

日立の研究領域

日立製作所の研究領域は、映像・音声・センサデータの解析技術から自然言語処理，ロボティクスまで多岐にわたります。

特に日立の研究所で開発された映像解析技術は、システムインテグレーションやクラウドサービス等のITシステム分野から、鉄道・産業・エネルギー等のインフラシステム分野まで日立が手掛ける幅広い事業領域に向けて展開されています。



適用事例：映像セキュリティ向け人物追跡技術



空港、駅などの大規模施設や街区などの公共空間では、安全確保のために防犯カメラによる監視や警備が行われていますが、事件の早期発見と解決には、施設利用者などの目撃情報をもとに不審者や迷子を、広域の防犯カメラ網からリアルタイムに把握する必要があります。

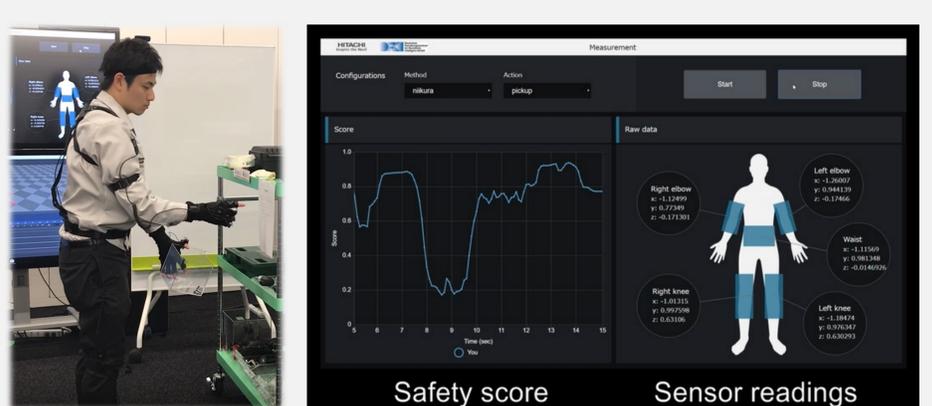
日立は、性別や年齢層、服装など多数の特徴情報をリアルタイムに判別することで、該当する人物を即座に発見する技術を開発しました。さらに、その人物がどのような足取りを取ったのかを、リアルタイムに広域の防犯カメラ映像の中から抽出することも可能です。



【関連ニュースリリース】

<https://www.hitachi.co.jp/New/cnews/month/2017/03/0327.html>

適用事例：製造・保守向け作業支援技術



センサを着用した作業者

身体負荷を部位ごとに定量化し、改善箇所をフィードバック

熟練労働力の減少と新たな労働力の確保が社会課題となっており、製造・保守現場における作業者の教育支援の必要性が高まっています。

日立はドイツ人工知能研究センタ（DFKI）と共同で、スーツ型のウェアラブルデバイスを用いて作業者の負荷を身体部位ごとに定量評価する技術を開発しました。ウェアラブルデバイスを利用することで、カメラを用いた手法よりも安定的かつ定量的な身体負荷評価を実現しました。またこの技術により評価した負荷を模範的な作業動作と比較し、その差異をフィードバックする教育支援も可能となりました。

なお、本システムは世界最大の産業見本市であるドイツハノーバーメッセに出展しました。

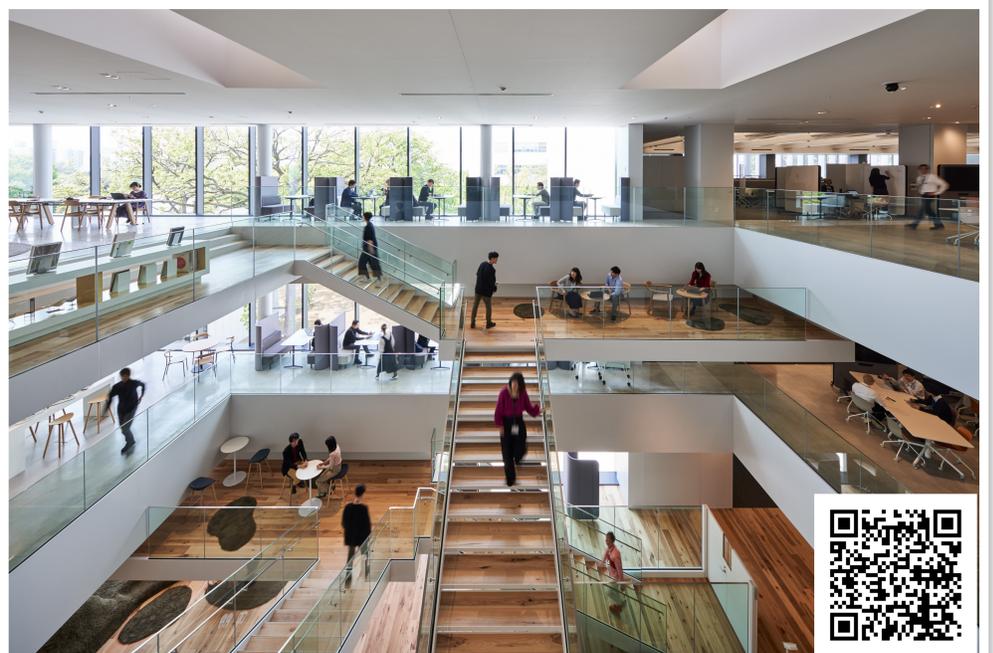
【関連ニュースリリース】

<https://www.hitachi.co.jp/New/cnews/month/2019/03/0320.html>



日立の研究環境

日立製作所の映像解析研究チームは、緑豊かな東京都国分寺市の中央研究所を主な拠点とし、研究開発を行っています。20年3月発足の新組織でオープンイノベーションを推進します。

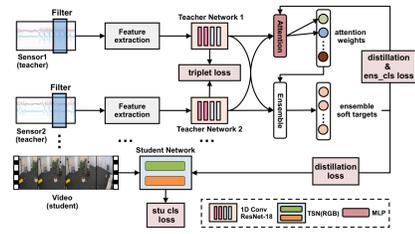


行動認識データセットMMAct [ICCV'19]

マルチモーダル行動認識データセットMMAct
37種類行動、4シーン、7種類センサ、36Kサンプル



映像、センサ間のクロスモーダル学習

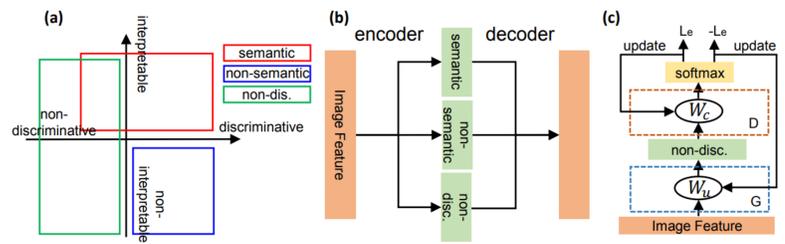


大規模マルチモーダル行動認識データセットMMActを構築、それを用いて、人の動きをセンシティブに捉える複数の身体装着センサーの情報を映像の行動認識モデルに学習させるクロスモーダル行動認識手法を提案し、遮蔽ありや微小な動きへの認識精度大幅な改善を確認した。



Dataset Download <https://mmact19.github.io/2019/>

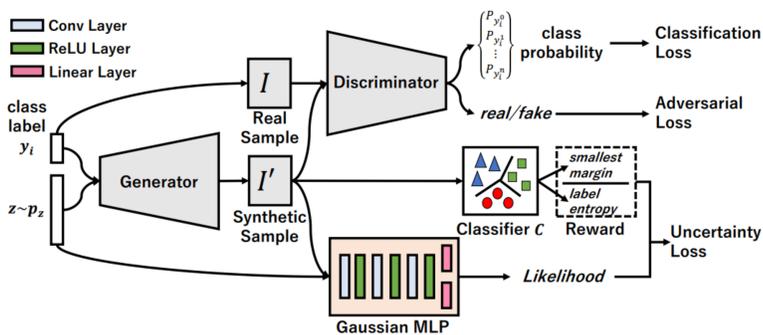
ゼロショット学習 [CVPR'19, AAAI'18]



ゼロショット学習において、未知の画像の分類に有用な特徴量を獲得する手法を提案した。本手法では、まず既知の画像のみを用いて識別モデルを学習する。その学習済みモデルから得られる特徴量から、さらに未知の画像の分類に有用と考えられる特徴量を抽出するような、オートエンコーダを用いたモデルを導入する。提案手法によって、従来手法を上回る性能を実現した。

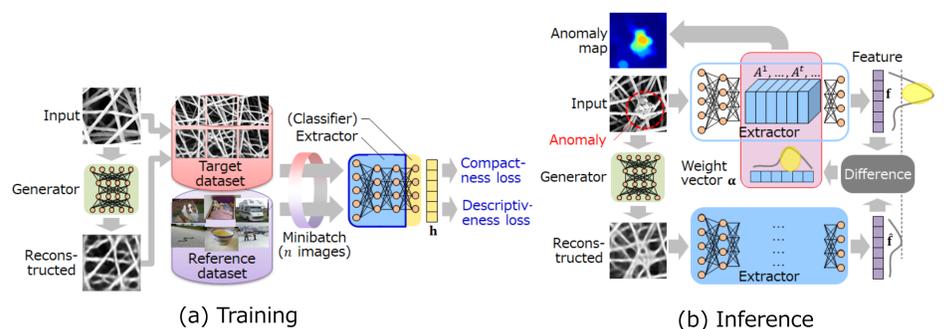
GANを用いた教師データ生成 [AAAI'19]

本研究は能動学習のコンセプトの元に、学習に有益 (informative) な教師付きデータを生成できる生成モデルActiveGANを提案した。ActiveGANは、条件付き生成モデルに基づき、さらに生成したサンプルの不確定性報酬を最大化するように、方策勾配法によってパラメータを更新するのを特徴とする。評価として、分類タスクにおける大規模データセットから生成したサンプルを用いて分類モデルの精度を向上させることに成功した。

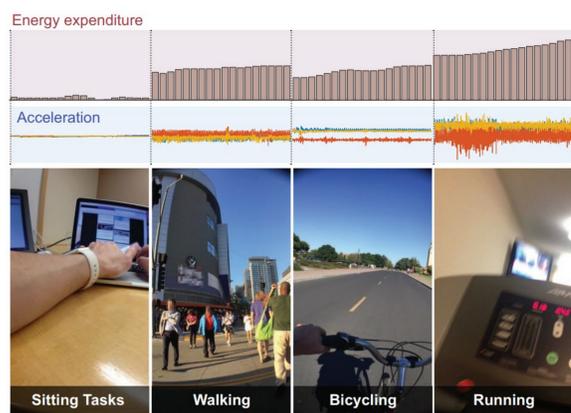


教師無し異常検知 [ICANN'19]

画像中の異常領域を特定する画像診断では、異常状態の定義及びデータ入手が困難という課題がある。一方、人間は過去に得た知見 (識別能力) を活用し、未知の異常であっても異常領域の特異的な特徴を捉え、異常領域を特定することが可能である。そこで、画像診断の対象画像とは別の外部データセットから知見を習得し、未知の異常に対しても特徴量レベルで異常領域を特定可能な手法を提案した。二種のデータセットで実験を行い、有用性を示した。

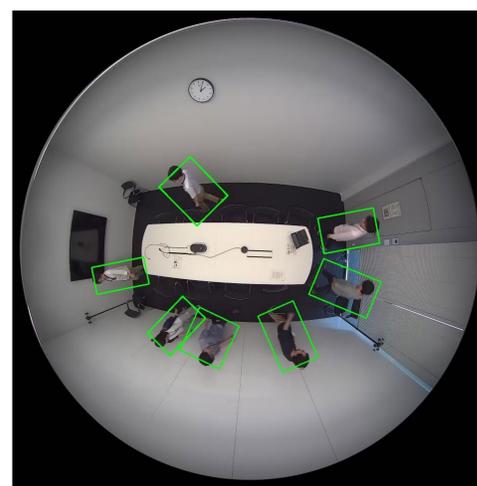


一人称視点映像からの動作認識・消費エネルギー予測 [CVPR'17]



ウェアラブルカメラ映像から運動による消費エネルギーを推定する新タスク"Visual Calorie Counter"を提案した。時系列マルチモーダル深層学習のフレームワークにおいて、消費エネルギー推定と行動認識とのマルチタスク学習を行うことでこれを解決した。学習時には心拍数を消費エネルギーの自己教師データとして用いた。

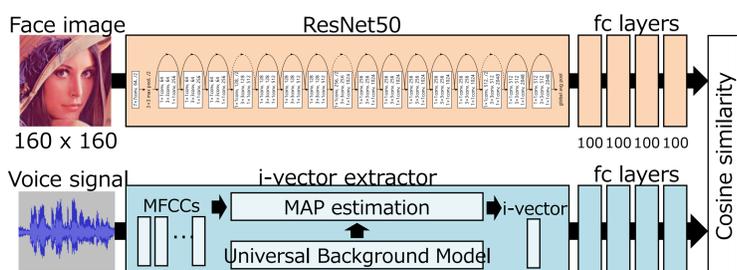
全方位カメラ人物検出 [WACV'19]



全方位カメラからの人物検出は、セキュリティ上重要な課題であるが、検出器の学習に用いるデータの少なさが課題となっていた。本研究では、透視投影カメラ画像をランダムに回転させたデータを用いることで、回転不変な人物検出器を学習する手法を提案した。これにより、全方位画像を一度透視投影画像に変換してから一般的な人物検出を行う場合よりも、高速かつ高精度な全方位カメラ画像からの人物検出を行うことが可能になった。

顔⇔声の予測 [ACMMM'18]

「人間は、声のみからその人の顔がある程度予測できる」ということが心理学分野の研究によって明らかになっている。本研究では、顔と声を同一の空間に射影するクロスモーダルなニューラルネットワークを提案し、この人間の性質を再現することを試みた。クロスモーダルな特徴空間において最近傍探索を行うことで、人間の顔から声、声から顔を人間と同程度の精度で予測することに成功した。



人物の骨格情報を用いた監視向け人物間インタラクション認識 [SSII'20]



公共施設のセキュリティ強化・監視業務支援を目的として、二名の連携行動 (インタラクション) の認識手法を提案した。特に本手法では、危険物等の受け渡し行動の認識を目的としている。本手法の特徴は、骨格情報から人物間の相対距離を表現する特徴量を算出し、低い学習コストでありながらも、受け渡し方・撮像位置の変化への高い頑健性を有する点にある。約100シーンの受け渡し行動を用いた学習で、約90%の認識率を達成した。