核セキュリティのための内部脅威者の自動検知技術の開発

Automatic Detection of Malicious Insider Behavior for Nuclear Security

東京大学	川崎　祐典	Yusuke Kawasaki	No-Member
東京大学	出町　和之	Kazuyuki Demachi	Member
東京大学	笠原　直人	Naoto Kasahara	Member

Abstract

After Fukushima accident, the threats of terrorism are increasing. However, countermeasures for insider terrorists are insufficient because the main target of Physical Protection System (PPS) is outsider. Thus development of automatic detection of insider terrorists is required. In this paper, focusing on the hand motion, final goal is image recognition by Convolutional Neural Network (CNN) and analysis of each finger's position and angle and detection of signs of malicious behavior. And For this, results of trying CNN and making dataset for hand motion recognition are showed. Future work is development of analysis method of time-series feature data and detection of malicious behavior.

Keywords: Nuclear security, Insider terrorists, Abnormal detection, Image analysis, Image recognition

1．緒言

1.1. 原子力発電所における核セキュリティの脅威

福島第一事故後重要機器の構造などの情報がメディアを通じて発信され、これは原子力発電所における重大事故に繋がるシナリオの開示を意味し、テロリズムの脅威が増加している可能性を示唆している。Fig. 1 は国際原子力機関（IAEA）による核テロリズムの分類であり、核テロリズムは核兵器の盗取、核物質の盗取、ダーティーボムの製造、妨害破壊行為の４つに分類される。原子力施設の安全という観点から考えると特に注目すべきは妨害破壊行為である。

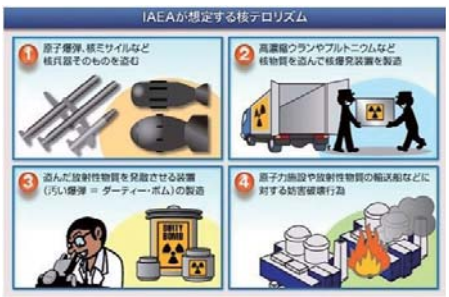

Fig.1 Threats of nuclear security

連絡先:、〒113-0033 東京都文京区本郷 7 丁目 3-1、東京大学大学院工学系研究科原子力国際専攻
E-mail : yu_ten@live.jp

1.2. 物理的防護システム（PPS）

原子力施設におけるテロ対策システムとして物理的防護システム（PPS）がある。PPS の主な目的はテロリストの侵入を防ぎ行為の完遂を阻止することである。侵入された場合、侵入後の位置を知る「検知」、行為完遂のための時間を稼ぐ「遅延」、対抗手段としての外部武力である「対応」の三つの段階がある。しかしこの PPS は問題を抱えている。

1.3. PPS の問題点

テロリストはアウトサイダーとインサイダー（内部脅威者）に分類され、アウトサイダーは組織外部に位置し武力装備しており、インサイダーは組織内部に潜伏し専門的な知識を有するという特徴がある。PPS の主な対象はアウトサイダーでありインサイダーに対する効力は低い。加えてインサイダーと通常作業員との区別は困難であり、原子力発電所内を人力で監視することが実質不可能である。また侵入防護としての身辺調査等は整備予定である。従って本研究ではインサイダーの自動検知技術の開発に焦点を当てる。

1.4. インサイダー検知へのアプローチ

前述した様にインサイダーの自動検知には通常作業員との区別が必要である。また、重大事故に繋がるシナリオを特に検知する必要からインサイダーの行為そのものを検知する必要がある。これに対する本研究の提案は以下

の通りである。初めにインサイダーによる妨害破壊行為をその構成要素(顔、ID、行動情報(全体姿勢、手の動き)、所持物、等)に分解する。その後それぞれの要素ごとの解析を行う。最後に解析結果を組み合わせることで多層的に行為の危険度を評価する。この手法により本来不可能である人間の行動の検知を間接的に可能とする。また、構成要素の情報の取得には監視カメラのような動画像による取得が効率的であると考え本研究では動画像を対象に解析を行うものとする。

1.5. 先行研究

一般的に画像の自動認識には機械学習、特に近年では深層学習が用いられる。また異常検知の手法は世の中に数多く存在している。しかしこれらの手法において動画を対象に扱っているものはない。従って本研究における解析対象は動画とする。これによりインサイダーの危険行動の変化の予兆を検知することが可能となりより早期の検知が可能となる。

1.6. インサイダーの自動検知の全体の流れ

本研究の最終的な目的であるインサイダーのけちシステムの全体の流れは以下の通り。撮影されたインサイダーの妨害破壊行為から特徴量を抽出する。次にそれらを時系列データとして解析する。その時系列データから変化の予兆を検知する。ここで、人の全体姿勢等の体全体の動きの情報と手の動きの情報が特に人の行動を決定づけていること、手の動きを取得・解析する手法が確率されていないことから、この中でもさらに手の動きに着目し機械学習法を用いた画像解析による特徴量の抽出のための基礎技術の開発を行うものとする。

２．目的

インサイダーの妨害破壊行為を自動検知できるシステムの開発を本研究の目的とする。特に手の動きに着目し、機械学習法を用いた画像解析により特徴量の抽出のための基礎技術の開発を目的とする。

3. 手法

3.1. Convolutional Neural Network (CNN)

一般的に画像の自動認識には機械学習法が用いられ、中でも画像認識の分野では Convolutional Neural Network (CNN) が非常に高性能である。本研究でもこの CNN を用いる。

CNN は畳み込み部とサブサンプリング部、全素子結合部の三つから成り、フィルタリング処理と識別の役割を担っている。この構造中には多くのパラメータがあり、事前の最適化が必要なハイパーパラメータと自動更新される学習パラメータがある。

４．結果

4.1. CNN による手書き文字認識

CNN の画像認識における性能の確認のため手書き文字認識 (MNIST) を行った。3000 枚の画像による学習を行い 500 枚の画像により検証を行ったところ、誤回答率は 5% 以下となり、非常に高い性能を示した。

4.2. 手の動きの特徴量抽出のためのデータセットの作製

事前学習のためのデータセットの作製を行った。このデータセットは 600 枚の画像とそれに対するラベルで構成されている。画像は元の画像から背景除去、トリミング後に手の中心を画像の中心に移動させ手以外のノイズを除去することで作製した。ラベルは各指がどの程度開いているかを０～１の値で評価したものである。Fig.2 に示すのは元画像とそこから作製されたデータセットの画像の例である。

Fig.2 Raw image and image for CNN

５．結論及び課題

CNN は画像認識、特に MNIST では非常に高性能な結果を示した。また、CNN による手の動きの特徴量抽出のためのデータセットを作製した。

今後の課題として CNN による手の動きの特徴抽出プログラムの作製、特徴量の解析手法の作製、解析結果からの予兆検知手法の作製がある。

参考文献

[1] Yann LeCun, et al.: Learning Methods for Generic Object Recognition with Invariance to Pose and Lighting, *CVPR, IEEE, 2004.*