

## 映像解析と深層学習による鉄道の閉扉判断 AI の研究

長澤 史記 田内 葉子 小平 孝之 奥村 誠司 菅谷 元典 (三菱電機)  
星子 遼 小澤 圭広 松田 卓也 早川 裕基 松本 耕輔 (東京地下鉄)

Research on door-closing control AI using video analysis and deep learning

Fuminori Nagasawa, Yoko Tanouchi, Takayuki Kodaira, Seiji Okumura, Motonori Sugaya, (Mitsubishi Electric)

Ryo Hoshiko, Yoshihiro Ozawa, Takuya Matsuda, Hiroki Hayakawa, Kosuke Matsumoto, (Tokyo Metro)

We propose a "door-closing control AI" that analyzes camera images for checking boarding and alighting at station platforms and determines the optimal door-closing timing. The door-closing control must take into account the departure time, but also avoid door pinching of passengers. To realize this task as a system, we developed a method that combines deep learning and optical flow analysis. The results of cross-validation showed that the accuracy was over 90% at the station where the training data was collected. In the future, this method will be used to reduce the workload of door-closing control making by train crews.

キーワード：機械学習、深層学習、映像解析、オプティカルフロー

(Machine Learning, Deep Learning, Video Analysis, Optical Flow Analysis)

### 1. はじめに

国内労働人口の減少や交通需要の変化により、様々な鉄道事業者で運転士一人のみが乗務するワンマン運転の導入検討が進んでいる。ワンマン運転では従来車掌が実施していた業務を運転士が担うため、運転士の業務負荷が増加する。このような環境で列車の定時運行を維持するためには、従来車掌が担ってきた業務を自動化して運転士の業務負荷を低減する事が求められる。

負荷が高い業務の一つとして、従来車掌が行ってきた閉扉判断業務がある。これは出発時刻を考慮しつつ、旅客の車両乗降状況からホームドアおよび車両ドアを閉じるタイミングを判断する業務である。

東京地下鉄と三菱電機では、運転士の負荷低減を目指し、閉扉判断業務を自動化する閉扉判断AIの開発に取り組んでいる。本稿では、ホームの乗降確認用カメラ(以下カメラ)の映像から閉扉の適切なタイミングを判断する映像解析技術

について解説する。

### 2. 閉扉判断業務自動化の課題

閉扉判断業務では、鉄道の定時運行を妨げないように出発時刻を考慮しつつ、旅客がドアに挟まれる等の事故を回避する必要がある。

現状、閉扉タイミングは運転士や車掌がホームのカメラ映像を目視確認して、旅客の乗降意図を予測しつつ、人流の切れ目や駆け込み乗車の有無、混雑状況などを考慮して総合的に判断されている。具体的には、車両近くに旅客がいる場合でも降車する方向に移動する場合はドアを閉めることができる一方で、車両から離れた位置にいる旅客でも、乗車方向に移動している場合は戸挟みの危険があるため、ドアを閉めることができない。このため、閉扉判断AIにおいても、旅客の位置と動きの両者を考慮した閉扉タイミングを判断することが技術課題となる。

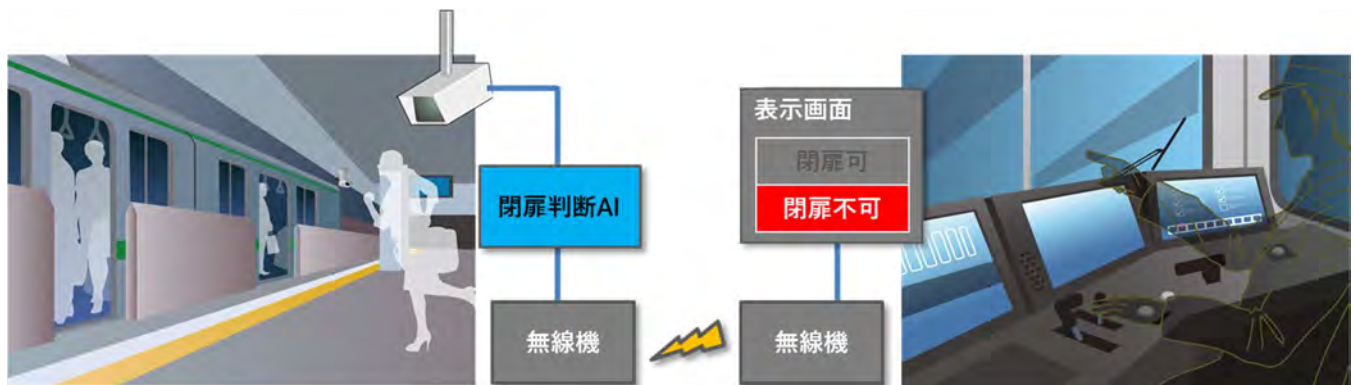


図 1：閉扉判断 AI の動作イメージ

また、閉扉判断 AI を実際に導入する場合には精度を高く保つために、導入するすべての駅で個別に AI の学習を行う事が望ましい。しかし、AI の学習には駅での映像収集やそのラベリングといった学習データの作成が必要となる。これらの作業は導入コストやリードタイムに大きく関わってくるため、導入するすべての駅の学習データを収集することなく、限られた駅のデータで学習しても性能を発揮できるように汎化性を高める技術の開発も課題である。

### 3.閉扉判断 AI

#### (3.1) 閉扉判断 AI の概要

本研究で開発した閉扉判断 AI の処理の概要を説明する。閉扉判断 AI はホーム上のカメラの映像を入力として、カメラごとに閉扉の適切なタイミングを判断する。(図 1 参照)。カメラ映像中に乗降中の旅客がいな場合は閉扉可とし、すべてのドアで閉扉可と判断された場合に、車両全体で閉扉可と判定する。

閉扉判断 AI は映像入力用のカメラと、2 つの映像解析処理 (3D-CNN による閉扉可否判断・動きベクトルを用いた旅客移動分析)、および映像解析処理の結果を統合する解析結果統合処理で構成する (図 2 参照)。

本論文では、今回開発した閉扉判断 AI の 2 つの映像解析処理 (3D-CNN による閉扉可否判断・動きベクトルを用いた旅客移動分析) についてそれぞれ説明する。

#### (3.2) 3D-CNN による閉扉可否判断

本機能ではカメラ映像から旅客の位置と動きを解析し、乗降動作がみられる場合は閉扉不可と判定する。

映像上の旅客の位置と動きの両方を考慮した解析が行えるように、本研究では閉扉可否判断モデルとして位置特徴と時系列特徴を同時に学習可能な深層学習モデルである 3D-CNN [1] を使用し、モデルの学習には汎化性能を高める効果のある AdaCos [2] を適用した。

#### (3.3) 動きベクトルを用いた旅客移動分析

本機能では、ホームの奥行きを考慮して旅客移動方向を判定するための方法として、極座標に基づいた旅客移動分析を行う。

カメラ映像では手前から奥へ延びる平行線が一点へ収束するように見えるため、カメラと旅客の位置関係によって、遠方で動きがあったときに動きの量が小さく見える性質と、カメラ光軸に平行な移動が斜め移動に見える性質がある。奥行き方向に広い駅のホームの映像ではこの性質が特に強く表れる。

3D-CNN では旅客の位置と動きの特徴を正確に検出可能だが、映像での見かけ上の動きを分析するため、カメラから遠く見かけ上の動きが小さい旅客の乗降動作を見落とす場合や、カメラの近くで静止した旅客の体の揺れを乗降動作と誤検知する問題がある。このため、動きベクトルを用いた旅客移動分析においては(I)距離による動きの大きさ変化に頑健な判定、(II)位置関係による角度変化に頑健な移動方向判定が課題となる。

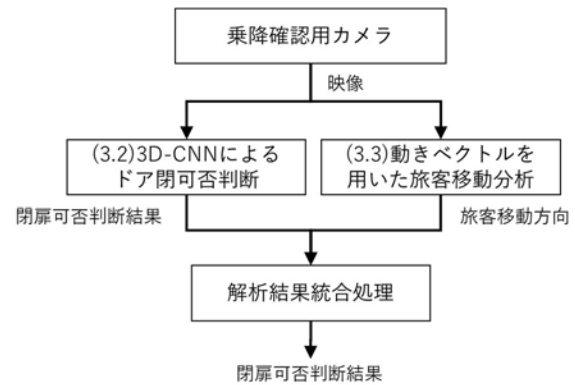
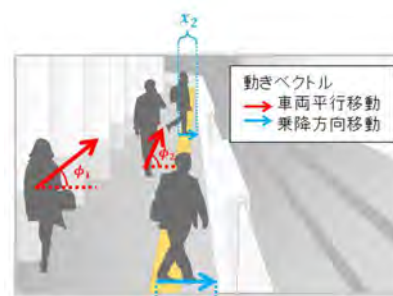
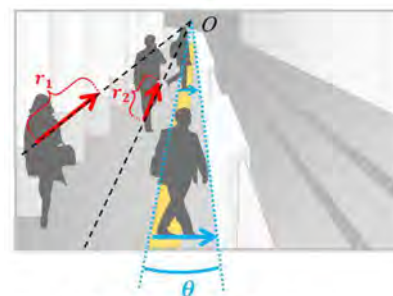


図 2: 閉扉判断 AI の構成



同じ方向の移動でも、カメラとの位置関係によって動きベクトルの角度や大きさが変化する。



ホーム最奥を原点とする極座標に変換し、 $\theta$ 成分(原点周りを回る方向の長さ)を用いて移動方向を判定する。

図 3: 極座標に基づく移動方向判定の概要

この 2 つの課題を解決するために、本機能では極座標に基づいた動きベクトル分析により旅客移動方向を判定する。カメラ映像の最新 2 フレーム間の動きベクトルを抽出し、ホーム上に車両乗降方向移動を行う旅客がいるか判定する。

極座標に基づく旅客移動分析では、動きベクトルを極座標成分で比較することによりホーム上の旅客の移動方向を分類する。具体的な処理としては、画像中の見かけの人流から抽出された各動きベクトルを動きの起点と終点の情報に基づきホーム最奥を原点とする極座標ベクトル( $\theta, r$ )に変換する。 $\theta$ 成分が最も大きなベクトルを代表ベクトルとして選択し、代表ベクトルの $\theta$ 成分が閾値以上の場合に「乗降方向移動」とし閾値未満の場合に「車両平行移動」とする(図 3 参照)。

4. 実験

今回開発した 2 つの映像解析処理技術の性能評価のため、実際に東京地下鉄の駅の乗降確認用カメラから映像を収集して精度評価を実施した。まず 3D-CNN による閉扉可否判断精度の評価を行い、次に動きベクトルを用いた旅客移動分析についての評価を行った。

(4.1) 使用した駅映像の概要

精度評価に用いる映像として、駅タイプ (ホーム形状[直線/曲線]・駅設置場所[地上/地下]) が類似する 5 駅と、前者とは駅タイプが異なる 2 駅の映像を収集した。映像収集を行った駅と各駅の駅タイプの分類を表 2 に示す。

映像は駅ホーム毎に 3~4 台分のカメラ映像を 1 週間分収集した。そして収集した映像から、車両が到着してから発車するまでを 1 シーンとして、朝夕のラッシュ時を含む様々な時間帯からシーンを切り出して使用した。切り出した映像のフレーム毎に、映像中の閉扉可否を人手によりラベル付けした。ラベルを付与したシーンの総数は約 750 シーンであり、約 130 万フレーム分の映像となった。

ラベル付与後のデータについて、学習に使用する教師データと、学習に使用せず評価に用いる試験データを分離した。試験データは駅毎に 30%のデータを無作為抽出した。

(4.2) 実験手順

4.2.1. 3D-CNN による閉扉可否判断

3D-CNN による閉扉可否判断精度を交差検定により評価した。分類モデルを学習する際は、駅タイプが共通する 5 駅のうち 1 駅を教師データから除外して学習した分類モデルを除外する駅ごとに作成した。評価の際は、各駅の試験データに対する人と AI による閉扉可否判断の一致率を精度 (accuracy)として式(1)により求めた。

$$accuracy(\%) = \frac{\text{閉扉可否正解フレーム数}}{\text{評価に使用した全フレーム数}} \dots(1)$$

閉扉可否正解フレーム数：AI の判定結果が人の付与したラベルと一致したフレーム数  
 評価に使用した全フレーム数：評価データ全フレーム数

また駅タイプが異なる駅での精度評価のため、学習駅と駅タイプが異なる 2 駅の試験データに対する各分類モデルの精度を求めた。

4.2.2. 動きベクトルを用いた旅客移動分析

A 駅のカメラ 1 台分の映像に対して、動きベクトル分析による人流移動方向分類精度の評価のために、乗降方向移動・車両平行移動する旅客が含まれると人がラベル付けしたフレームについて、正しく乗降方向移動・車両平行移動を判定できた割合として、乗降方向移動判定精度と車両平行移動判定精度を求めた。

$$\text{乗降方向移動判定精度}(\%) = \frac{\text{乗降方向移動判定正解フレーム数}}{\text{乗降方向移動の全フレーム数}} \dots(2)$$

乗降方向移動の全フレーム数：「乗降方向移動する乗客が含まれる」と人がラベル付けしたフレームの総数  
 乗降方向移動判定正解フレーム数：乗降方向移動ありと判定した結果が人の付与したラベルと一致したフレーム数

$$\text{車両平行移動判定精度}(\%) = \frac{\text{車両平行移動判定正解フレーム数}}{\text{車両平行移動の全フレーム数}} \dots(3)$$

車両平行移動の全フレーム数：「車両平行方向移動する乗客が含まれる」と人がラベル付けしたフレームの総数  
 車両平行移動判定正解フレーム数：車両平行移動ありと判定した結果が人の付与したラベルと一致したフレーム数

人流移動方向分類は閾値を基準とした 2 項分類であるため、2 つの精度 (乗降方向移動判定精度・車両平行移動判定精度) はトレードオフの関係にある。このため、分類性能の指標としては、2 つの精度で高い精度を両立できることが重要である。そこで、代表ベクトル  $\theta$  成分の閾値を変化させた際の 2 つの精度の関係を求めた。

また、開発した手法が課題(I)(II)の両方を解決できているか評価するための比較対象として、以下 2 つの手法での移動方向判定精度を同様に求めた。

- (i) 遠い動きが小さく見える性質の影響を受ける手法：  
 $x$  成分の大きさが閾値以上の動きベクトルが検出された場合に「乗降方向移動」とし、それ以外を「車両平行移動」とする。課題(I)を解決していない手法。

表 2：映像を収集した駅と駅タイプの対応

		駅設置場所	
		地下	地上
ホーム形状	直線	A, B, C, D, E	G
	曲線	F	

表 1：3D-CNN による閉扉可否判断精度評価結果

学習駅	駅毎の精度						
	地下駅					地上駅	
	ホーム形状直線					ホーム形状曲線	
	A	B	C	D	E	F	G
B+C+D+E	89.6%	94.4%	97.5%	97.7%	95.9%	93.4%	89.5%
A +C+D+E	93.2%	94.6%	97.9%	97.8%	96.2%	93.4%	92.2%
A+B +D+E	94.0%	96.4%	95.9%	97.6%	95.5%	88.4%	84.0%
A+B+C +E	94.0%	95.2%	92.0%	96.3%	96.5%	92.0%	77.0%
A+B+C+D	91.9%	95.4%	97.8%	97.7%	92.3%	85.0%	81.8%

(ii) 平行な移動が斜めに見える性質の影響を受ける手法：  
見かけ上最も横向きの動きベクトルの角度 $\phi$ が閾値未満の場合に「乗降方向移動」とし、閾値未満の場合には「車両平行移動」とする。課題(II)を解決していない手法。

## 5. 結果

4章で述べた実験の結果を示す。

### 5.1. 3D-CNN による閉扉可否判断

4.1 節の結果を表 1 に示す。表 1 の結果から、学習した駅では最大 97.9%の精度が得られ、学習に使用した駅での精度はいずれも 91%以上であった。学習に使用していなかった駅での精度では、駅タイプが同等の駅では、最も低い精度のものでも 89.6%と、90%近い精度が得られた。一方で、駅タイプが異なる駅(D, E)では精度が最も低い事例で 77.0%と、精度が低下する事例がみられた。

### 5.2. 動きベクトルを用いた旅客移動分析

提案手法および比較した 2 つの手法について、判定閾値を変化させた際の 2 つの精度（乗降方向移動判定精度・車両平行移動判定精度）の関係を図 4 に示す。2 つの精度はトレードオフの関係にあるため、両者の精度が同時に 100%となる理想的な分類モデルではグラフは右上(100%,100%)を通り、ランダムな結果を返すモデルではグラフは図の中央(50%,50%)を通る。

本研究で開発した極座標 $\theta$ 成分判定手法の精度変化を示すグラフは、(i)遠い動きが小さく見える性質の影響を受ける場合や(ii)平行な移動が斜めに見える性質の影響を受ける場合の精度変化を示すグラフと比較して右上(100%,100%)に近い点を通過した。この結果から、開発した手法が乗降方向移動判定精度と車両平行移動判定精度の両者を最も高い精度で両立できることがわかった。

## 6. 考察

表 1 の結果から、学習していない駅でも駅タイプが同様であれば 90%以上の精度で正しく閉扉可否を分類できた。一方で、駅タイプが異なる場合には分類精度が 77%程度まで低下したことから、高精度な閉扉判断のためには、駅タイプ毎の閉扉判断モデル作成が必要であると考えられる。

図 4 の結果から、(i)遠い動きが小さく見える性質の影響を受ける場合や(ii)平行な移動が斜めに見える性質の影響を受ける場合では精度が低くなった一方で、開発した手法では高い精度が得られた。この結果から、本研究で対象とする駅ホームのように奥行の大きい映像では、極座標に基づく移動方向判定手法が有効であると考えられる。

3D-CNN による閉扉可否判断の判断失敗事例を個別分析したところ、以下の傾向が散見された(図 5 参照)。

- (a) 旅客の移動方向(乗降方向移動/車両平行移動)の誤り
  - (b) カメラから遠方にいる旅客の乗降移動の見落とし
  - (c) カメラ近くの旅客の体の揺れを乗降移動と誤検知
- これらの点については、動きベクトル分析による移動方向分析によって高精度に判定可能である。このことから、3D-

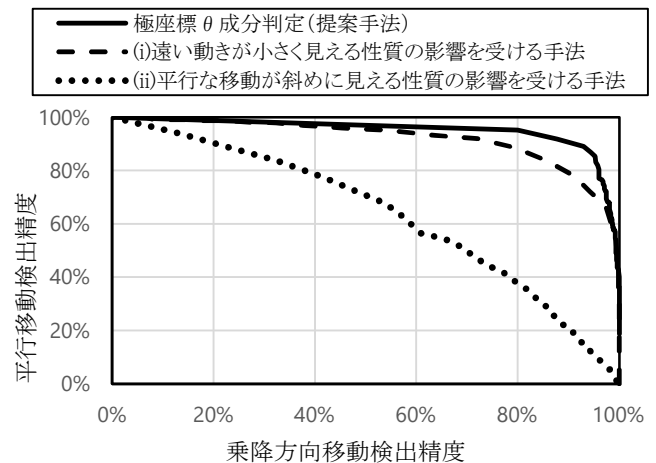


図 4：動きベクトル分析による人流移動分析  
閾値を変化させた際の人流方向判定精度



図 5：3D-CNN による判定誤り事例イメージ

CNN による閉扉可否判断結果と動きベクトルによる人流移動方向分析結果の統合により、高精度な閉扉判断が可能であることがわかった。

## 7. おわりに

本論文では、閉扉判断業務を自動化するための映像解析技術である閉扉判断 AI の実現に向けて開発した 2 つの映像解析処理(3D-CNN による閉扉可否判断・動きベクトルを用いた旅客移動分析)について説明した。

駅映像での評価を行い、3D-CNN による閉扉可否判断では学習を行った駅で最大 97.9%の精度で閉扉可否を正しく推定できたほか、極座標に基づく手法では遠い動きが小さく見える性質の影響を受ける場合や平行な移動が斜めに見える性質の影響を受ける場合を用いて判定した場合と比較して高い精度で人流移動方向を推定できた。

今後、本技術を用いた閉扉判断装置の実現に向けて、本論文で説明した 3D-CNN による閉扉可否判断結果と動きベクトル分析による人流移動方向分析結果を統合するためのアルゴリズム検討や、出発時刻を考慮した閉扉判断制御手法についての開発を行う予定である。

## 文 献

- [1] R. Chengping, L. Yang, "Three-dimensional convolutional neural network (3D-CNN) for heterogeneous material homogenization," *Computational Materials Science*, 第 卷 184, 2020.
- [2] Z. Xiao, Z. Rui, Q. Yu, W. Xiaogang and L. Hongsheng, "AdaCos: Adaptively Scaling Cosine Logits for Effectively Learning Deep Face Representations," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019.