

# Identificação de Objetos Ilícitos em Visitantes de Penitenciárias Utilizando Aprendizado Não Supervisionado

Bruno A. S. Oliveira\* Hugo E. Ziviani\*\* Luana Gonçalves\*  
Alan M. Viegas\*\* Daniel Calvo\*

\* Fundação para Inovações Tecnológicas - FITec, Belo Horizonte, MG,  
(e-mail: brunooliveira, lgoncalves, dcalvo {@fitec.org.br}).

\*\* VMI Sistemas de Segurança, Lagoa Santa, MG, (e-mail:  
hugo.ziviani, alan {@vmis.com.br}).

---

**Abstract:** Prisons have strict rules regarding the entrance of employees, visitors, and inmates, which are checked for the presence of hidden objects. Several facilities perform these security inspections using full-body x-ray images. The inspection is carried out by a security agent who screens the x-ray images looking for forbidden objects or contraband, in a process that is slow, repetitive, and prone to errors. This work proposes a solution based on unsupervised machine learning, which implements a Convolutional Autoencoder, aimed at assisting the operator in the identification of forbidden items. The Convolutional Autoencoder was trained on a dataset containing 2,621 images with no hidden objects and 130 images considered suspect, all collected and classified by an expert. The results obtained indicate that the solution is promising and can be used to assist the agent at identifying people trying to enter the facility with forbidden items. By using the system, a greater number of illicit objects can be identified and prevented from being carried into the facilities.

**Resumo:** Com o objetivo de impedir a entrada de objetos ilícitos que comprometam a segurança de funcionários, visitantes e detentos em instalações penitenciárias, a inspeção a partir de imagens radiológicas é amplamente utilizada para o monitoramento de ingressantes a complexos carcerários. Atualmente essa fiscalização é realizada por um operador que analisa imagens obtidas por equipamentos de raio-x, buscando identificar algum objeto suspeito. No entanto, tal processo é lento, dispendioso e repetitivo. Nesse sentido, o objetivo desse trabalho foi desenvolver uma solução baseada em aprendizado de máquina não supervisionado, denominada Autoencoder Convolutacional, cuja finalidade é auxiliar o operador na identificação de objetos ilícitos em imagens de raio-x de corpo inteiro. Para isso, foram coletadas e classificadas por um especialista 2.621 de imagens sem indicações de objetos suspeitos e 130 imagens com suspeita de objetos ilícitos. Em seguida, um Autoencoder Convolutacional foi treinado e avaliado, para detecção de anomalias. Os resultados obtidos indicam que a solução é promissora e pode ser utilizada no auxílio à tomada de decisão do operador do equipamento de raio-x. Com o uso do sistema, um maior número de objetos ilícitos pode ser identificado e conseqüentemente impedir que materiais não permitidos sejam transportados para dentro das penitenciárias.

*Keywords:* Deep Learning; Convolutional Autoencoder; Anomaly Detection; Identification; X-ray images; Monitoring; Illicit Objects; Security.

*Palavras-chaves:* Aprendizado Profundo; Autoencoder Convolutacional; Detecção de Anomalias; Identificação; Imagens de Raio-x; Monitoramento; Objetos Ilícitos; Segurança.

---

## 1. INTRODUÇÃO

A análise de imagens radiológicas é um dos métodos mais utilizados para monitoramento do tráfego de pessoas em aeroportos e penitenciárias com restrição ao porte de objetos específicos. Nesses locais é comum a proibição de itens que comprometam a segurança interna das instalações monitoradas. Tais objetos incluem armas de fogo, lâminas, explosivos, narcóticos e, nos casos de instituições de reclusão de pessoas, também celulares e objetos metálicos. Além disso, em fronteira de países é imprescindível o

controle rígido de entrada, com o intuito de fazer cumprir as políticas migratórias do país.

No entanto, o monitoramento encontra diversos fatores adversos, tais como a invasão de privacidade e a exposição à radiação ionizante. Apesar de autorizar o uso de imagens de raio-x, a Comissão Nacional de Energia Nuclear (CNEN), responsável por todos os aspectos que envolvem a utilização de material radioativo e aparelhos que geram radiação ionizante, regulamenta doses de radiação relativamente baixas, quando comparadas às imagens de finalidade médico-diagnóstica. Isso se deve ao fato de

frequentemente um indivíduo precisar ser exposto a múltiplas exposições em curtos períodos. O baixo nível de radiação é um fator adverso, na medida em que afeta a qualidade visual da imagem. Devido o fluxo de pessoas ser intenso, o operador tem um curto período para inspecionar as imagens, dispondo de aproximadamente dez a quinze segundos. Isto posto, estudos recentes de processamento de imagem e aprendizado de máquina apresentam técnicas para auxiliar o operador na tarefa de inspecionar tais imagens.

As técnicas que objetivam uma melhoria das imagens (Chen et al., 2005; Abidi et al., 2006; Lu and Connors, 2006), classificação (Caldwell et al., 2017; Akçay et al., 2018b; Miao et al., 2019; Yang et al., 2019), detecção (Liang et al., 2018; Steitz et al., 2018; Liu et al., 2018), segmentação (Singh and Singh, 2004; Heitz and Chechik, 2010; Gaus et al., 2019) e detecção de anomalias, empregam métodos de Inteligência Artificial e aprendizado de máquina não-supervisionado (Andrews et al., 2016; Akçay et al., 2018a; Akçay et al., 2019; Griffin et al., 2018) e têm como foco principal auxiliar a identificação de ilícitos. Contudo, ainda existem poucos trabalhos que abordam o tema por meio de sistemas automatizados, em virtude da ausência de bancos públicos de imagens (Akçay and Breckon, 2020).

Para amenizar o problema da falta de especialistas treinados para uma tomada de decisão correta (Oliveira et al., 2018, 2021) e diminuir o tempo necessário para a identificação de objetos ilícitos, é necessário o desenvolvimento de tecnologias que visam automatizar esse processo. O auxílio de um sistema computacional poderá facilitar a tomada de decisão pelo operador, aumentando a assertividade do processo e dando uma maior segurança para o especialista.

Nos últimos anos, os estudos em aprendizado profundo vêm se difundindo e suas aplicações tem se tornando cada vez mais presentes no dia a dia da sociedade. O aprendizado profundo pode ser entendido como uma família de métodos de aprendizado de máquina que são baseados em redes neurais artificiais (RNA). Esse aprendizado pode ser supervisionado, semi-supervisionado ou não supervisionado. Algumas pesquisas (Mouton and Breckon, 2015; Rogers et al., 2017) dividem o campo de estudo nessa área em duas partes: processamento digital de imagens e reconhecimento de imagem. Os primeiros trabalhos tentam encontrar objetos ilícitos usando técnicas convencionais, enquanto a segunda geração de estudos tenta cumprir o mesmo objetivo, porém por meio de aprendizado de máquina e aprendizado profundo (Miao et al., 2019; Steitz et al., 2018; Gaus et al., 2019; Akçay et al., 2019; Griffin et al., 2018).

Dessa forma, o objetivo deste estudo é propor uma solução que utiliza a técnica de aprendizado não supervisionado Autoencoder Convolutacional para detecção de anomalias e suporte da identificação de objetos ilícitos em imagens de raio-x de corpo inteiro. A contribuição dessa pesquisa consiste em explorar o uso da técnica de aprendizado não supervisionado, uma vez que existe uma grande dificuldade para obtenção de um conjunto significativo de imagens rotuladas.

O texto está organizado em seções. Na seção 2 é apresentado o referencial teórico e as aplicações na área de

detecção de anomalias, assim como a técnica Autoencoder Convolutacional. Na seção 3 são apresentados os materiais empregados e a metodologia da solução proposta. A seção 4 apresenta os resultados obtidos do estudo. Por fim, a seção 5 mostra as conclusões, limitações e trabalhos futuros.

## 2. REFERENCIAL TEÓRICO

### 2.1 Detecção de Anomalias

A detecção de anomalias é o processo de identificação de eventos inesperados dentro de um conjunto de dados e geralmente é aplicada em dados não rotulados, por meio do aprendizado não supervisionado. O processo pode ser composto por duas suposições básicas: (1) raridade na ocorrência de anomalias e (2) os recursos das ocorrências anômalas diferem significativamente das instâncias consideradas comuns (Chandola et al., 2009).

Em algumas aplicações, adquirir um conjunto de dados representativo, balanceado e rotulado é difícil e extremamente caro. Diferentemente da análise de imagens de cenas naturais, determinar se uma radiografia apresenta anormalidades é uma tarefa que depende de especialistas com anos de treinamento e estudo. Outro fator a ser analisado é se apenas a classificação da imagem é suficiente ou se seria necessário utilizar outros processos de análise. Dessa forma, a aplicação de técnicas de aprendizado não supervisionado para a detecção de anomalias pode ser uma alternativa promissora.

### 2.2 Autoencoder Convolutacional

Autoencoder é um tipo de Rede Neural Artificial que recebe como entrada um conjunto de dados e os comprime para uma representação em um espaço latente com os melhores padrões daqueles dados. Em seguida, com o espaço latente definido, o Autoencoder reconstrói os dados de entrada a partir da representação latente, gerando a saída do método (Baldi, 2012).

Normalmente, um Autoencoder possui dois componentes: codificador e decodificador. O primeiro recebe os dados de entrada e os compacta no espaço latente, enquanto o segundo é responsável por reconstruir a entrada original a partir da representação latente. Em aplicações com imagens, são especificados os Autoencoders convolutacionais, por possuírem camadas de convolução na rede.

De modo geral, em aplicações de visão computacional os Autoencoders são empregados para: (1) redução de dimensionalidade; (2) remoção de ruído; (3) detecção de anomalias. Neste trabalho tem-se o uso na identificação de valores que fogem da normalidade para o caso de exemplos que não são tão bem reconstruídos pelo algoritmo, pelo simples fato de serem anômalos.

## 3. MATERIAL E MÉTODOS

### 3.1 Coleta das Imagens

O *Spectrum Bodyscan* é um equipamento para inspeção de corpo inteiro, com capacidade para auxiliar o operador na detecção de objetos ocultos, incluindo objetos engolidos ou

introduzidos no corpo. O equipamento é capaz de prover alto nível de controle e segurança para penitenciárias, aeroportos, postos de fronteira e instalações governamentais. Na Figura 1 é apresentado um exemplo de duas imagens obtidas pelo *Spectrum Bodyscan*. As imagens são salvas no formato *Portable Network Graphics* (PNG) e possuem o tamanho de  $1024 \times N$ , onde  $N$  depende da massa corporal do indivíduo.



Figura 1. Exemplo de imagens de duas pessoas diferentes capturadas pelo *Spectrum Bodyscan*.

Com o auxílio de um especialista treinado na identificação de objetos ilícitos, foram separadas manualmente 2.621 imagens de pessoas que não apresentam objetos suspeitos e 130 imagens de pessoas que chamaram a atenção do operador do equipamento de raio-x.

### 3.2 Pré-processamento

Para o desenvolvimento do estudo foi empregada a linguagem de programação Python 3.6, no ambiente de desenvolvimento Colab, disponibilizado pela Google. O Colab possui uma GPU Tesla T4 com 15.109MB de memória, fazendo seu uso pertinente para o treinamento e validação de algoritmos baseados em aprendizado profundo. Além disso, conta com uma interface na Web, fazendo com que não seja necessária a instalação de ambientes ou programas localmente (Bisong, 2019).

Como apresentado na Figura 1, é praticamente impossível que seja detectado algum objeto suspeito na imagem, até mesmo para um especialista treinado. Por isso, um filtro de equalização de histograma (Pizer et al., 1987) foi aplicado, com o objetivo de melhorar o contraste das imagens. Dessa forma, torna-se possível perceber alguns objetos (anel, arma, colar, drogas) caso estejam presentes na imagem.

Além disso, como pode ser percebido na Figura 1, o número de linhas da matriz na imagem é fixo, não dependendo da altura do indivíduo. Para que houvesse menos perda no processo de redimensionamento da imagem, todas as linhas e as colunas que tivessem valores repetidos foram removidas, resultando em uma imagem cuja área se limitava à região com presença da pessoa.

Em seguida, todas as imagens foram redimensionadas para o tamanho de  $256 \times 256$  e os valores dos pixels foram invertidos. A inversão dos pixels é utilizada para que o fundo da imagem tenha valor zero e seja desconsiderado pela rede. Dessa maneira, os pixels do fundo possuem o valor zero e os pixels que representam a pessoa são próximos de 255.

Na sequência, as imagens normais foram separadas em duas partes, de forma aleatória: 2.482 imagens para o treinamento e 139 para teste. Além disso, foi aplicada a operação de inversão horizontal em cópias das imagens de treinamento, com o objetivo de aumentar o número de imagens disponíveis. Ressalta-se que as imagens de pessoas suspeitas foram mantidas sem alterações, visto que somente as imagens de pessoas normais seriam utilizadas para o treinamento do Autoencoder Convolutivo.

Por fim, foi utilizado o método *ImageDataGenerator* disponível na biblioteca Keras (Chollet et al., 2015) para carregar as imagens do diretório, sendo o tamanho do lote definido como 64, reescalando os valores dos pixels por 255 para acelerar o processo de treinamento, especificando que as imagens estão em tons de cinza e em um tamanho de  $256 \times 256$ , além de definir que as imagens de saída são idênticas às imagens de entrada. A Figura 2 ilustra dois exemplos de imagens pré-processadas que foram empregadas para o treinamento do modelo.



Figura 2. Exemplo de imagens pré-processadas de duas pessoas diferentes.

### 3.3 Detecção de ilícitos

O Autoencoder Convolutivo foi implementado da seguinte forma: na camada de entrada, é esperada uma matriz  $256 \times 256$  com apenas um canal de profundidade. A primeira parte do modelo, também chamada de codificador, possui 32, 64, 128 e 256 filtros, respectivamente, de forma que, para cada camada de convolução 2D, cada filtro possui tamanho  $3 \times 3$ , com um passo de dois e com um preenchimento do tipo *same*. Além disso, em cada operação descrita é utilizada a função de ativação *Leaky-ReLU* com um *alpha* de valor 0,2, tendo sido empregada a operação para normalização em lote.

Definida a arquitetura do codificador, espera-se como saída um vetor latente com 128 recursos, valores que teoricamente devem conter as características mais relevantes da imagem original, cujo tamanho é  $256 \times 256$ . Já na segunda parte do Autoencoder Convolucional, também conhecida como decodificador, é utilizada a mesma arquitetura do codificador, porém, de maneira inversa. A Figura 3 ilustra a arquitetura do Autoencoder Convolucional implementado, obtida por meio da função disponível na biblioteca Keras `keras.utils.vis_utils`. Como pode ser observado, a entrada do modelo recebe uma imagem original com o tamanho de  $256 \times 256$  e gera em sua saída a imagem reconstruída com a mesma dimensão da entrada,  $256 \times 256$ .

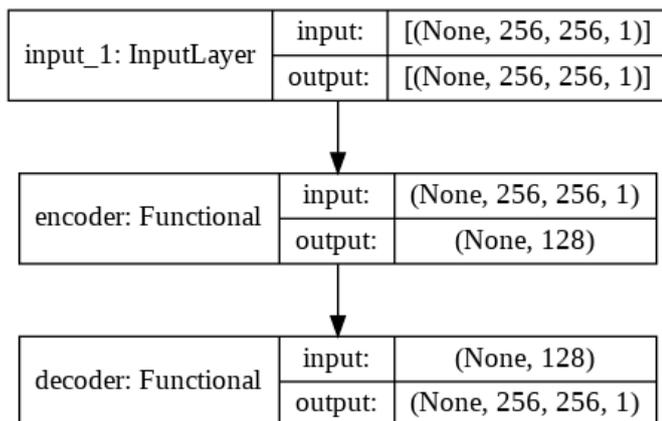


Figura 3. Arquitetura do Autoencoder Convolucional.

Para o treinamento foram definidas 25 épocas e para a função de perda a métrica *mean squared error* (MSE). Além disso, o otimizador especificado foi o *Adaptive Moment Estimation* (Adam) (Kingma and Ba, 2014) com uma taxa de aprendizado em  $1e-3$  e um decaimento sendo a razão entre a taxa de aprendizado e o número de épocas. Por fim, foi posto que o valor de saída esperado é o mesmo que o valor de entrada, uma vez que o objetivo do Autoencoder Convolucional é reconstruir a mesma imagem de entrada.

Definida a arquitetura do Autoencoder Convolucional e realizada a configuração para o treinamento do modelo, foi realizado o treinamento propriamente dito. Em seguida, o modelo foi avaliado nas imagens de teste e nas imagens suspeitas, para que se pudesse avaliar a capacidade de reconstrução da rede.

Com a imagem reconstruída, torna-se possível calcular o erro em cada um dos pixels, através da equação 1:

$$|(real - previsto)^2| \quad (1)$$

aonde é retornado o valor absoluto da subtração entre o valor do pixel real e o valor do pixel previsto elevado ao quadrado. O erro calculado pela subtração é elevado ao quadrado com o objetivo de destacar valores discrepantes.

Com isso, torna-se possível gerar um mapa de calor da matriz de erros calculada pela equação 1. Assim, espera-se que os maiores valores de erros gerados serão nos locais com presença de objetos suspeitos, uma vez que o modelo treinado não conseguiria reconstruí-los corretamente, pelo fato de não terem sido incluídos no treinamento.

Além disso, considerando que os erros de reconstrução são pontuais nas imagens, calcular um único erro global não é suficiente para classificar a imagem como um todo em suspeita ou não suspeita. Sendo assim, foi definido que os 100 maiores erros seriam utilizados para essa classificação da imagem, calculando uma média desses valores.

Dessa forma, no momento da inferência de novas imagens, o mapa de calor gerado pode ser apresentado ao operador chamando sua atenção nos locais em que houve uma maior discrepância na reconstrução da imagem, auxiliando o agente em sua tomada de decisão para a presença ou não de um objeto ilícito ali contido.

#### 4. RESULTADOS E DISCUSSÃO

A Tabela 1 apresenta o resultado médio e o desvio padrão da imagem como um todo, considerando a média dos 100 maiores erros calculados para cada imagem.

Tabela 1. Erro médio calculado entre as imagens de pessoas suspeitas e não suspeitas.

Classe	Total de imagens	Erro médio	Desvio padrão
Não suspeita	139	$\approx 8707$	$\approx 2054$
Suspeita	130	$\approx 22866$	$\approx 6506$

Como mostrado na Tabela 1, fica evidente que o erro médio (considerando os 100 maiores erros) dentre as imagens de pessoas não suspeitas é bem menor que o de pessoas suspeitas. Com o desvio padrão de  $\approx 6506$  para a classe de pessoas suspeitas, percebe-se que, em algumas imagens, pelo fato de o objeto ilícito ser pequeno e pontual, o modelo obtém um erro global próximo ao obtido em imagens de pessoas não suspeitas. Apesar disso, nota-se que o Autoencoder treinado com a métrica de cálculo de erro bem empregada, conseguiu distinguir as imagens entre as duas classes.

A Figura 4 apresenta duas pessoas consideradas sem objetos suspeitos pelo especialista, sendo um adulto na parte superior e uma criança na parte inferior. Na primeira coluna tem-se a imagem original e logo na sequência a imagem reconstruída pelo Autoencoder Convolucional. Já na última coluna é mostrado o mapa de calor gerado através da equação 1.

Assim como esperado, o Autoencoder Convolucional não iria conseguir reconstruir perfeitamente a imagem original, mesmo ela não possuindo nenhum ilícito. Entretanto, chama-se a atenção para os valores mais discrepantes de erros obtidos que, no geral, não ultrapassou o valor de 15.000 para as imagens de teste que não foram consideradas suspeitas pelo especialista.

Como mostrado na Figura 4, poucos erros com um alto valor foram calculados, mesmo assim não ultrapassaram o valor de 15.000, em relação à imagem do adulto (parte superior da Figura 4). Para o caso da criança (parte inferior da Figura 4), pôde-se perceber que a reconstrução do Autoencoder Convolucional foi ainda melhor, não possuindo erros acima do valor de 10.000.

A Figura 5 apresenta duas pessoas devidamente separadas pelo especialista como suspeitas de portarem algum objeto ilícito. Na primeira coluna tem-se a imagem original e logo

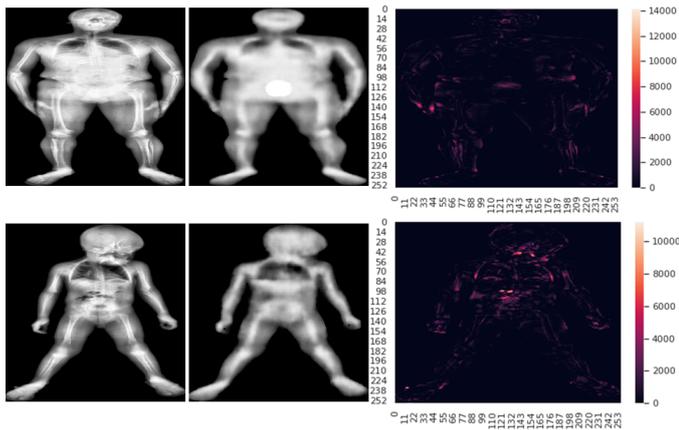


Figura 4. Exemplo com duas imagens originais de pessoas não suspeitas e suas reconstruções, além do mapa de calor de erros obtido.

na sequência a imagem reconstruída pelo Autoencoder Convolutivo. Já na última coluna é mostrado o mapa de calor gerado através da equação 1.

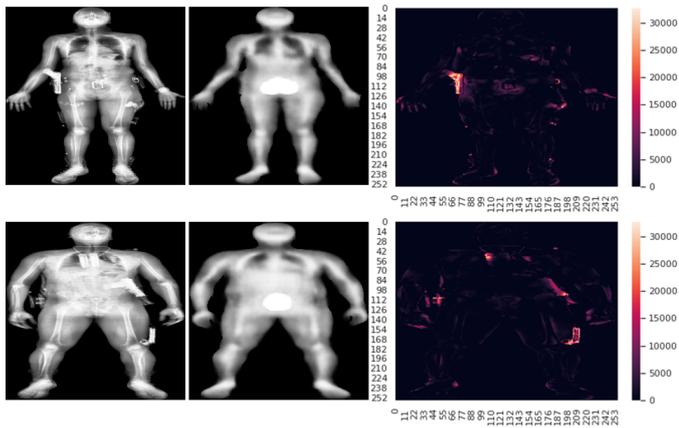


Figura 5. Exemplo com duas imagens originais de pessoas suspeitas e suas reconstruções, além do mapa de calor de erros obtido.

Diferentemente dos erros obtidos e apresentados nos mapas de calor na Figura 4, nos mapas de calor mostrados na Figura 5 é possível perceber um valor de erro muito alto para algumas regiões, sendo próximos a 30.000. Além disso, é notório nos mapas de calor algumas regiões com um tom mais claro, ressaltando que ali houve um maior erro de reconstrução da imagem.

Ao observar com cautela a Figura 5, pode-se comparar a região mais clara no mapa de calor com a mesma região na imagem original e na imagem reconstruída. Ao fazer isso, percebe-se que um objeto ilícito não foi reconstruído, o que ocasionou um maior erro naquela região.

Na imagem original na parte inferior da Figura 5 nota-se entre os pulmões algum objeto suspeito. Com o mapa de calor gerado, o operador do equipamento de raio-x terá sua atenção chamada para tomar um maior cuidado na avaliação daquela determinada região, fazendo com que a solução proposta tenha auxiliado na tomada de decisão do especialista.

Além disso, pelo fato de a solução ter sido desenvolvida com uma abordagem não supervisionada e treinada com imagens de pessoas que aparentemente não carregavam objetos suspeitos, a técnica apresentada pode ser capaz de capturar novos objetos que ainda não estão no banco de dados, ou mesmo desconhecidos pelas autoridades.

## 5. CONSIDERAÇÕES FINAIS

Nesse artigo foi apresentada uma solução para a identificação de objetos ilícitos por meio de imagens de raio-x de corpo inteiro. Sabendo-se do custo e da dificuldade para se conseguir um conjunto de imagens com milhares de objetos ilícitos e rotulados, empregou-se o método de aprendizado não supervisionado Autoencoder Convolutivo como forma de contornar esse problema.

Os resultados dos experimentos realizados indicam uma proposta promissora para auxiliar a tomada de decisão do operador do equipamento *Spectrum Bodyscan* em selecionar quem pode ou não ter acesso à penitenciária, implicando em uma maior agilidade e assertividade ao processo e reduzindo eventuais desconfortos aos indivíduos submetidos à inspeção.

Como limitações do trabalho, tem-se que a solução ainda deve ser testada em um ambiente real e experimentada em um maior número de imagens. Como trabalhos futuros, é possível que o sistema seja adequado também para outras aplicações como, por exemplo, aeroportos. Métodos de aprendizado supervisionado tais como YOLO, SSD, Faster e Mask R-CNN, entre outros, também podem ser testados, desde que exista disponibilidade de imagens com um número considerável de objetos rotulados.

Além disso, outras formas de se calcular o erro gerado entre a imagem real e a imagem prevista podem ser exploradas. A utilização da média dos 100 maiores erros de acordo com a métrica MSE pode ser a responsável pelo alto valor do desvio padrão. Dessa forma, pode ser válido investigar em trabalhos futuros formas de avaliação menos sensíveis a grandes variações.

## REFERÊNCIAS

- Abidi, B.R., Zheng, Y., Gribok, A.V., and Abidi, M.A. (2006). Improving weapon detection in single energy x-ray images through pseudocoloring. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 36(6), 784–796.
- Akçay, S., Atapour-Abarghouei, A., and Breckon, T.P. (2018a). Ganomaly: Semi-supervised anomaly detection via adversarial training. In *Asian conference on computer vision*, 622–637. Springer.
- Akçay, S., Atapour-Abarghouei, A., and Breckon, T.P. (2019). Skip-ganomaly: Skip connected and adversarially trained encoder-decoder anomaly detection. In *2019 International Joint Conference on Neural Networks (IJCNN)*, 1–8. IEEE.
- Akçay, S. and Breckon, T. (2020). Towards automatic threat detection: A survey of advances of deep learning within x-ray security imaging. *arXiv preprint arXiv:2001.01293*.
- Akçay, S., Kundegorski, M.E., Willcocks, C.G., and Breckon, T.P. (2018b). Using deep convolutional neural

- network architectures for object classification and detection within x-ray baggage security imagery. *IEEE transactions on information forensics and security*, 13(9), 2203–2215.
- Andrews, J.T., Morton, E.J., and Griffin, L.D. (2016). Detecting anomalous data using auto-encoders. *International Journal of Machine Learning*, 6(1), 21.
- Baldi, P. (2012). Autoencoders, unsupervised learning, and deep architectures. In *Proceedings of ICML workshop on unsupervised and transfer learning*, 37–49. JMLR Workshop and Conference Proceedings.
- Bisong, E. (2019). Google colaboratory. In *Building Machine Learning and Deep Learning Models on Google Cloud Platform*, 59–64. Springer.
- Caldwell, M., Ransley, M., Rogers, T.W., and Griffin, L.D. (2017). Transferring x-ray based automated threat detection between scanners with different energies and resolution. In *Counterterrorism, Crime Fighting, and Surveillance Technologies*, volume 10441, 104410F. International Society for Optics and Photonics.
- Chandola, V., Banerjee, A., and Kumar, V. (2009). Anomaly detection: A survey. *ACM computing surveys (CSUR)*, 41(3), 1–58.
- Chen, Z., Zheng, Y., Abidi, B.R., Page, D.L., and Abidi, M.A. (2005). A combinational approach to the fusion, de-noising and enhancement of dual-energy x-ray luggage images. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)-Workshops*, 2–2. IEEE.
- Chollet, F. et al. (2015). Keras. <https://github.com/fchollet/keras>.
- Gaus, Y.F.A., Bhowmik, N., Akçay, S., Guillén-Garcia, P.M., Barker, J.W., and Breckon, T.P. (2019). Evaluation of a dual convolutional neural network architecture for object-wise anomaly detection in cluttered x-ray security imagery. In *2019 International Joint Conference on Neural Networks (IJCNN)*, 1–8. IEEE.
- Griffin, L.D., Caldwell, M., Andrews, J.T., and Bohler, H. (2018). Anomaly detection in x-ray security images. *IEEE Transactions on Information Forensics and Security*, 14(6), 1539–1553.
- Heitz, G. and Chechik, G. (2010). Object separation in x-ray image sets. In *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2093–2100. IEEE.
- Kingma, D.P. and Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.
- Liang, K.J., Heilmann, G., Gregory, C., Diallo, S.O., Carlson, D., Spell, G.P., Sigman, J.B., Roe, K., and Carin, L. (2018). Automatic threat recognition of prohibited items at aviation checkpoint with x-ray imaging: a deep learning approach. In *Anomaly Detection and Imaging with X-Rays (ADIX) III*, volume 10632, 1063203. International Society for Optics and Photonics.
- Liu, Z., Li, J., Shu, Y., and Zhang, D. (2018). Detection and recognition of security detection object based on yolo9000. In *2018 5th International Conference on Systems and Informatics (ICSAI)*, 278–282. IEEE.
- Lu, Q. and Conners, R.W. (2006). Using image processing methods to improve the explosive detection accuracy. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 36(6), 750–760.
- Miao, C., Xie, L., Wan, F., Su, C., Liu, H., Jiao, J., and Ye, Q. (2019). Sixray: A large-scale security inspection x-ray benchmark for prohibited item discovery in overlapping images. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2119–2128.
- Mouton, A. and Breckon, T.P. (2015). A review of automated image understanding within 3d baggage computed tomography security screening. *Journal of X-ray science and technology*, 23(5), 531–555.
- Oliveira, B.A.S., Nametala, M.S., Castoril, B.R., Catalano, M.D., and Nametala, C.A.L. (2018). Avaliação de uma rede neural artificial como estimador temporal pluviométrico no sistema de abastecimento cantareira. *Revista de Informática Aplicada*, 14(1).
- Oliveira, B.A.S., Neto, A.P.D.F., Fernandino, R.M.A., Carvalho, R.F., Fernandes, A.L., and Guimarães, F.G. (2021). Automated monitoring of construction sites of electric power substations using deep learning. *IEEE Access*, 9, 19195–19207.
- Pizer, S.M., Amburn, E.P., Austin, J.D., Cromartie, R., Geselowitz, A., Greer, T., ter Haar Romeny, B., Zimmerman, J.B., and Zuiderveld, K. (1987). Adaptive histogram equalization and its variations. *Computer vision, graphics, and image processing*, 39(3), 355–368.
- Rogers, T.W., Jaccard, N., Morton, E.J., and Griffin, L.D. (2017). Automated x-ray image analysis for cargo security: Critical review and future promise. *Journal of X-ray science and technology*, 25(1), 33–56.
- Singh, M. and Singh, S. (2004). Image segmentation optimisation for x-ray images of airline luggage. In *Proceedings of the 2004 IEEE International Conference on Computational Intelligence for Homeland Security and Personal Safety, 2004. CIHSPS 2004.*, 10–17. IEEE.
- Steitz, J.M.O., Saeedan, F., and Roth, S. (2018). Multi-view x-ray r-cnn. In *German Conference on Pattern Recognition*, 153–168. Springer.
- Yang, J., Zhao, Z., Zhang, H., and Shi, Y. (2019). Data augmentation for x-ray prohibited item images using generative adversarial networks. *IEEE Access*, 7, 28894–28902.