の効果と検知精度への影響について、検証を行った。

(1) 推論速度の高速化

一般的に 32bit 浮動小数の範囲で計算される DNN のパラメータ (畳み込み係数、バイアス係数、など) を低 bit 化する量子化⁸⁾ が提案されている。16bit 浮動小数の範囲で計算する量子化を YOLOv3 に適用し、高速化の効果を検証する。

(2) モデルサイズ削減

モデルの軽量化手法として、DNN の量子化 (2 値化⁹⁾、3 値化¹⁰⁾、など)、スパース化¹¹⁾、ウェイトプルーニング¹²⁾ などが提案されている。これらの手法は、DNN のモデルサイズを大幅に小さくすることが可能であるが、高速に推論するために専用のハードウェアが必要であるため、車載機のコストが高くなるという問題がある。

専用のハードウェアを必要としない手法として、チャンネルプルーニング¹³⁾ が提案されている。チャンネルプルーニングは、DNN の演算量で支配的な畳み込み層の演算のうち、検出結果に影響しないチャンネルを除去することで、精度低下を抑えつつ、高速化、モデルのパラメータ数削減を実現する手法

である。チャンネルプルーニングを前項の量子化後のYOLOv3に適用し、高速化とモデルサイズの削減の効果を検証する。

6. 検証方法と結果

6.1 検証方法

検証の系を図5に示す。車載カメラとしてLogicool社製USBカメラ、車載エッジデバイスとして、AIプロセッサを搭載した組み込み機器であるNVIDIA社製のJetson Xavier、クラウド側としてPC上で動作する試作ソフトを開発し検証した。

また、軽量化手法は検知対象（顔、ナンバープレート）によらないため、本稿では、ナンバープレートのデータセットを用い、さらに推論速度と検知性能とのバランスが良いDNN入力サイズ416×416のモデルを用いて検証を進める。

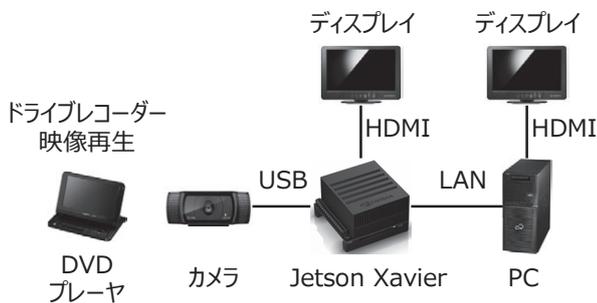


図5 検証系

6.2 結果

ナンバープレートの学習済みモデルに対し、次に示す(1)と(2)の評価を行った。

(1)量子化による高速化

表7にDNNパラメータを16bitに量子化したときのナンバープレートの推論速度、適合率、モデルサイズを示す。推論速度は、1万回の平均値である。DNNパラメータを量子化しても適合率を維持したまま推論速度を向上させられることを確認した。

表7 ナンバープレート（量子化）

	推論速度 [FPS]	適合率	モデルサイズ [MB]
量子化なし	13.6	0.94	246.3
量子化あり	18.6	0.94	246.3

(2)プルーニングによるモデルサイズ削減

16bit量子化後のYOLOv3モデルにチャンネルプルーニングを適用した結果を表8に示す。DNNのモデルサイズは、DNNのパラメータ数により決まる。プルーニングを繰り返し適用することにより、約6,130万個のパラメータを徐々に削減していく様子を確認できる。表9より、プルーニングを8回繰り返し適用したモデルでは、適合率を維持したまま推論速度向上（1.4倍高速化）とモデルサイズ削減（1/7.5）ができることを確認した。

表8 チャンネルプルーニングの適用結果

プルーニング回数	パラメータ数 [M]	モデルサイズ [MB]	推論速度 [FPS]	適合率
なし	61.3	246.3	18.6	0.94
1回	33.3	134.1	19.7	0.91
2回	20.0	80.4	22.2	0.93
3回	13.5	54.6	23.8	0.94
4回	10.6	42.9	22.7	0.94
5回	9.5	38.2	24.5	0.95
6回	8.9	36.0	23.1	0.94
7回	8.6	34.5	24.2	0.95
8回	8.2	32.9	25.3	0.94
9回	7.7	31.1	25.2	0.84

表9 ナンバープレート（量子化+プルーニング）

	推論速度 [FPS]	適合率	モデルサイズ [MB]
量子化	18.6	0.94	246.3
量子化 + プルーニング	25.3	0.94	32.9

6.3 考察

32bit 浮動小数点の範囲で計算される DNN のパラメータを低 bit の範囲に限定する量子化と、推論結果への影響が小さいパラメータを削減するプルーニング手法を用いて YOLOv3 の軽量化を行った。その結果、精度を維持したまま、推論速度を 1.9 倍高速化、モデルサイズを 1/7.5 まで削減できることを確認した (表 10)。さらなる軽量化の方式としては量子化の低ビット化(8bit 化、4bit 化など)や、YOLOv3 の特徴抽出器である Backbone network をより軽量のネットワークに変更すること、などがある。今後さらなる軽量化を実施し、より安価な車載機でも実行可能な DNN 軽量化技術の実現を目指す。

表 10 YOLOv3 と軽量化後の YOLOv3 の比較

	推論速度 [FPS]	適合率	モデルサイズ [MB]
YOLOv3	13.6	0.94	246.3
軽量化後 YOLOv3	25.3	0.94	32.9

次章では、個人情報のタグ情報を利用した情報匿名化への適用について説明する。

7. 個人情報保護への適用

個人情報の保護処理については、個人情報対象の特定にタグ情報を活用する。タグ情報は、YOLOv3 で検出される物体種別や物体数などをもとに付与される。顔とナンバープレートを検出した場合は個人情報のタグを付与し、画像中の当該物体領域をマスクする。

DNN 入力サイズ 416 × 416 のモデルを用いて、車室内のドライブレコーダー映像に対し、顔の検知を行った結果を図 6 に示す。また、ナンバープレートの検知を行った結果を図 7 に示す。

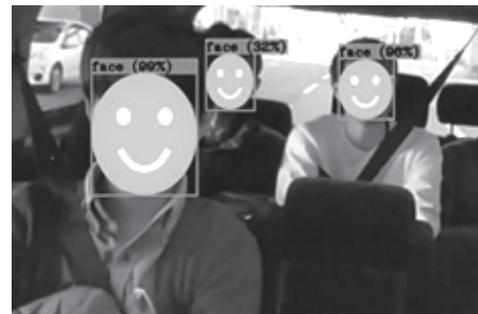


図 6 顔の匿名化結果



図 7 ナンバープレートの匿名化結果

顔やナンバープレートの検知枠の位置や大きさに大きな誤差はなく、個人情報の保護に適用できる精度があることを確認した。

また、検出位置や大きさにズレが生じて、マスクの大きさを拡大することでずれを吸収することが可能となる。個人情報対象の特定にタグ情報を活用することで、個人情報の保護処理に適用できることが分かった。

8. おわりに

本稿では、DNN による画像 AI を車載機へ適用するためのモデル軽量化と個人情報保護へ適用した結果について述べた。

まず、物体検知モデルに YOLOv3 を採用し、量子化やプルーニングによるモデル軽量化を行い、検出性能を低下させることなく、パラメータ数の削減と推論速度向上が可能であることを確認し、これら軽量化のアプローチが有効であることを示した。

次に、歩行者の顔やナンバープレートを個人情報のタグを利用してマスクできることを示した。

これらから、収集した画像を特定するタグ付けとタグ情報を活用した個人情報保護すべき領域の

特定を DNN による画像 AI で実現できることを確認した。

今後、通信機能を持つドライブレコーダーやフリースペースへの展開を進めていくため、車載向け組み込み評価ボードへの実装を進めるとともに、更なる軽量化に取り組んでいく。

参考文献

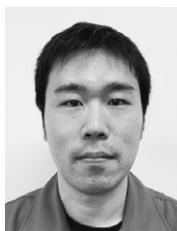
- 1) 業務用ドライブレコーダー G500
<https://www.denso-ten.com/jp/biz-recorder/g500/products/index.html>
- 2) データロガー
<https://www.denso-ten.com/jp/biz-recorder/g500/products/function03/index.html>
- 3) Ren, Shaoqing, et al. "Faster R-CNN: Towards real-time object detection with region proposal networks." *Advances in neural information processing systems.* (2015).
- 4) Liu, Wei, et al. "SSD: Single Shot MultiBox Detector." *arXiv preprint arXiv:1512.02325* (2015).
- 5) Joseph Redmon et al. "YOLOv3: An Incremental Improvement" *arXiv:1804.02767* (2018)
- 6) Xinlei Chen, et al. "Microsoft coco captions: Data collection and evaluation server" *arXiv preprint arXiv:1504.00325* (2015).
- 7) Shuo Yang, et al. "WIDER FACE: A Face Detection Benchmark" *arXiv:1511.06523* (2015).
- 8) Raghuraman Krishnamoorthi "Quantizing deep convolutional networks for efficient inference: A whitepaper." *arXiv:1806.08342* (2018)
- 9) M. Courbariaux, et al. "Binarynet: Training deep neural networks with weights and activations constrained to +1 or -1." *arXiv preprint arXiv:1602.02830* (2016).
- 10) Kyuhyeon Hwang, et al. "Fixed-point feedforward deep neural network design using weights +1, 0, and -1." *Signal Processing Systems* (2014).
- 11) Baoyuan Liu, et al. "Sparse Convolutional Neural Networks." In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2015).
- 12) Song Han, et al. "Learning both Weights and Connections for Efficient Neural Networks." *arXiv:1506.02626* (2015)
- 13) Zhuang Liu, et al. "Learning Efficient Convolutional Networks through Network Slimming." *arXiv:1708.06519* (2017)

筆者紹介



岡田 康貴
おかだ やすたか

共通技術推進部
先行開発企画室



関 竜介
せきりゆうすけ

共通技術推進部
先行開発企画室



村下 君孝
むらした きみたか

共通技術推進部
先行開発企画室



山田 浩
やまだ ひろし

共通技術推進部