

# ドライブレコーダーメーカーの視点 モノ+サービスの時代と人工知能

2018年6月26日(火)

矢崎総業株式会社

技術研究所

知能・情報技術研究部

杉田 明宏

1. 経歴
2. なぜ AI に注目しているのか
3. クラウドの活用
4. 教師データセット
5. 静止画のクラス分類
6. 静止画のオブジェクト検出
7. 次元数の小さい時系列データのパターン認識
8. 次元数の大きい時系列データのパターン認識の構想

年	職
2001	京都大学工学研究科分子工学専攻博士後期課程修了
2001	京都大学工学研究科分子工学専攻 日本学術振興会 (JSPS) 特別研究員 (PD)
2003	日本原子力研究所関西研究所光量子源システム研究グループ JSPS 特別研究員 (PD)
2004	日本原子力研究所関西研究所光量子源システム研究グループ 任期付研究員
2005	日本原子力研究開発機構レーザー物質制御研究グループ 任期付研究員
2008	矢崎総業株式会社 技術研究所 センシング技術研究部

アカデミック時代の専攻は、物理化学（専門は、レーザーを使う化学反応動力学）でした。

基本は実験屋ですが、必要に応じて量子計算も。

原子・分子の量子状態を操作し、直接、化学反応を操作する！

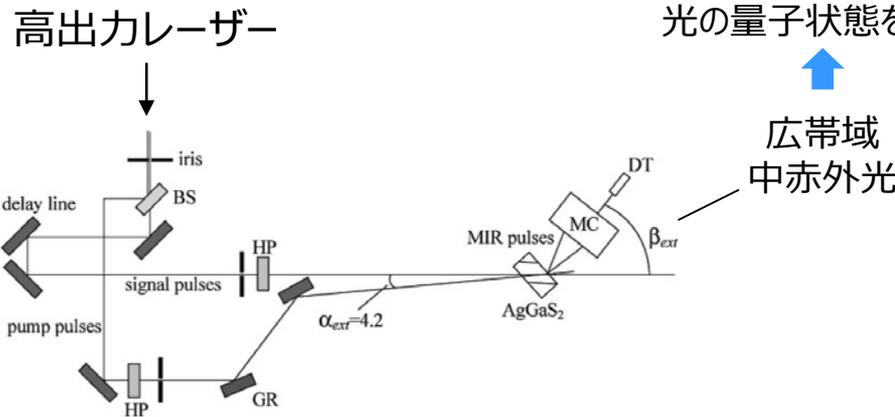
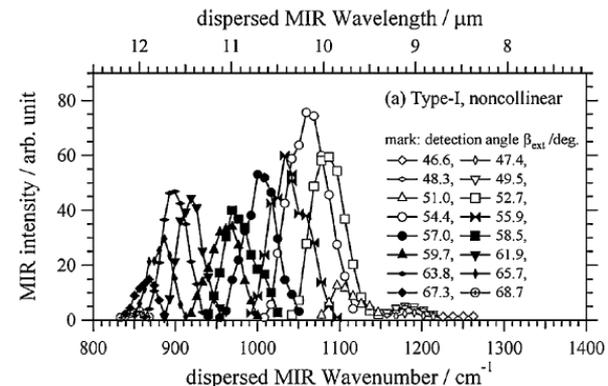


Fig. 4. Schematic layout of optics and beam lines. BS, beam splitter; HP, half-wave plate; GR, grating; MC, monochromator; and DT, HgCdZnTe detector.  $\beta_{ext}$  is the detection angle between signal and idler pulses.



A. Sugita et al., "Generation of Broadband Mid-Infrared Pulse by Noncollinear Difference Frequency Mixing", Jpn. J. Appl. Phys. 46, 226-228 (2007)

# 経歴 2/2

年	職	業務
2008	矢崎総業株式会社 技術研究所 センシング技術研究部 ⇒ 電子デバイス研究部	大気環境
2011	矢崎総業株式会社 技術研究所 ICT応用技術研究部	ディスプレイ
2017	矢崎総業株式会社 技術研究所 知能・情報技術研究部	AI

企業に就職しても、化学反応動力学の応用の一つである（光化学反応の）大気環境分野や光に関係するディスプレイの仕事に携わっていました。

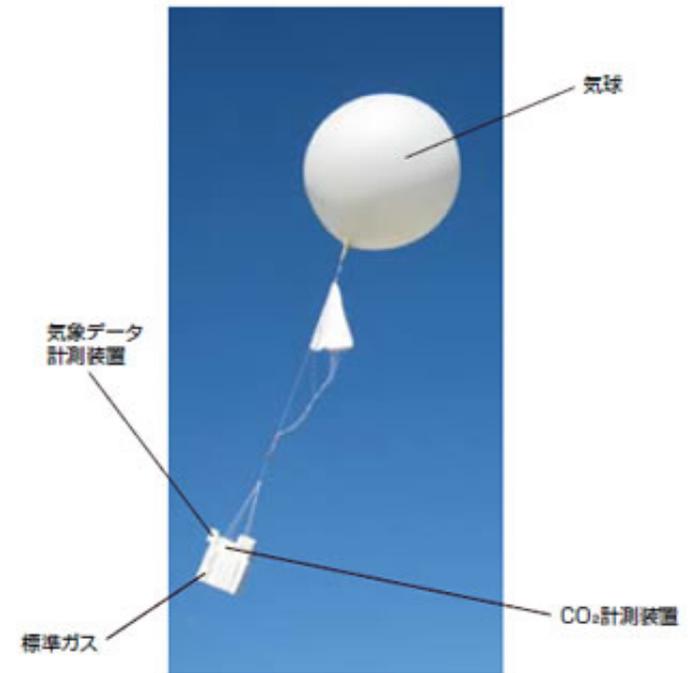


紆余曲折を経て、昨年2017年1月から AI に関する業務に従事しています。

ただし、Linux, HPC と数値計算の経験は、そこそこある。

(AI 初心者が担当者になり)

モノ視点に捕らわれているモノづくりメーカーが、いかに AI に取り組んでいくのか、という参考例を紹介する話と考えてください。



成層圏や対流圏における二酸化炭素の濃度を測定する CO<sub>2</sub> ゾンデ [1]

[1] 科学技術振興機構 (JST), 先端計測分析技術・機器開発プログラム, “高精度な二酸化炭素計測装置の開発”, [http://www.jst.go.jp/sentan/result/seika2012/03\\_05.html](http://www.jst.go.jp/sentan/result/seika2012/03_05.html)

# 企業紹介

## 矢崎総業株式会社

(<https://www.yazaki-group.com/>)

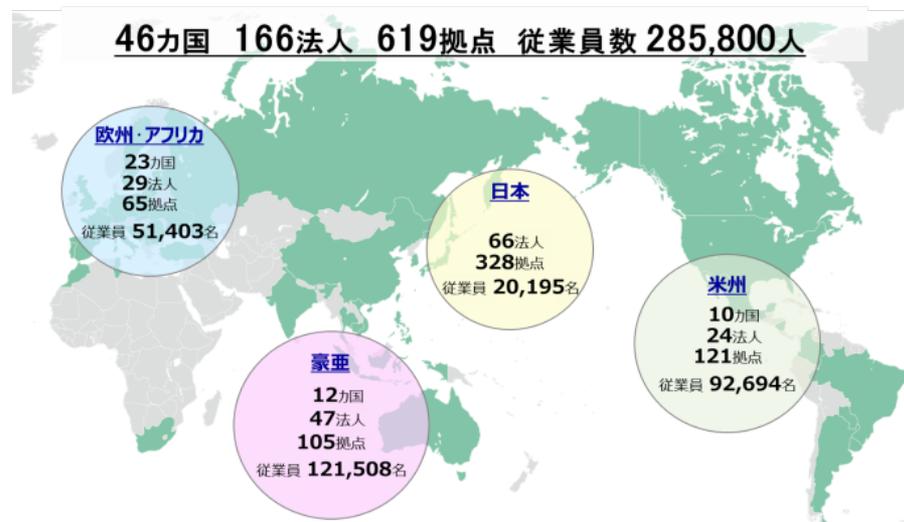
- グローバルな事業展開

## 事業

- 自動車関連
  - ワイヤーハーネス（世界トップクラスのシェア）
  - 自動車計器
  - 充電コネクタ [1,2]
  - 計装機器（ドライブレコーダーなど）
- 生活環境関連
  - 電線
  - ガス機器
  - 空調機器
  - 太陽熱利用機器
- 新規事業
  - 環境・リサイクル
  - 介護
  - 農業・食品



主に、新幹線の車内および静岡県内の新幹線の改札内に広告を掲載させて頂いております。



(2017年6月20日時点)

[1] 電気自動車 (Electric Vehicle, EV)

[2] プラグインハイブリッド車 (Plug-in Hybrid Electric Vehicle, PHEV)

# 自動車関連

## 最主力製品

- ワイヤーハーネス (Wire harness, 組み電線)

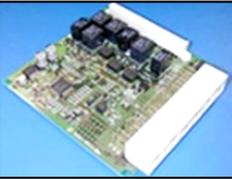
国内シェア= 約 50% [1]  
海外シェア= 約 30% [1]



メーター



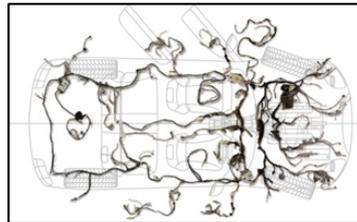
電源分配ボックス



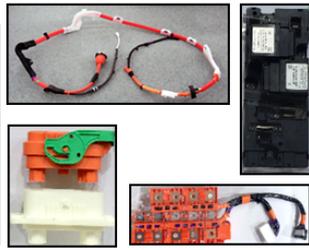
ECU



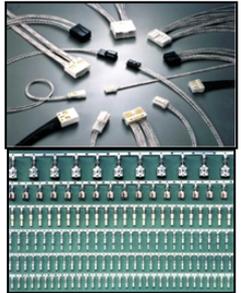
センサ



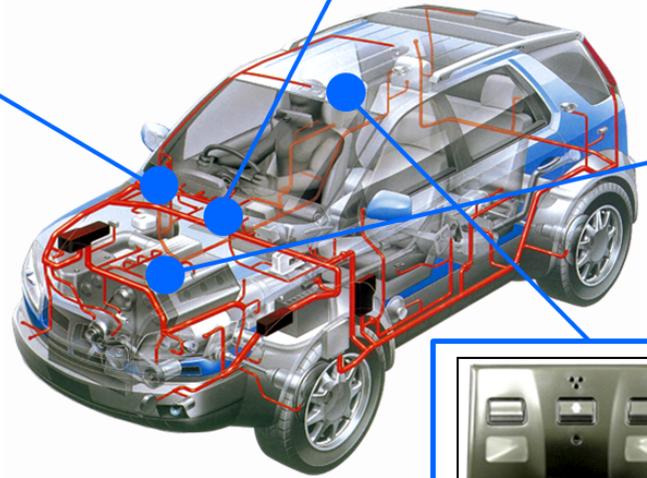
ワイヤーハーネス



高電圧  
ワイヤーハーネス部品



コネクタ、ターミナル



AC充電  
車両側インレット



室内ランプ



AC充電  
インフラ側コネクタ

[1] <http://wireharness.jp/feature/2017/03/31/157/>

弊社は、BtoB<sup>[1]</sup>を基本とするモノづくりメーカーです。

## 計装機器



タクシーメーター



デジタルタコグラフ



ドライブレコーダー

## 電線



住宅内の配線や  
電車の架線など  
企業向けに生  
産・販売

電線・ケーブル

## ガス機器



ガスメータ



L P ガス警報器

## 空調機器



吸収式冷温水器

オフィスビルや旅館・ホ  
テルなどの大規模建  
築向けの空調機器

[1] 企業間取引 (Business-to-business, BtoB, B2B)

# 一般市場向け事業

## 太陽熱



太陽熱温水器

## 農業・食品



ベビーリーフ



シャーベット



甘酒



施肥機



ゲル被覆種子

## 環境・リサイクル



紙リサイクル



ガラスリサイクル



トナーカートリッジリユース

## 介護



デイケアサービス「紙ふうせん」

一般市場向けもなくはないが、基本的に BtoB のモノづくりメーカーです。  
これまでは「質が良いモノ（製品）を作れば、売れる」という発想だった。

# モノ+サービスで注目した自社製品

弊社は、BtoB [1] を基本とするモノづくりメーカーです。

## 計装機器



タクシーメーター

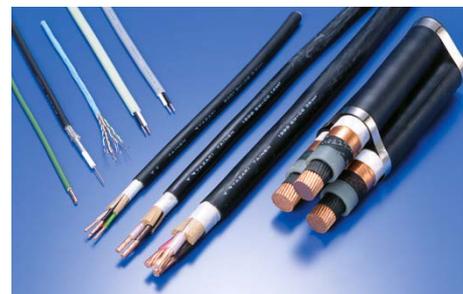


デジタルタコグラフ



ドライブレコーダー

## 電線



電線・ケーブル

住宅内の配線や  
電車の架線など  
企業向けに生  
産・販売

## ガス機器



ガスメータ



L P ガス警報器

注目した！

## 空調機器



吸収式冷温水器

オフィスビルや旅館・ホ  
テルなどの大規模建  
築向けの空調機器

[1] 企業間取引 (Business-to-business, BtoB, B2B)

# タクシーメーターの（弊社の）危機

## スマートメーター（スマートタクシーメーター）

タクシーメーター本体部とスマートフォンやタブレットなどのデバイスを Bluetooth などで接続したスマートタクシーメーターの新規参入が盛んである。



スピーディで効率的な配車 [1]

タクシー配車管理システム Smart to TAXI [1]

弊社の国内シェアは約 51% (約 123,000 台に装備) だが

スマートフォンの脅威にさらされている、と言える。

モノづくりメーカーも、「質の良いモノづくり」に加え、「ユーザー目線」のサービスを提供しなければ、お客様から見向きされなくなる。



タクシー配車システム SMART [2]

メーター連携、動態管理、配車、ナビの全てをスマートフォン1台で。

[1] 株式会社エコー・システム, [http://www.echosystem.co.jp/products/smart\\_taxi/](http://www.echosystem.co.jp/products/smart_taxi/)  
 [2] リアライズ・モバイル・コミュニケーションズ株式会社, [https://www.realize-mobile.co.jp/case/fuji\\_smart/](https://www.realize-mobile.co.jp/case/fuji_smart/)

# タコグラフ

## 商用車の法定三要素

- ・ 時間
- ・ 距離
- ・ 速度

乗務員管理のため、この三要素の測定が**義務付け**られている。

タコグラフ (運行記録計)

年	義務対象
1962	貸切バス, 路線バス (片道 100 km 超), 路線トラック
1967	総重量 8 t 以上 and/or 最大積載量 5 t 以上のトラック, 国内 15 都市のハイヤー・タクシー
1990	特別積合せ貨物運送 (宅配便など)
2015	車両総重量 7 t 以上 and/or 最大積載量 4 t 以上のトラック



アナログタコグラフ 大型



デジタルタコグラフ DTG7

国内シェア  
アナログ = 100%  
デジタル = 約 50%

国内合計で約 240,000 台  
に装備されている。

# ドライブレコーダー

## ドライブレコーダー

- 事故前後の映像を含む情報を記録する装置。

近年、需要が増している。

弊社製ドライブレコーダーは、国内合計で約 130,000 台に装備されている。



デジタルタコグラフ・ドライブレコーダー一体型 DTG7

ドライブレコーダー単体製品も販売しています。

## 弊社ドライブレコーダーの特徴

- 信頼性 (頑強)
  - 常時記録
  - 高価
  - サイズが大きい
- 事故時に壊れないで、問題なく記録することができます。

スマートフォンのドライブレコーダーアプリなど続々と新規参入が増えている。

しかも、ネットを活用した便利なサービス付きアプリもある。

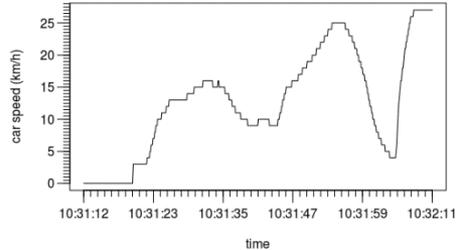


「記録するだけ」から、AI を活用して、「計測データを使って、お客様にサービスを提供する」へ。

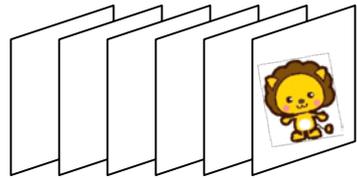
# 我々が求める AI

## 我々が求める AI と、その基盤

- a. 分析したいデータを、ネットワークを經由して、時系列データとして集められる。
- b. 複数の 1 次元の時系列データやカメラ撮影データなど多次元の時系列データを取り扱うことができる。
- c. 時系列データを入力すると、解を出力する。
- d. 最適解を、ネットワークを經由して、展開することができる。



1次元の時系列データ



動画は (m × n) 次元の時系列データ

## AI 作りで注目したドライブレコーダーが計測するデータ

- ① 動画 (前方)
- ② 自動車速度
- ③ 自動車加速度

- GPS センサーを接続すれば、GPS 情報も計測することができる。
- 車内カメラで人を撮影できるだけでなく、前方カメラでも人が撮影される。
- 前方カメラにより、前方の車も撮影される。

GPS 情報や人の外見、ナンバープレートなどは個人情報であり、その情報を (特に欧州で) 保存することはリスク<sup>[1]</sup>を生む可能性がある。

現状、なるべく個人情報は取り扱わないで提供することができるサービスを検討している。

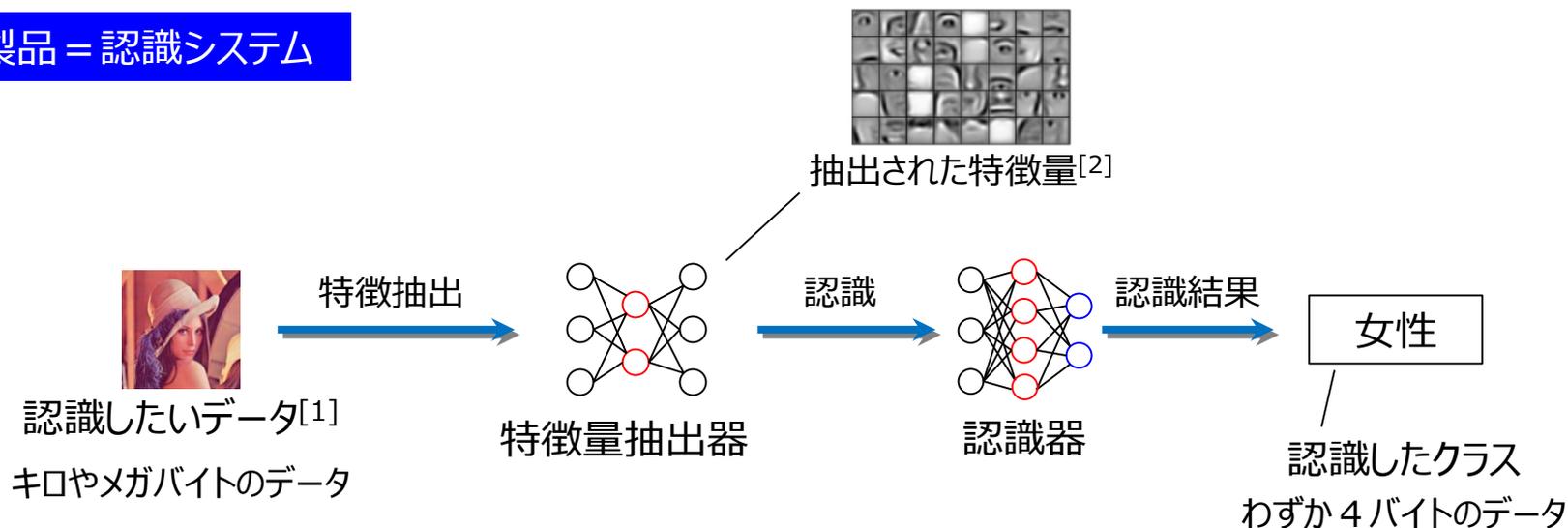
[1] 2018年5月25日より、EU 一般データ保護規則 (General Data Protection Regulation, GDPR) の本格適用が始まり、顕在リスクになった。

# AI の理解

## 我々が考える AI とは

- 計測したデータを自分達が求める形に情報量を低下させる変換器である。
- 画像などクラス分類 (Classification) の AI の中には, 「特徴量抽出器」と「認識器」で構成されるタイプがある。

製品 = 認識システム



どうやって目的通りに変換する「特徴量抽出器」と「認識器」を作るのか？

学習

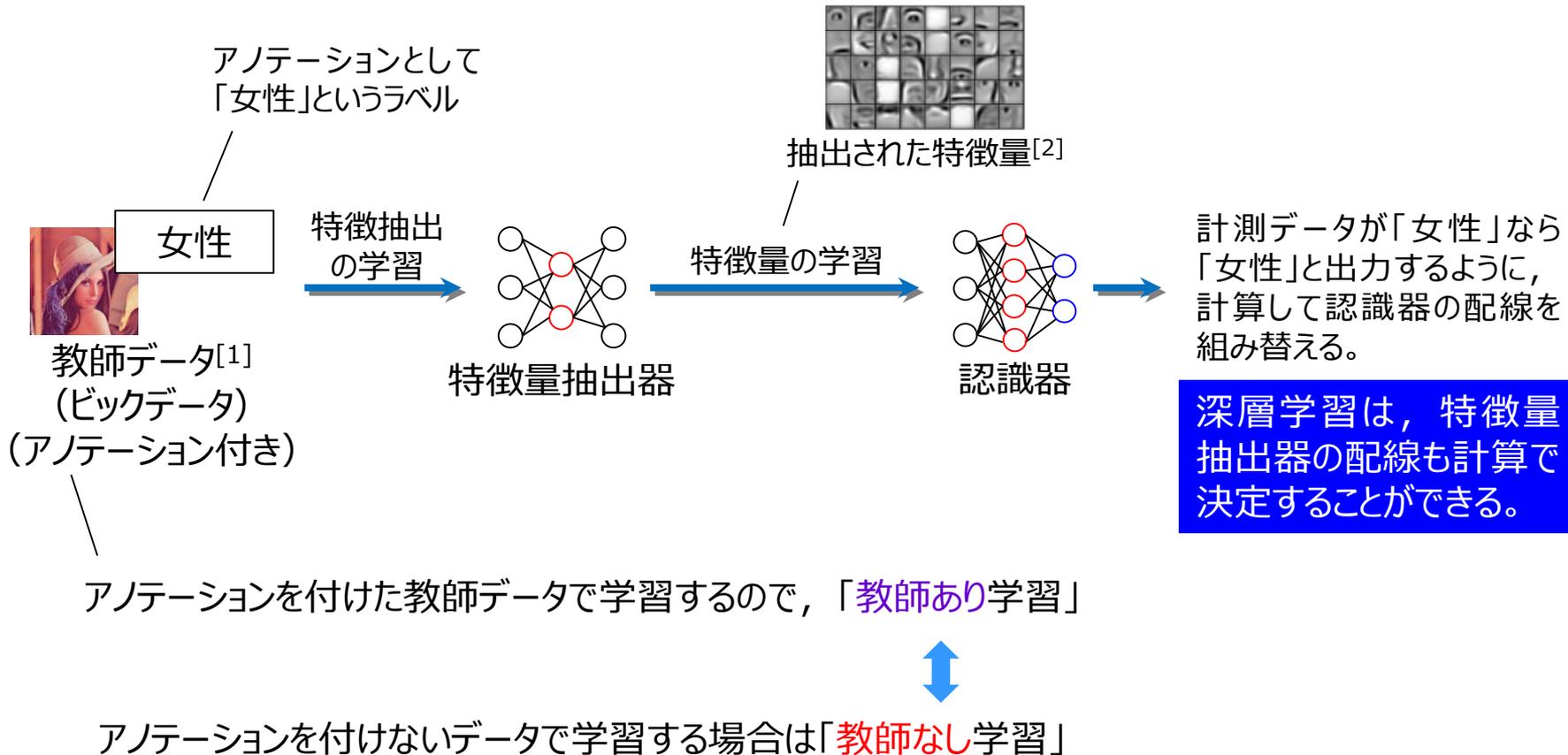
[1] Lenna, <http://www.lenna.org/>

[2] Honglak Lee, Roger Grosse, Rajesh Ranganath, Andrew Y. Ng, "Unsupervised learning of hierarchical representations with convolutional deep belief networks.", Communications of the ACM, Vol. 54, No. 10, pp. 95-103, 2011.

# AI の学習

## 我々が求める AI の学習とは

変換器を成す回路の配線の抜き差しや太さをコンピューターが計算して、決定する。



[1] Lenna, <http://www.lenna.org/>

[2] Honglak Lee, Roger Grosse, Rajesh Ranganath, Andrew Y. Ng, "Unsupervised learning of hierarchical representations with convolutional deep belief networks.", Communications of the ACM, Vol. 54, No. 10, pp. 95-103, 2011.

# AI 研究に使用する計算資源

## AI 計算資源の問題

- AI の計算は高負荷で、高価なコンピューター・リソース（資源）を必要とする。
- AI の研究途上では、どれだけの性能が必要なのか、わかりようがない。



所属部門では設備投資が**難しい**。特に、設備を追加することは極めて困難である。

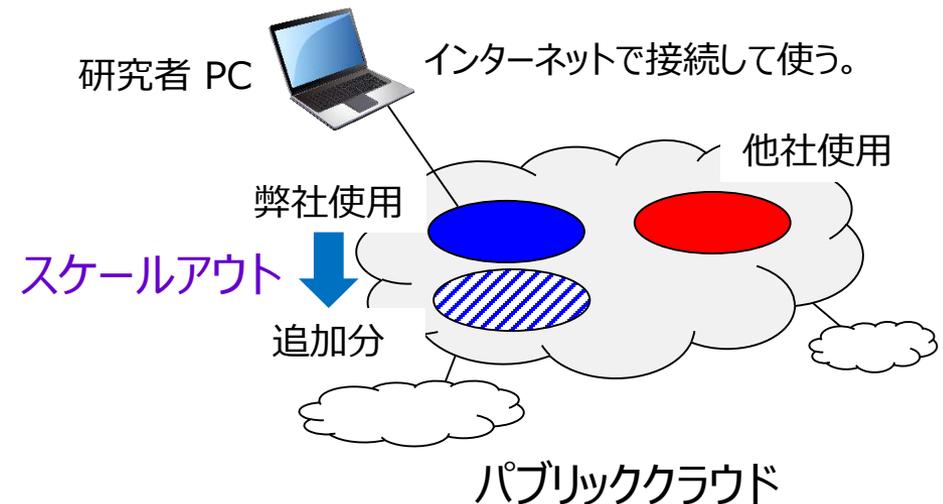
## オンプレミス<sup>[1]</sup>（スケールアップ）

性能の高い部品（CPU やメモリなど）を購入して更新するか、サーバーを追加購入する。



## パブリッククラウド（スケールアウト）

簡単に性能の高いサーバーに切り替えることやサーバーの使用数を増やすことができる。



[1] オンプレミス (on-premises), 自社運用, 基本的に買い切りでハードウェアを購入し, 構築する。

# クラウドの利用

## パブリッククラウドの利点

- パブリッククラウドは、スケールインも容易で、使わない時に解除することで費用を抑えられる。
- 数年後に陳腐化したハードウェアを持て余すことなく、常に最新のテクノロジーを利用することができる。

## GPU の利用

- CPU<sup>[1]</sup>ではなく、グラフィックを描画するプロセッサである GPU<sup>[2]</sup> を使って計算させる GPGPU<sup>[3]</sup> が普及している。
- AI 計算でも GPGPU が多用されている。

## 計算資源の見積もり

ストレージ	データベース	使用量 (月当たり)		ワークステーション (買い切り)	パブリッククラウド (月当たりの料金)	
		GPU	AI 計算		Amazon	Microsoft
1 TB	100 GB	8 GPU 320 時間	500 時間 2 百万件	~¥10,000,000 (別途, 諸経費)	¥350,097	¥248,884
1 PB	1 TB	8 GPU 320 時間	500 時間 2 百万件	数億円? (別途, 諸経費)	¥3,445,683	¥2,676,746

パブリッククラウドは、CPU および GPU のグレードが異なるので、一概に比較することはできない。

[1] 中央処理装置 (Central Processing Unit, CPU)

[2] 画像処理用演算プロセッサ (Graphics Processing Unit)

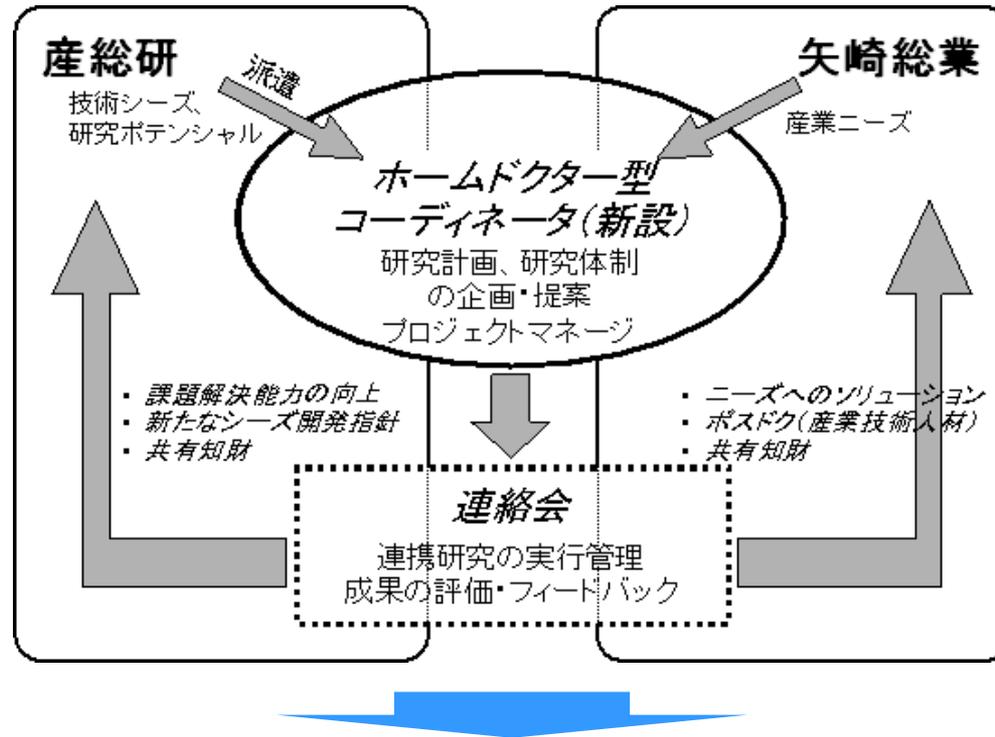
[3] GPU による汎用計算 (General-Purpose computing on Graphics Processing Units, GPGPU)

# 産総研様と共同研究

AI と縁のなかったモノづくりメーカーがゼロから AI に取り組むのは難しい。

## 協力協定

- 2008年7月24日、産総研様と弊社の間で、研究開発と人材育成を目的とする協力協定を締結<sup>[1]</sup>させて頂きました。



2017年1月より、産総研 人工知能研究センター様と共同研究をさせて頂いております。

[1] [http://www.aist.go.jp/aist\\_j/news/pr20080724.html](http://www.aist.go.jp/aist_j/news/pr20080724.html)

# 産総研様が開発された共同使用が可能な計算資源

計算資源	稼働開始	GPU	仮想マシン	コンテナ
ASGC [1]	2014年7月		✓	
Nairobi [2]	2016年6月	✓		
AAIC [3]	2017年6月	✓	(CPU インスタンスのみ)	✓ [5]
ABCI [4]	2018年8月 (予定)	✓		✓ [5]

- 仮想マシンやコンテナに対応していない計算資源では OS を共同で使用する。そのため、デバイスドライバーや画像処理ライブラリなどの導入は、管理者に適切に依頼しなければならない。共同使用の兼ね合いで要望に対応してもらえないこともある。
- 仮想マシンやコンテナを使うことができると、ある程度、自由にソフトウェアを導入することができるので、日々、更新され続けている AI ソフトウェアを試しやすい。

例えば、弊社ドライブレコーダーが計測した動画は FFMPEG デコーダーではデコードできるが、OpenCV 標準デコーダーではデコードできない。OpenCV 3.3 以前では、設定しても FFMPEG デコーダーを使えなかった。2017年12月23日にリリースされた OpenCV 3.4 から、設定すれば FFMPEG デコーダーを使うようにできる。

結局、FFMPEG ラッパーになるパッケージを使って、デコードしている。(Python では PyAV, R 言語では magick)

[1] AIST Super Green Cloud Portal Site (ASGC)  
 [2] Development of Next-Generation AI and Robot Core Technologies (NAIRobi)  
 [3] AIST Artificial Intelligence Cloud (AAIC)  
 [4] AI Bridging Cloud Infrastructure (ABCI)  
 [5] コンテナ間の MPI 通信が難しく、クラスターの対応が難しい Docker ではなく、HPC (High-Performance Computing,高性能計算) 向けに開発された Singularity を利用することができる。

# 世界最大級，超省電力 オープンAIインフラストラクチャ



Univ. Tokyo / Kashiwa Campus



- 世界トップレベルの計算処理能力とデータ処理能力
- AIとビッグデータのアルゴリズム，ソフトウェア，応用開発のための，オープンかつ専用の計算機システム
- 我が国における産学官共同のAI研究開発を加速するプラットフォーム

- 0.55 EFLOPS (HP), 37 PFLOPS (DP)
- 2.3 MW
- < 1.1 Avg. PUE



2018年8月運用開始



## スクラッチから超高密度データセンタを設計

- 低コスト, 軽量の「倉庫」を作り, その中に「やぐら」を組む二重構造を設計
- 通常のデータセンタの20倍の熱密度を実現

## 超省電力

- スパコン向けの冷却システムの民生利用
- 温水冷却と空冷の併用により冷却に要する電力を削減, ラックあたり70kWの冷却を可能に

## デファクトHW&SWアーキテクチャの採用

- デファクト&コモディティのHW, AI向けアクセラレータを採用
- 最先端のHPC・AIソフトウェア資産の活用, 迅速な技術移転を可能に

## ソフトウェアエコシステムの活用

- コンテナによるソフトウェア資産の蓄積・利用, グローバルコミュニティでの相互活用を促進

## データの安全な活用の促進

- マルチペタバイトスケールの高速ストレージ, データの共有・公開向けデータ基盤を提供
- 通信路とデータをクラウドのデファクト技術を使って暗号化



数百の研究機関・大学・企業の利用・協業,  
数千の研究者・エンジニアによる利用を促進

# 我々が求める AI に要求される要素

## 我々が求める AI 基盤の要素

データの用意	(A)	計測データの収集
	(B)	計測データのアノテーション追加 (教師データ)
	(C)	AI が読み取りやすい教師データセットの作成
学習	(D)	多次元の時系列データから解を出力する AI 作成
認識	(E)	AI を用いて, 計測データからお客様にサービスを提供

Amazon Mechanical Turk のようなアウトソーシングでは不特定多数にデータを公開して, アノテーションを追加してもらうことになる。

情報漏洩を強く懸念している。

## セキュアな教師データ作成環境の迅速な用意

- 手間を少なく, 迅速に環境を用意するため, オープンソースのアノテーションツール [1] を利用した。
- 簡単にセキュアな環境を構築することができるように, ローカルネットワーク内で運用できる環境を用意した。

## パブリッククラウド (AWS) 運用へ

- ① ストレージ (EBS, EFS) 暗号化
- ② グローバル IP アドレスの接続制限
- ③ 通信の暗号化 (SSL)
- ④ ユーザー認証 (自己署名証明書, OpenSSH)
- ⑤ ログ監視
- ⑥ クラウド API レベルのログ監視

情報漏洩を防ぐアノテーションツール利用環境を整備した。

[1] LabelMe ソースコード, <https://github.com/CSAILVision/LabelMeAnnotationTool>

# 教師データセット

## アノテーションツール LabelMe

- 静止画にアノテーションを作成する。
- ウェブサーバーアプリで、**多人数が同時に作業**できる。
- 矩形だけでなく、多角形（**ポリゴン**）でオブジェクトの位置の範囲を指定することができる。
- アノテーション情報は、xml ファイルとして自動的に保存する。

## 静止画ベース

- 動画を視聴し、オブジェクトが映っているフレームを静止画ファイルとして保存し、その静止画ファイルを LabelMe で開き、アノテーションを作成した。

## アノテーションを追加したオブジェクト

- ポリゴンで位置の範囲も追加したオブジェクト
  - 信号機（青，黄，赤も区別）
  - 最高速度標識（速度の値も区別）
  - 一時停止標識
- 分類ラベルだけを追加したオブジェクト
  - 一時停止標示（路面上のペイント）
  - 踏切



LabelMe を用いたアノテーション作成

# AI 作りの進め方

最初から多次元の時系列データを分析する AI 作りに取り掛かるのではなく、ステップを踏んで進める。

ステップ	要素
①, ②	2次元データのクラス分類
③, ④	2次元データのオブジェクト検出
⑤, ⑥	次元数が小さい多次元の時系列データ（センサーデータ）のパターン認識
⑦	次元数が大きい多次元の時系列データ（センサーデータ+動画）のパターン認識



静止画の抽出



①, ② クラス分類

クラス（速度標識 40 km/h など）を推定する。

① = 小サイズの静止画のクラス分類

② = 大サイズの静止画のクラス分類

③, ④ オブジェクト検出

画像の一部にあるオブジェクトのクラスだけでなく、オブジェクトの位置の範囲も推定する。

③ = 静止画, ④ = 動画

# AI 作りのプロセス

ステップ	項目	注釈	備考
①	静止画の クラス分類	小サイズの静止画	深層学習を用いた静止画クラス分類の基本として CNN [1] を利用する。
②		大サイズの静止画	定評のある AlexNet [2] を利用する。
③	2次元画像中の オブジェクト検出	静止画	深層学習の手法 [3,4] を利用する。
④		動画	
⑤	時系列データの パターン認識	次元数の小さい多次元 (複数の1次元)	音声認識などで定評のある GMM-HMM [5] を利用する。
⑥			深層学習の手法を用いた GMM-DNN [6] を利用する。
⑦		次元数の大きい多次元 (複数1次元+動画)	深層学習の手法である RNN [7] などを利用することを検討する。

深層学習フレームワークは、Chainer および MXNet を使用している。

(今後、AWS とマイクロソフトが共同開発した深層学習インターフェース Gluon を試す予定)

[1] 畳み込みニューラルネットワーク (Convolutional Neural Network)

[2] AlexNet, [http://vision.stanford.edu/teaching/cs231b\\_spring1415/slides/alexnet\\_tugce\\_kyunghee.pdf](http://vision.stanford.edu/teaching/cs231b_spring1415/slides/alexnet_tugce_kyunghee.pdf)

[3] YOLO v2 (You only look once, YOLO), <https://pjreddie.com/darknet/yolov2/>

[4] Single Shot MultiBox Detector (SSD), <https://arxiv.org/abs/1512.02325>

[5] 混合ガウスモデル - 隠れマルコフモデル (Gaussian Mixture Model - Hidden Markov Model, GMM-HMM)

[6] 混合ガウスモデル - 深層学習 (Gaussian Mixture Model - Deep Neural Network, GMM-DNN)

[7] 再帰型ニューラルネットワーク (Recurrent Neural Network, RNN)

MXNet を使用しているのは、講演者が R 言語に馴染んでいるため。

# ステップ ① 静止画クラス分類

## 3層 CNN を用いた二値分類

- 教師データの情報から該当部分を切り出し、画像解像度を 16X16 に変更した画像を学習する。
- 学習に使っていない 16X 16 画像を認識し、性能を評価する。

他の解像度も試した結果、選んだ値



画像解像度 16X16 の青信号画像

## 評価結果

判定 \ 実際	Positive	Negative
Positive	TP = 309.0	FP = 1900.0
Negative	FN = 0.0	TN = 1194.0

信号機などのオブジェクトを P とする。

TP = P を P と推定

TN = P ではないものを P ではないと推定

FP = P ではないものを P と推定 = 誤認識

FN = P を P ではないと推定 = **取りこぼし**

誤認識があっても良いが、取りこぼしはゼロにしたい課題。

## 性能

評価指標	計算結果
Recall (取りこぼし)	1.000
Precision (誤認識)	0.134
Accuracy (正確度)	0.442

**取りこぼしが無い!**

(その代わりに、誤認識が多い。)

$$Recall = \frac{TP}{TP + FN}$$

$$Precision = \frac{TP}{TP + FP}$$

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}$$

# ステップ ② 一般画像認識技術

## 目的

位置の範囲を特定することが難しく、動画から抽出した静止画にラベルを追加しただけの一時停止標示（路面標示）と踏切を定評のある AlexNet を利用して、識別することを試みた。

## 二値分類（踏切か一時停止線, それ以外）の結果

認識結果\実際	踏切か一時停止線	その他
正解	163	699
不正解	5 (取りこぼし)	1 (誤認識)

Precision (誤認識の評価)	0.994
Recall (取りこぼしの評価)	0.970
F 尺度 (総合評価)	0.982

$$F - measure = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall}$$

## 取りこぼし例

大きな歪み



近すぎが原因



遠すぎが原因

やむを得ないと思われる画像を取りこぼした。

# AI 作りのプロセス

ステップ	項目	注釈	備考
①	静止画の クラス分類	小サイズの静止画	深層学習を用いた静止画クラス分類の基本として CNN [1] を利用する。
②		大サイズの静止画	定評のある AlexNet [2] を利用する。
③	2次元画像中の オブジェクト検出	静止画	深層学習の手法 [3,4] を利用する。
④		動画	
⑤	時系列データの パターン認識	次元数の小さい多次元 (複数の1次元)	音声認識などで定評のある GMM-HMM [5] を利用する。
⑥			深層学習の手法を用いた GMM-DNN [6] を利用する。
⑦		次元数の大きい多次元 (複数1次元+動画)	深層学習の手法である RNN [7] などを利用することを検討する。

[1] 畳み込みニューラルネットワーク (Convolutional Neural Network)

[2] AlexNet, [http://vision.stanford.edu/teaching/cs231b\\_spring1415/slides/alexnet\\_tugce\\_kyunghee.pdf](http://vision.stanford.edu/teaching/cs231b_spring1415/slides/alexnet_tugce_kyunghee.pdf)

[3] YOLO v2 (You only look once, YOLO), <https://pjreddie.com/darknet/yolov2/>

[4] Single Shot MultiBox Detector (SSD), <https://arxiv.org/abs/1512.02325>

[5] 混合ガウスモデル - 隠れマルコフモデル (Gaussian Mixture Model - Hidden Markov Model, GMM-HMM)

[6] 混合ガウスモデル - 深層学習 (Gaussian Mixture Model - Deep Neural Network, GMM-DNN)

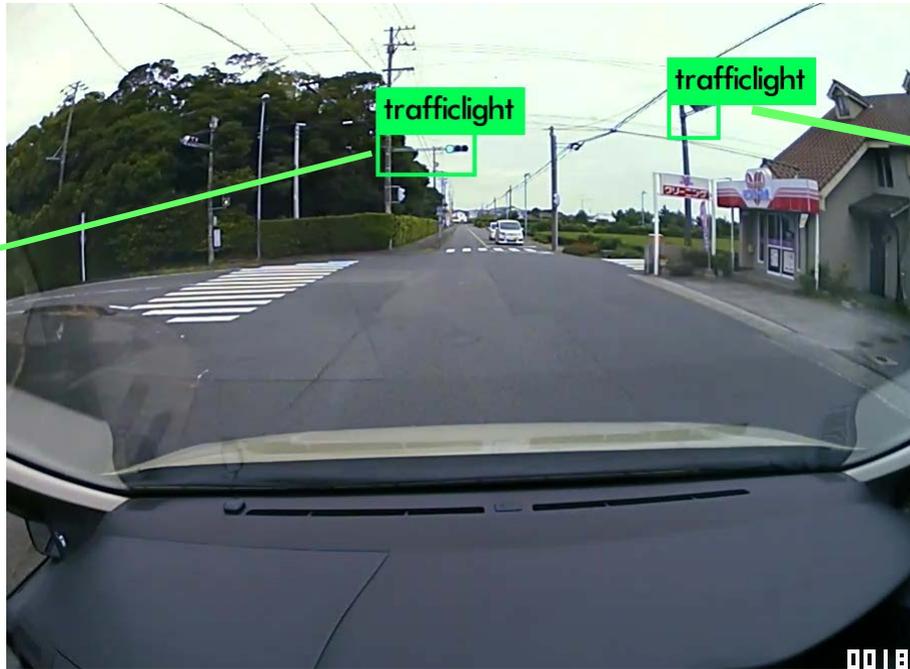
[7] 再帰型ニューラルネットワーク (Recurrent Neural Network, RNN)

# ステップ③ 静止画オブジェクト検出

## オブジェクト検出

- オブジェクトのクラスだけでなく、オブジェクトの位置も特定する。
- 不要な標識および標示の認識結果を取り除くためにオブジェクトの位置を知ることは重要である。

必要な認識  
結果



オブジェクト検出の結果の一例



横向きの信号機  
(不要な認識結果)

現在、動画の各フレームに対する「ステップ④ 動画オブジェクト検出」を進めている。

2018年2月27日リリースの OpenCV 3.4.1 では YOLO v3 を実行させた時、エラーが生じるらしいので、OpenCV 3.4.0 を使っている。

動画のデコードは、FFMPEG ラッパーとして使えるパッケージで行っている。

# AI 作りのプロセス

ステップ	項目	注釈	備考
①	静止画の クラス分類	小サイズの静止画	深層学習を用いた静止画クラス分類の基本として CNN [1] を利用する。
②		大サイズの静止画	定評のある AlexNet [2] を利用する。
③	2次元画像中の オブジェクト検出	静止画	深層学習の手法 [3,4] を利用する。
④		動画	
⑤	時系列データの パターン認識	次元数の小さい多次元 (複数の1次元)	音声認識などで定評のある GMM-HMM [5] を利用する。
⑥			深層学習の手法を用いた GMM-DNN [6] を利用する。
⑦		次元数の大きい多次元 (複数1次元+動画)	深層学習の手法である RNN [7] などを利用することを検討する。

[1] 畳み込みニューラルネットワーク (Convolutional Neural Network)

[2] AlexNet, [http://vision.stanford.edu/teaching/cs231b\\_spring1415/slides/alexnet\\_tugce\\_kyunghee.pdf](http://vision.stanford.edu/teaching/cs231b_spring1415/slides/alexnet_tugce_kyunghee.pdf)

[3] YOLO v2 (You only look once, YOLO), <https://pjreddie.com/darknet/yolov2/>

[4] Single Shot MultiBox Detector (SSD), <https://arxiv.org/abs/1512.02325>

[5] 混合ガウスモデル - 隠れマルコフモデル (Gaussian Mixture Model - Hidden Markov Model, GMM-HMM)

[6] 混合ガウスモデル - 深層学習 (Gaussian Mixture Model - Deep Neural Network, GMM-DNN)

[7] 再帰型ニューラルネットワーク (Recurrent Neural Network, RNN)

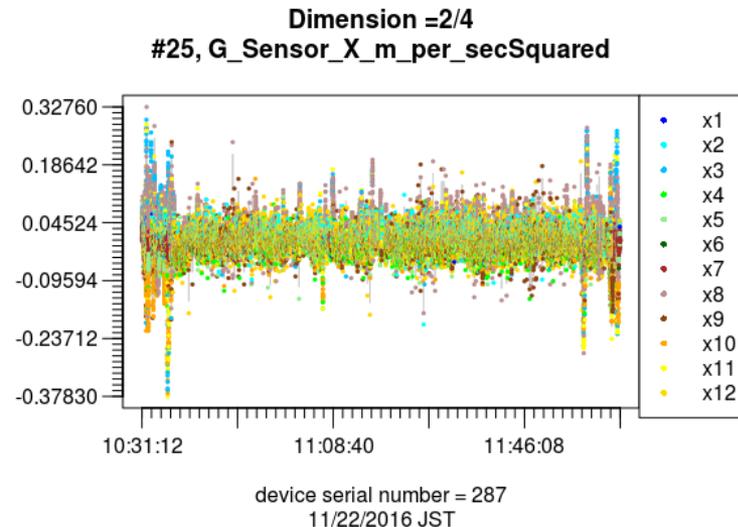
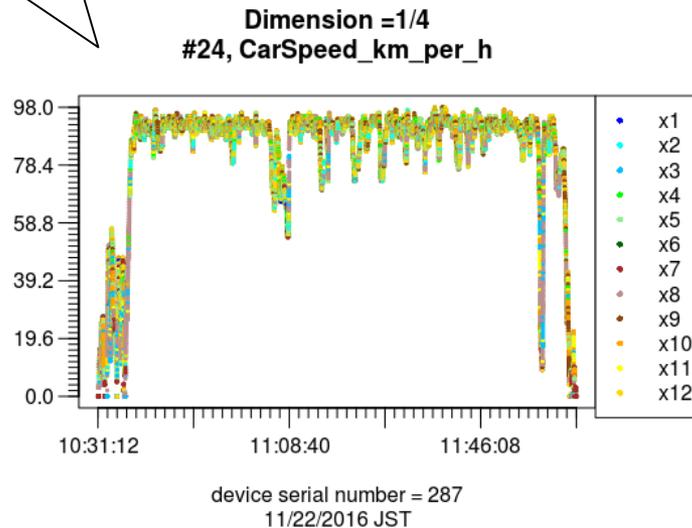
# ステップ ⑤ 時系列データのパターン認識

## GMM-HMM

- 教師あり学習が可能な手法であるが、教師データがないので、**教師なし学習**で進めた。
- 混合ガウスモデル (Gaussian Mixture Model, GMM) で観測できる物理状態を表す。
- 隠れマルコフモデル (Hidden Markov Model, HMM) で**隠れた状態**を推定する。

R 言語と RcppHMM  
パッケージを使って計算

e.g.  
隠れた状態 = もうすぐ雨が降る  
観測できる状態 = 温度や湿度など



GMM-HMM 教師なし学習モデルを用いたパターン認識の結果

次元数の大きくない多次元の時系列データのパターンを認識することができた。

# 時系列データのクラス分類

## 教師なし学習の問題

隠れ状態を推定 (estimation) することに成功したが、**教師なし**で学習したため、それが何のパターンを意味するのか、そのクラスを分類 (classification) することができない。

- パターン時系列データの表示と同時に動画を表示させるツールを開発した。
- パターン時系列データと動画の推移を**人手**で確認し、パターンのクラスを調べた。

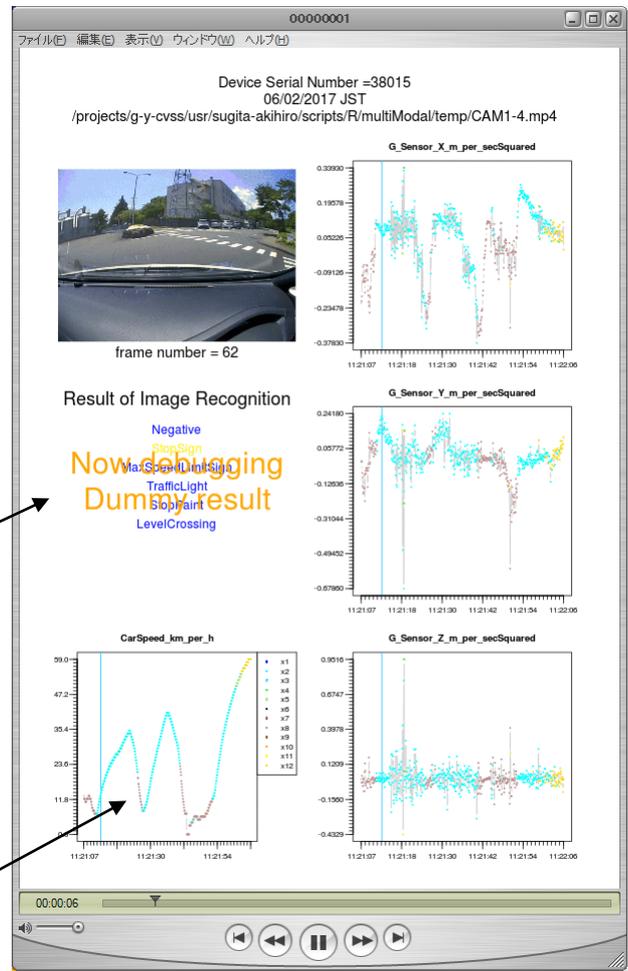
結果を別の動画で報告

次元数の小さい多次元の時系列データを使った AI で運転パターンを推定することができた。

パテント出願

ステップ ⑥, ⑦  
を見越して開発  
した部分

茶色パターン  
が右左折



ドライブレコーダー動画とパターンの同時再生ツールが出力した動画

# AI 作りのプロセス

ステップ	項目	注釈	備考
①	静止画の クラス分類	小サイズの静止画	深層学習を用いた静止画クラス分類の基本として CNN [1] を利用する。
②		大サイズの静止画	定評のある AlexNet [2] を利用する。
③	2次元画像中の オブジェクト検出	静止画	深層学習の手法 [3,4] を利用する。
④		動画	
⑤	時系列データの パターン認識	次元数の小さい多次元 (複数の1次元)	音声認識などで定評のある GMM-HMM [5] を利用する。
⑥			深層学習の手法を用いた GMM-DNN [6] を利用する。
⑦		次元数の大きい多次元 (複数1次元+動画)	深層学習の手法である RNN [7] などを利用することを検討する。

[1] 畳み込みニューラルネットワーク (Convolutional Neural Network)

[2] AlexNet, [http://vision.stanford.edu/teaching/cs231b\\_spring1415/slides/alexnet\\_tugce\\_kyunghee.pdf](http://vision.stanford.edu/teaching/cs231b_spring1415/slides/alexnet_tugce_kyunghee.pdf)

[3] YOLO v2 (You only look once, YOLO), <https://pjreddie.com/darknet/yolov2/>

[4] Single Shot MultiBox Detector (SSD), <https://arxiv.org/abs/1512.02325>

[5] 混合ガウスモデル - 隠れマルコフモデル (Gaussian Mixture Model - Hidden Markov Model, GMM-HMM)

[6] 混合ガウスモデル - 深層学習 (Gaussian Mixture Model - Deep Neural Network, GMM-DNN)

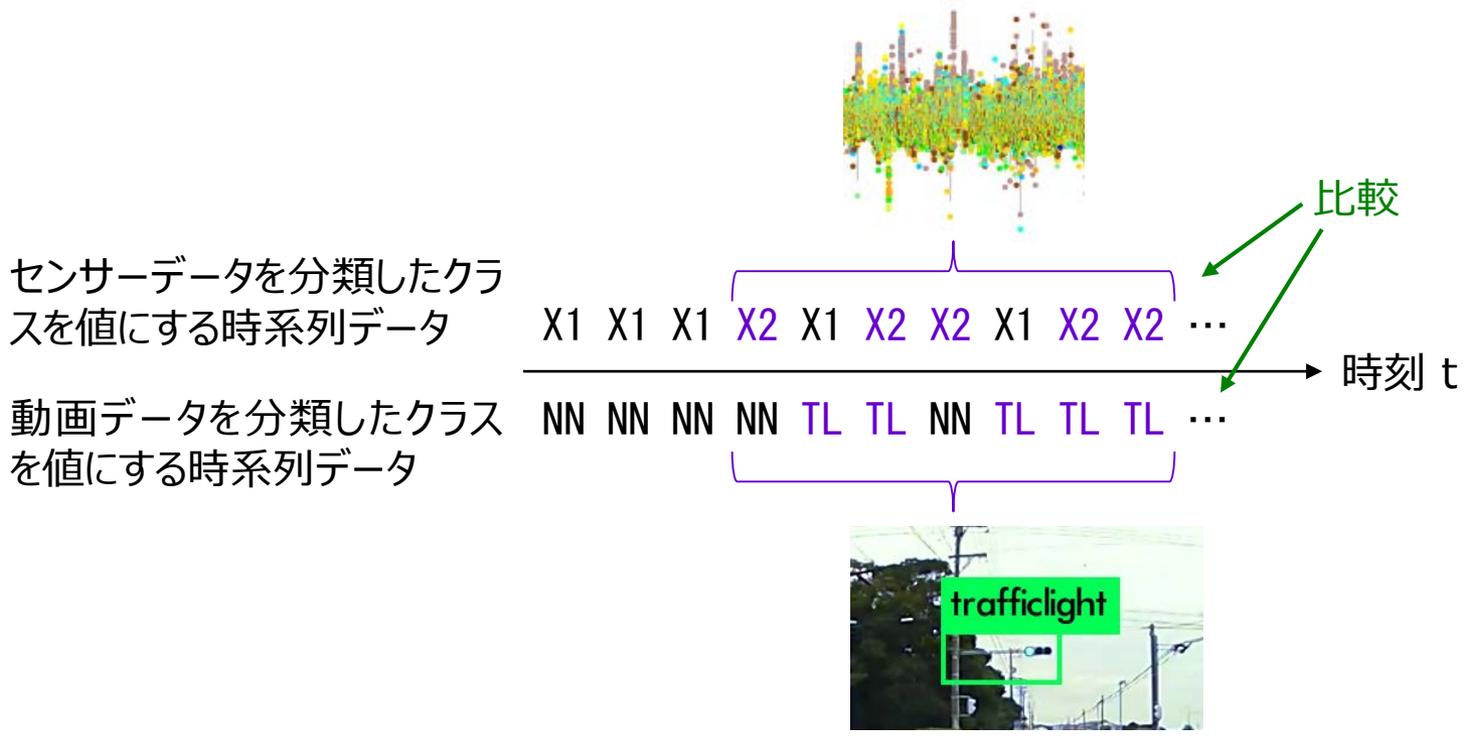
[7] 再帰型ニューラルネットワーク (Recurrent Neural Network, RNN)

# ⑥ ハイブリッド時系列パターン認識

## センサーデータと動画データのハイブリッド時系列パターン認識

2種類の時系列データを GMM-DNN モデルの教師なし学習で学習する。

- i. センサーデータの時系列データまたはセンサーデータを分類したクラスを値にする時系列データ
- ii. 動画データの各フレームを画像認識することにより分類したクラスを値にする時系列データ



運転操作のパターンを検出できる可能性がある。

# 教師なし学習および情報不足の問題

## 教師なし学習ではクラス分類が難しい

- 教師なし学習では、教師がないので、人が求める形に変換するように変更することができない。例えば、飲酒運転の運転パターンを見つけたい、と考えても、そのように変更することはできない。
- 意図したクラスで推定できるようにするためには、意図するクラスの教師データが必要である。

## 情報の不足による教師あり学習の問題

- 教師データがあっても、クラス分類で動画の次元を著しく低下させたデータは、速度や加速度などのセンサーデータと相関する情報を失っているかもしれない。



今後、ステップ ⑦ 教師あり学習で、（動画を含む）次元数の大きい多次元の時系列データのクラス分類に取り組んでいく。

これまでの次元 = 1 次元 × 4 種類 × 時刻 = 4 × 時刻

検討していく次元 = [(M × N) + 4] × 時刻

動画 1 フレーム  
の画面解像度



今後

# AI 作りのプロセス

ステップ	項目	注釈	備考
①	静止画の クラス分類	小サイズの静止画	深層学習を用いた静止画クラス分類の基本として CNN [1] を利用する。
②		大サイズの静止画	定評のある AlexNet [2] を利用する。
③	2次元画像中の オブジェクト検出	静止画	深層学習の手法 [3,4] を利用する。
④		動画	
⑤	時系列データの パターン認識	次元数の小さい多次元 (複数の1次元)	音声認識などで定評のある GMM-HMM [5] を利用する。
⑥			深層学習の手法を用いた GMM-DNN [6] を利用する。
⑦		次元数の大きい多次元 (複数1次元+動画)	深層学習の手法である RNN [7] などを利用することを検討する。

[1] 畳み込みニューラルネットワーク (Convolutional Neural Network)

[2] AlexNet, [http://vision.stanford.edu/teaching/cs231b\\_spring1415/slides/alexnet\\_tugce\\_kyunghee.pdf](http://vision.stanford.edu/teaching/cs231b_spring1415/slides/alexnet_tugce_kyunghee.pdf)

[3] YOLO v2 (You only look once, YOLO), <https://pjreddie.com/darknet/yolov2/>

[4] Single Shot MultiBox Detector (SSD), <https://arxiv.org/abs/1512.02325>

[5] 混合ガウスモデル - 隠れマルコフモデル (Gaussian Mixture Model - Hidden Markov Model, GMM-HMM)

[6] 混合ガウスモデル - 深層学習 (Gaussian Mixture Model - Deep Neural Network, GMM-DNN)

[7] 再帰型ニューラルネットワーク (Recurrent Neural Network, RNN)

# 最後に

- 質の良いモノを作れば売れる、と思い込んでいたモノづくりメーカーも、サービスを提供しなければ、お客様に目を向けてもらえない時代である。
- 打開策を生む技術として、我々は、非常に次元数の大きい多次元の時系列データから解を導出することができる AI を追求する。
- モノづくりメーカーに勤める AI 初心者は、計算資源から考えなければならなかった。
- パブリッククラウドは AI の計算資源として非常に高い性能を持っている。
- 産総研様は、オープン AI インフラストラクチャーを標榜する ABCI<sup>[1]</sup> など計算資源を含め、オープンイノベーションの受け皿となってくださる。
- 少なくとも、静止画や動画のオブジェクトを認識する AI を作るためには、計測データを数多く集めるとともに、十分な量と質の教師データを用意することが重要である。
- 信号機、最高速度標識、一時停止標識、一時停止標示（一時停止線の路面標示）、踏切の教師データセットを作った。
- 近年の AI は、初心者でも進められるくらい、開発環境が整っており、既存のフレームワークおよびパッケージを利用して、ドライブレコーダーが計測した動画から抽出した静止画のクラス分類およびオブジェクト検出を行い、速度および加速度の時系列データから運転パターンを推定することができた。
- 今後、極めて次元数の大きい多次元の時系列データからパターンを認識する AI を作っていこう、と考えている。

研究パートナーおよびビジネスパートナーを募集しています。

[1] AI Bridging Cloud Infrastructure (ABCI)