

## 深層学習によるホテルの客室清掃時間推定手法の検討

大谷 紀子 研究室

2172103 松田 拓士

### 1. 背景と目的

ホテルの客室は、チェックアウトからチェックインまでの限られた時間内で清掃する必要がある。また、ホテルの経営において客室を高品質に保つことが重要である。したがって、客室清掃員はミスが許されないというプレッシャーを感じながら、時間に追われ清掃しなければならない。客室清掃の業務内容としては、ベッドメイク、客室全体の清掃、リネン類交換、アメニティ確認と交換・補充などがある。客室清掃の業務内容に対して、ホテルの客室清掃員として働いている人からは、「清掃作業が退屈である」という声が挙がっている。

プレッシャーや退屈さを解消する手段として、音楽を聞くことが考えられる。音楽は、ストレス、不安、うつ病の軽減に効果的であることが示されている。また、モチベーションを高め、退屈な作業を楽しくさせることも示されている[1]。したがって、清掃時間に合わせた楽曲を提供することで、客室清掃員の負担軽減が見込まれる。しかし、不要物の量やアメニティの交換の有無により客室ごとの清掃時間が異なるため、清掃時間に合わせた楽曲を提供することができない。以上のことから、清掃時間に合わせて音楽生成をするためには、客室清掃にかかる時間推定が必要である。

本研究では、清掃時間に合わせた音楽生成に活用することを目的とし、CNNに基づく客室清掃時間推定手法における不要物の数と固定物体の移動情報の必要性を検討する。

### 2. 清掃時間の推定手法

本研究では、ビジネスホテルの客室における机

上の片付けを対象とする。清掃前の画像データに対し、Convolutional Neural Network（以下 CNN）と物体検出モデル YOLO[2]を用いて時間を推定する。CNNは、画像認識や画像分類のために使用されるニューラルネットワークの一種であり、データに含まれる特徴を学習することで、分類や予測が可能である。YOLOは、リアルタイム物体検出に特化したモデルであり、犬、猫、車など一般的な物体を検出できる既存モデルが存在する。また、自分が検出させたい物体のデータをトレーニングさせ、特定の物体を検出できるオリジナルモデルを作成することが可能である。

ホテルの机上では、不要物や固定物体など、検出させたい対象が決まっているため、オリジナルモデルを用いる。固定物体は、ティッシュやリモコンなど、客室使用前から置いてある既存の物である。本研究では、固定物体としてリモコン、TV、時計、ティッシュを検出対象とする。オリジナルモデルによる物体検出例を図 1 に示す。

CNN と物体検出を用いた清掃時間の推定手順

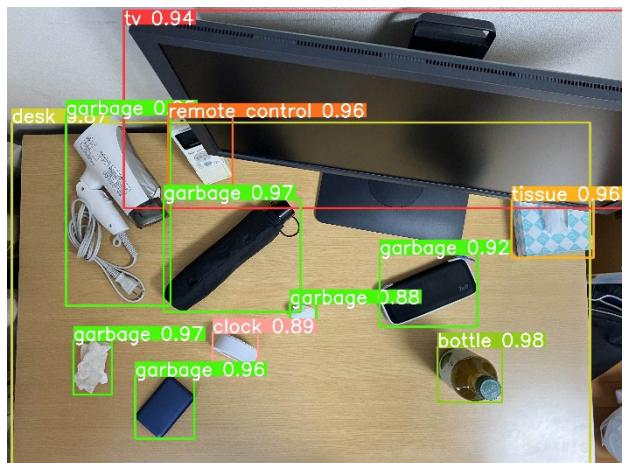


図 1 : YOLO による物体検出

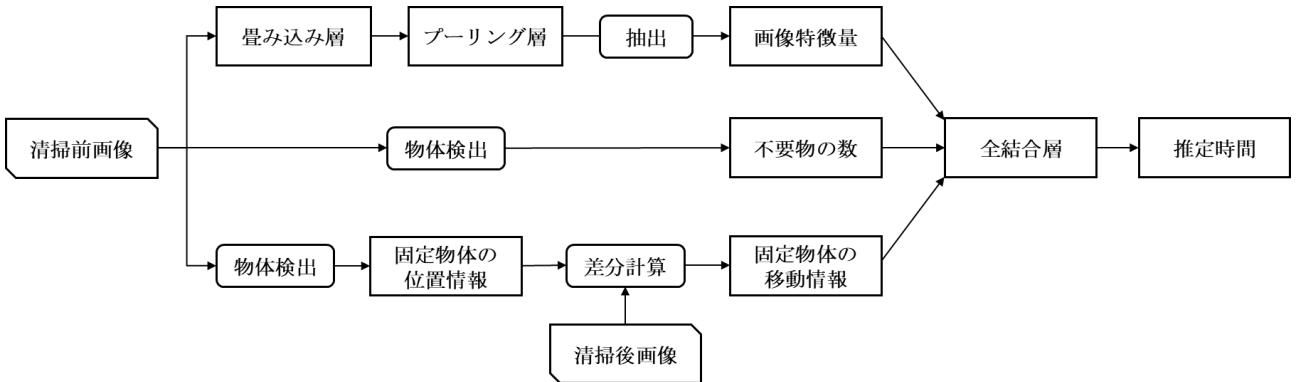


図 2 : 清掃時間の推定手順

を図 2 に示す。清掃前画像に対し、事前学習済みの CNN モデルを用いることで、画像の特徴量を抽出する。さらに、物体検出を用いることで、不要物の数と固定物体の位置情報を取得する。物体検出によって得られた固定物体の位置情報を清掃後画像と比較することで、固定物体の移動情報の差分を計算する。最後に、画像特徴量、不要物の数、固定物体の移動情報を統合し、全結合層に入力することで清掃時間を推定する。

### 3. 評価実験

実際のホテルの客室画像データを十分に収集することが困難であったため、本来のホテルの客室清掃を想定した環境に近い条件として、ホテルの机上環境を模擬して作成したデータセットを用いて評価実験を実施した。①画像特徴量のみを用いた手法、②画像特徴量と不要物の数を用いた手法、③画像特徴量と不要物の数に加えて固定物体の移動情報を用いた手法について、学習データ数を 50, 100, 150 の 3 通りに変化させ、計 9 通りの条件下で 10 分割クロスバリデーションを行った。二乗平均平方誤差 RMSE と決定係数  $R^2$  を表 1 に示す。

### 4. 考察

手法②の結果は、手法①より良い、あるいは同

程度であることから、不要物の数を追加情報として用いることで、予測の安定性が高まることがわかる。特に、少ない学習データ数 50 の条件では、改善度合いが高く、モデルがより正確な清掃時間を推定できるようになっている。しかし、本研究は、ホテルの机上環境を模擬したデータセットで物体検出モデルを学習したため、実際の客室へ適用するには、モデルの再学習や改変が必要である。

一方、手法③は固定物体の移動情報を追加してもかかわらず、すべての学習データ数で  $R^2$  が負の値を示し、予測性能が低下した。特に、少ない学習データ数 50 の条件では、RMSE が 26.68 と大幅に増加し、 $R^2$  が -45.39 と極端に低い値を示している。現時点の学習データ数や特徴量の扱い方では、清掃前後で動かされた固定物体の位置情報が、ノイズ化すると考えられる。したがって、固定物体の移動情報は十分なデータ数や適切な前処理がなければ、モデルにとって逆効果になるといえる。今後の課題は、学習データ数の拡張と固定物体の移動情報に関する統合方法の再検討である。

### 参考文献

- [1] T. Lesiuk, "The Effect of Music Listening on Work Performance," *Psychology of Music*, vol. 33, no. 2, pp. 173-191, 2005.
- [2] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 779-788, 2016.

表 1 : 清掃時間推定モデルの評価値

学習 データ数	①		②		③	
	RMSE	$R^2$	RMSE	$R^2$	RMSE	$R^2$
50	12.89	-5.93	7.58	-1.19	26.68	-45.39
100	5.77	0.38	5.85	0.44	8.85	-0.20
150	5.09	0.60	4.99	0.60	7.97	-0.28