Unisanta Science and Technology, 2026, 25, January
Published Online 2026 Vol.15 Nº1 http://periodicos.unisanta.br/index.php/sat)

UNISANTA
Science and Technology
ISSN 2317-1316

Desenvolvimento de Modelo de IA para Detecção de Situações de Risco em Imagens de Câmeras de Vigilância

Derek Antônio de Lima [1], Arnaldo de Carvalho Junior [1], Walter Augusto Varella [1], João Inácio da Silva Filho [2]

[1] **IFSP** – Inst. Fed. de Educação, Ciência e Tecnologia de São Paulo - Embedded Artificial Intelligence Lab. (EAILAB)
Rua Maria Cristiana, 50, Jardim Casqueiro – Cubatão/SP CEP: 11533-160

[2] **UNISANTA** – Universidade Santa Cecília - Grupo de Lógica Paraconsistente Aplicada
Rua Oswaldo Cruz, 277 - Boqueirão - SANTOS-SP, Brasil - CEP: 11045-100

E-mail: adecarvalhojr@ifsp.edu.br
Received Dec., 2025

Resumo: Câmeras de segurança tornaram-se progressivamente mais comuns no cotidiano da população mundial. Paralelamente ao aumento do número de dispositivos, houve significativa evolução das tecnologias aplicadas, resultando em melhor qualidade de imagem e na incorporação de funcionalidades como visão noturna, captação de áudio, armazenamento em nuvem e comunicação remota, permitindo o monitoramento remoto em tempo real. Grandes capitais passaram a integrar sistemas de videomonitoramento a seus planos de segurança pública. Apesar desses avanços, o monitoramento tradicional ainda apresenta limitações relevantes, sobretudo a necessidade de grande contingente de operadores humanos. Em muitos casos, as câmeras são utilizadas apenas para registro de imagens, com análise posterior para fins de identificação ou produção de evidências. Neste contexto, este trabalho utiliza a plataforma Edge Impulse para avaliar de forma exploratória a implementação de inteligência artificial voltada ao reconhecimento automático de situações de risco. O sistema proposto visa alertar operadores e autoridades de segurança, além de possibilitar a expansão de programas de monitoramento sem a necessidade de ampliação das salas de controle ou do quadro de pessoal.

Palavras-chave: Sistema de Vigilância; Inteligência Artificial; Edge Impulse; Aprendizado Profundo; Segurança Pública.

Development of an AI Model for Detecting Risk Situations in Surveillance Camera Images

Abstract: Security cameras have become increasingly common in everyday life worldwide. Along with the growth in the number of devices, there has been significant technological advancement, resulting in improved image quality and the incorporation of features such as night vision, audio capture, cloud storage, and remote communication, enabling real-time remote monitoring. Major cities have integrated video surveillance systems into their public safety strategies. Despite these advances, traditional monitoring still presents notable limitations, particularly the reliance on a large number of human operators. In many cases, cameras are used solely for image recording, with subsequent analysis conducted for identification purposes or as evidentiary support. In this context, this study employs the Edge Impulse platform to evaluate the implementation of artificial intelligence aimed at the automatic recognition of hazardous situations. The proposed system seeks to alert operators and security authorities, while enabling the expansion of monitoring programs without the need to increase control room capacity or personnel.

Keywords: Surveillance System; Artificial Intelligence; Edge Impulse; Deep Learning; Public Security.

1. INTRODUÇÃO

Segundo estimativas de 2023, existem algo em torno de 1,5 bilhão de câmeras de vigilância em operação no mundo [1]. Em um esforço para tornar os centros urbanos mais seguros, sistemas que empregam inteligência artificial (IA) têm sido cada vez mais utilizados dentro do conceito de cidades inteligentes (*smart cities*), de diferentes formas, auxiliando na detecção e prevenção de situações

de risco. Essas aplicações abrangem desde a análise de sons e imagens até a integração com outros dados e sensores disponíveis no ambiente urbano [2].

No Brasil inciativas públicas já existem para ampliar o uso de câmeras no combate ao crime. Em São Paulo um programa estadual chamado Muralha Paulista, que através de uma série de convênios interliga os sistemas de monitoramento de prefeituras, associações de moradores e empresas, é uma dessas apostas [3].

Este artigo explora a melhor forma de treinar um modelo de IA, para apresentar os melhores resultados. O objetivo é desenvolver e testar uma inteligência artificial para reconhecimento de pessoas em situação de perigo, utilizando os recursos disponibilizados pelo Edge Impulse© [4], uma plataforma pública usada para criar, testar e integrar IAs.

O intuito é de que a IA seja capaz de identificar pessoas em uma situação de risco, tais como; assalto ou sequestro; através do gesto de levantar as duas mãos, que é uma ação comum, e muitas vezes automática, em situações desse gênero. A Figura 1 exibe um exemplo de sinal enquadrado nessa categoria.

Figura 1. Pessoa realizando gesto que identifica situação de perigo.

Nota: Os rostos apresentados neste artigo foram descaracterizados para preservar as identidades das pessoas.

2. METODOLOGIA

O processo de treinamento do modelo de IA de reconhecimento de imagens do Edge Impulse, consiste em três fases principais [5]: a) desenvolvimento de uma biblioteca, com imagens das situações ou objetos que se quer identificar, b) Identificação dos elementos de interesse através do processo de demarcação via *labels*, e c) treinamento da IA. Estes itens estão mostrados no fluxograma da Figura 2.

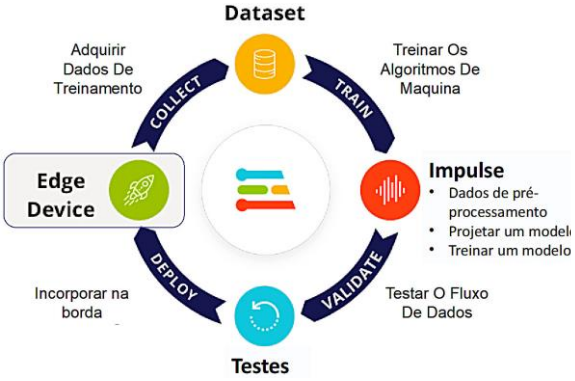

Figura 2. Fluxo de trabalho do *Edge Impulse*.

Para treinar uma IA em reconhecimento de imagens se recomenda, que a biblioteca ao ser montada, leve em conta a aplicabilidade do projeto. Em casos mais sérios como pesquisa se entende que o melhor, é montar uma coleção de imagens com maior qualidade e um nível de detalhes que ajude no estudo [6].

As imagens primeiramente usadas para ensinar o modelo de IA em como identificar situações de risco, foram coletadas fotografando cenas de filmes e programas de televisão disponíveis no Youtube, além de outras fontes da internet como *Google* e *Bing* imagens.

Durante o procedimento de treinamento o objeto de interesse é identificado através de 1 (um) *label* que é a forma como a IA aprende. Para que a IA identifique a situação de risco, foi utilizado um *label* nomeado Rendido". Nas Figura 3 e 4 se pode observar o processo de classificação através do *label* "Rendido". Como as IAs funcionam de forma binária (verdadeiro ou falso), para treinar de forma eficiente é preciso fornecer para o seu sistema referências de como seriam situações de não risco.

Figura 3. Imagem utilizada para montagem da biblioteca situação "Rendido".

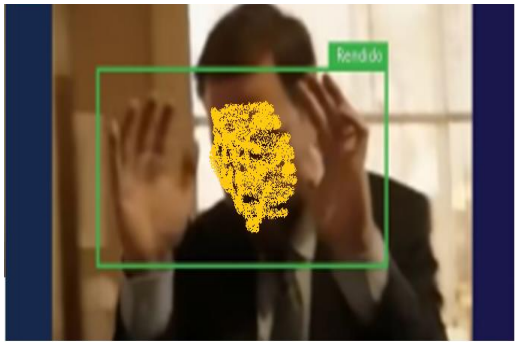

Figura 4. Imagem utilizada para montagem da biblioteca situação rendido com *label* aplicado.

Para que a IA tivesse essas referências, foram coletadas imagens de filmes e outras fontes da internet que também passaram pelo processo de fixação do *label*, nesse caso foi criado um outro *label* chamado "Normal". As Figuras 5 e 6 mostram o processo de implementação do *label* "Normal", apresentando a imagem original e rotulada para treinamento.

Após a coleta das imagens e processo de designação dos *labels* a IA foi treinada. Os primeiros testes mostraram uma confiabilidade baixa, inferior a 30%. A má performance deveu-se ao fato da biblioteca inicialmente usada conter um número muito pequeno de amostras. Portanto, quanto mais a biblioteca foi sendo expandida, observou-se maior confiabilidade e precisão do modelo de IA. O melhor resultado alcançado pelos testes foi com uma biblioteca de **437** imagens sendo **197** representando situações de não risco e **244** representando situações de risco.

Figura 5. Imagem utilizada para montagem da biblioteca situação normal.

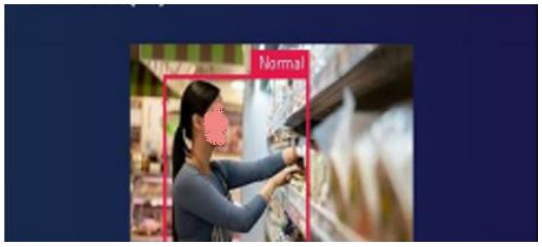

Figura 6. Imagem utilizada para montagem da biblioteca situação normal com *label* aplicado.

Como as imagens podem conter mais de um label, ao final foram atribuídos 607 labels, sendo 288 "Rendido" e 319 "Normal". Nenhuma imagem com múltiplos labels recebeu categorias diferentes. A *performance* da IA alcançou 58,6%, conforme apresentado na Figura 7.

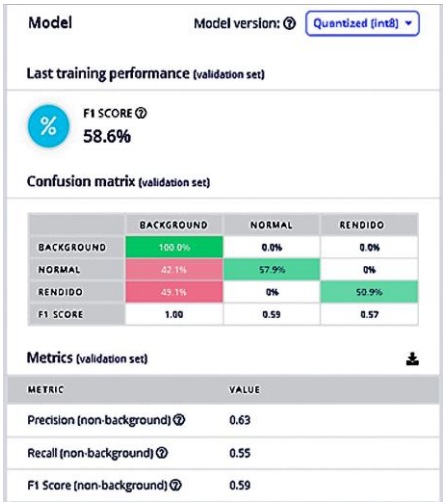

Figura 7. *Performance* do modelo de IA.

A Figura 8 apresenta o digrama de dispersão, que mostra o índice de dispersão entre as imagens, ou seja, o quão bem a IA conseguiu diferenciar os cenários apresentados. É possível observar que existem imagens pertencentes a categoria do *label* rendido que são óbvias para IA, enquanto outras possuem muitas similaridades com imagens da categoria normal.

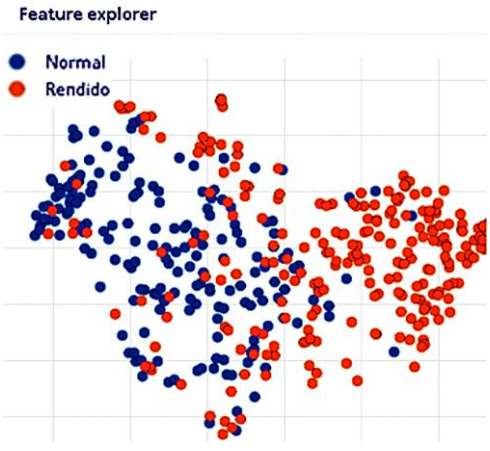

Figura 8. Diagrama de dispersão da IA no *Edge Impulse*.

A IA desenvolvida foi testada utilizando a câmera de um *smartphone*. Na Figura 9 se pode ver como a IA performou durante os testes. Durante tais testes, o primeiro modelo de IA desenvolvido, falhou em identificar a situação de perigo. Foi levantada a hipótese, de que a IA não conseguiria identificar propriamente a situação, por conta do fundo poluído, que era diferente em cada imagem de treino da biblioteca.

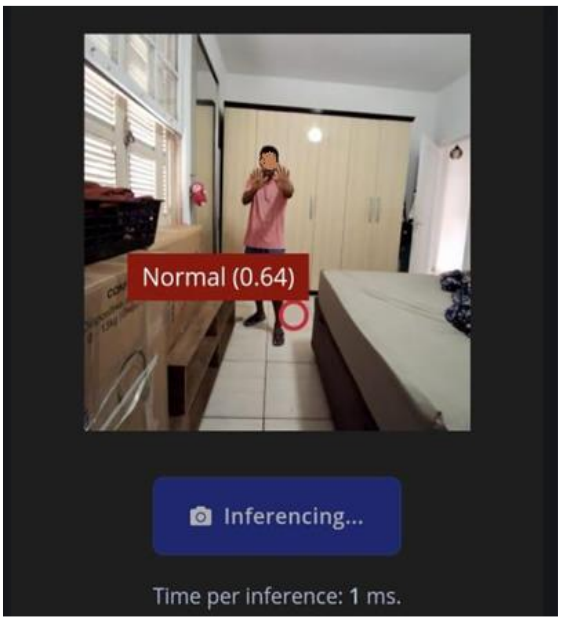

Figura 9. Resultado do teste de *performance* da IA.

Para tentar melhorar a qualidade da interpretação, foi então desenvolvida um novo modelo de IA com a mesma função, mas usando uma biblioteca diferente. E assim, o *label* Cabeça foi desenvolvido para identificar um aspecto comum, de ambos os tipos de imagens, que é a cabeça das pessoas que podem estar ou não em situação de risco. Este novo *label* funciona em adição aos outros 2 (dois) *labels* já existentes (Rendido, Normal).

As imagens dessa biblioteca tiveram o fundo excluído e foram reinseridas em um fundo neutro, que corresponde a uma parede cinza. Este procedimento foi feito com o auxílio de uma ferramenta *online* chamada Erase.bg©. O processo pode ser acompanhado de forma completa nas Figuras 10 a 12.

Assim como foi feito no primeiro modelo de IA gerado, para essa também foi necessário preencher a biblioteca com cenários de não risco. Parte da biblioteca foi preenchida com imagens que passaram pelo mesmo processo de edição efetuado nas Figuras 9, 10 e 11, exibidas anteriormente.

Figura 10. Imagem base original para *label* "Rendido".

Figura 11. Imagem de pessoa em situação de "Rendido", já com o fundo retirado.

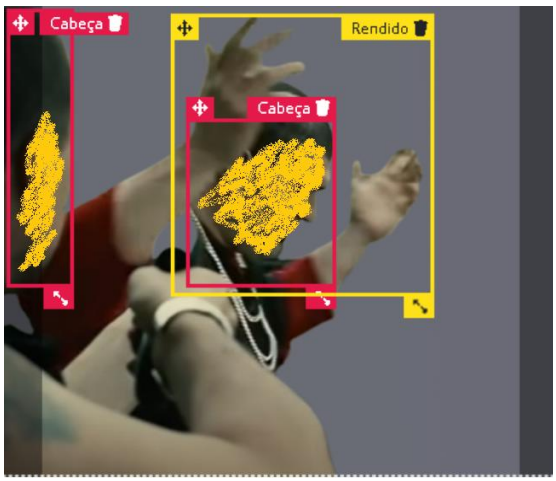

Figura 12. Imagem utilizada para montagem da biblioteca com pessoa em situação de rendido, já com o fundo retirado e os *labels* posicionados

Nas Figuras 13 e 14 é possível observar respectivamente, a imagem formatada e com os *labels* (Normal e Cabeça) já inseridos. Foi feito um *Upload* de 200 (duzentas) imagens com metade envolvendo situações de risco nas quais foram aplicados os *labels* "Rendido" e "Cabeça". E a outra metade com situações cotidianas em que foram aplicados os *labels* "Normal" e "Cabeça".

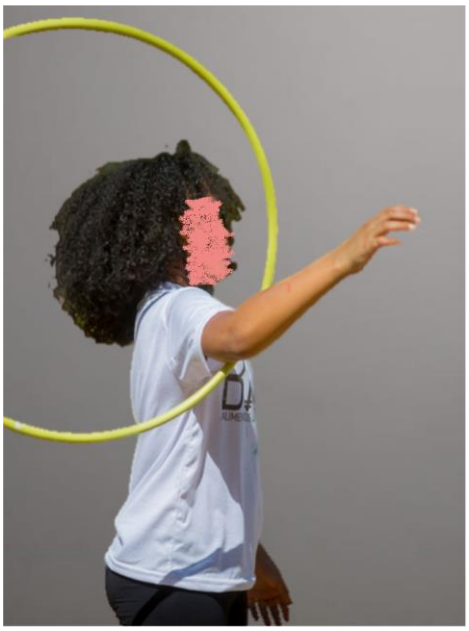

Figura 13. Imagem para montagem da biblioteca com pessoa em situação normal, com o fundo retirado.

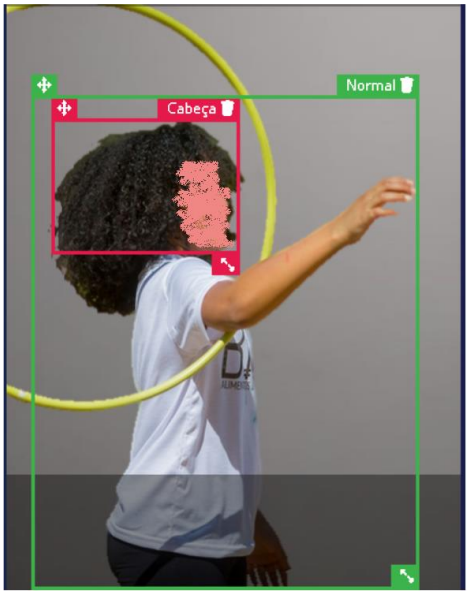

Figura 14. Imagem utilizada para montagem da biblioteca com pessoa em situação normal, já com o fundo retirado e os *labels* posicionados.

Ainda nas Figuras 11 e 13, pode-se observar que assim como na IA anterior, os *labels* inseridos para designar situações de risco foram aqueles nos quais o indivíduo se encontra com ambas as mãos levantadas.

Situações normais são quaisquer situações na quais as mãos dos indivíduos não realizam o sinal de risco (ambas as mãos para o alto). Para determinar os cenários de normais, basta selecionar, a área envolvendo as pessoas. Enquanto para o cenário de rendido, se seleciona a área com as mãos levantadas, incluindo a cabeça do indivíduo, que é demarcada em ambos os cenários.

O *label* cabeça é um denominador comum em ambos os cenários que têm a intenção de reduzir a confusão da IA.

A Figura 15, mostra o diagrama de dispersão apontando a capacidade da IA de diferenciar entre os cenários estipulados na biblioteca.

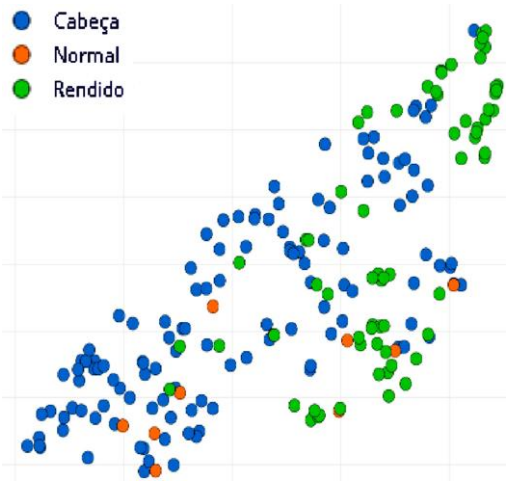

Figura 15. Diagrama de dispersão da IA no Edge Impulse.

Em seguida o modelo de IA foi testada de forma prática, com o auxílio de um *smartphone*. A Figura 16 mostra um *print* da tela durante o teste. Nesta figura, é possível observar que a pose é similar a aplicada no teste do primeiro modelo de IA. Já o fundo é similar ao utilizado no treinamento do segundo modelo de IA.

3. RESULTADOS

O primeiro modelo de IA falhou em detectar a situação de risco. A causa mais provável é que o fundo poluído das imagens de teste tenha dificultado a leitura, comprometendo a precisão do modelo.

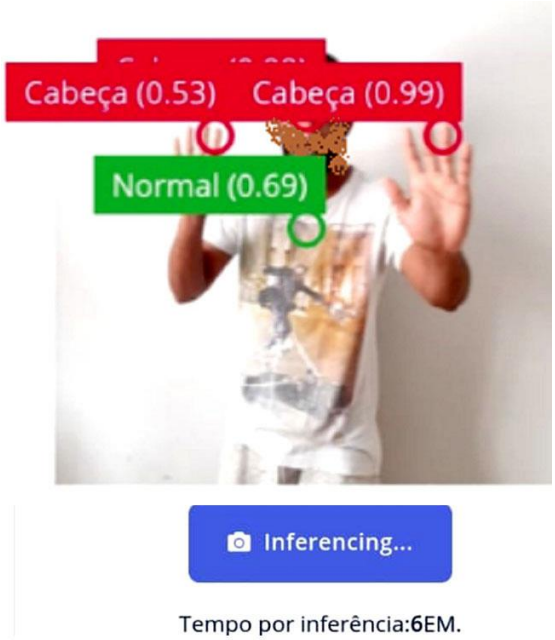

Figura 16. Resultado do teste de *performance* da IA.

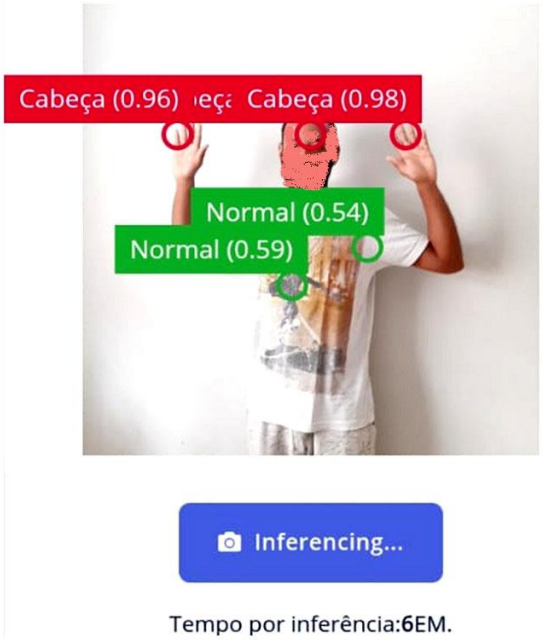

Figura 18. Teste de validação no *smartphone* do segundo modelo de IA.

Na Figura 17, é possível observar que a IA identificou pontos que não correspondem aos *labels* definidos. A hipótese levantada é que o modelo tenha associado essas detecções a padrões de cores presentes no conjunto de dados utilizado durante o treinamento.

O segundo modelo, também falhou em reconhecer a situação de perigo, designada pelo *label* rendido. Apesar de falhar no propósito principal é possível observar uma melhora nos resultados.

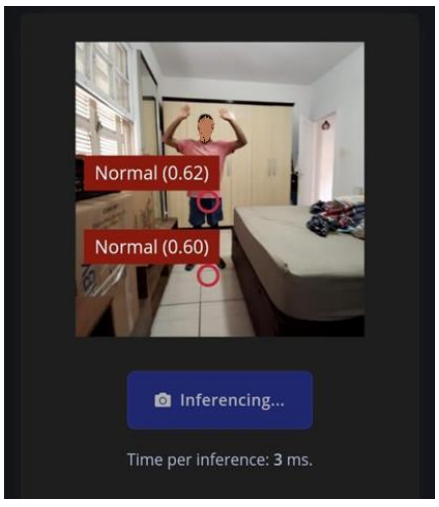

Figura 17. Teste de validação no smartphone do primeiro modelo de IA.

Comparando a Figura 16 com a Figura 18, vê-se que todos os *labels* focaram na pessoa usada no teste, enquanto nas Figuras 9 e 17 se percebe que o solo também foi reconhecido pela IA.

A falta de um fundo na segunda IA, permitiu ao modelo focar na pessoa com menos distrações. Mas a IA, ainda falha em ver o cenário de forma mais aberta, identificando mais cores e detalhes, ao invés dos contextos desejados.

A Figura 19, exibe a configuração final do modelo de IA no Edge Impulse. É possível observar o número de ciclos de treinamento (60), valor alto o bastante para garantir convergência sem gerar sobre ajuste (*overfitting*) e o registro de 9216 *features* identificados pelo modelo de IA na camada de entrada.

Também é possível observar a taxa de aprendizado utilizada, correspondendo a 0,001 que está no ponto de equilíbrio entre velocidade de convergência e estabilidade de treinamento.

4. CONCLUSÃO

O principal desafio de desenvolver um modelo de IA, para função objeto deste estudo reside na dificuldade de fazer o modelo "entender o contexto" de uma situação em vez de focar nos detalhes da imagem visualizada. Mesmo sendo situações completamente diferentes, uma situação

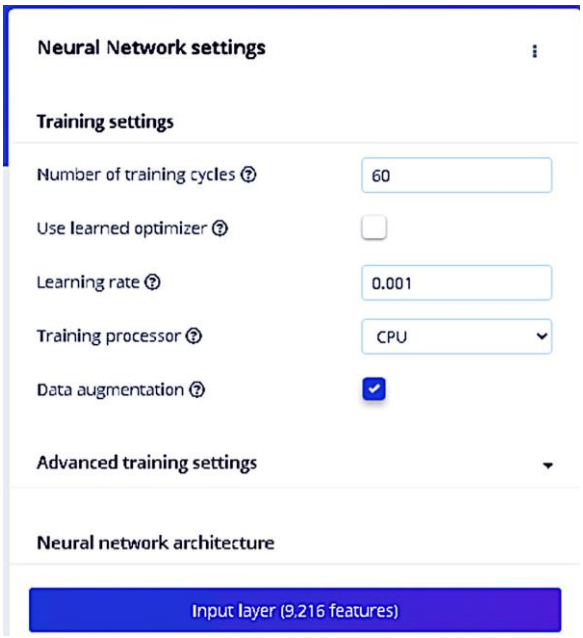

Figura 19. Configuração da IA no *Edge Impulse*.

de risco e uma de normalidade têm várias coisas em comum. Assim como imagens de situações que deveriam se enquadrar no mesmo *label* podem ser muito diferentes. Aparência das pessoas envolvidas, coloração das roupas e posição delas em relação a câmera (frontal, costas ou perfil). Entre a composição da biblioteca e os testes, existe um problema de perspectiva e qualidade das imagens, que também precisa ser levado em consideração. As imagens foram obtidas a partir de fontes diferentes, com qualidades variadas e filmadas de ângulos distintos.

Para trabalhos futuros, recomenda-se aproximar os testes das condições reais de operação. A definição de um cenário específico, com função real como uma loja ou escritório, gera uma melhora dos dados coletados. Sugere-se também a fixação da câmera em uma posição determinada, operando por um período prolongado, de modo a registrar o ambiente em condições rotineiras. Posteriormente, imagens contendo situações de risco devem ser inseridas mantendo o mesmo fundo e proporções dos cenários normais, garantindo maior consistência entre os dados de treinamento e teste. A simulação de situações de risco com o auxílio de atores pode contribuir para aumentar o realismo do conjunto de dados. Além disso, a ampliação do número de *labels*, para incluir a detecção de armas brancas e de fogo, é uma possibilidade a ser considerada para o aprimoramento do sistema.

REFERÊNCIAS

[1] JUDE, M. *IDC MarketScape: Worldwide Video Surveillance Camera 2023 Vendor Assessment*. Framingham, MA: International Data Corporation, 2023.

[2] SRIVASTAVA, Shweta; BISHT, Aditua; NARAYAN, Neetu. Safety and security in smart cities using artificial intelligence — a review. In: *Proceedings of the 2017 IEEE CONFLUENCE: The Next Generation Information Technology Summit*, Noida, India, 2017. p. 1–6. DOI: https://doi.org/10.1109/CONFLUENCE.2017.7943136. Disponível em: https://ieeexplore.ieee.org/document/7943136/authors#authors. Acesso em: 12 jan. 2026.

[3] BOL. São Paulo avança em integração de sistema de hipervigilância. *BOL Notícias*, São Paulo, 27 jan. 2024. Disponível em: https://www.bol.uol.com.br/noticias/2024/01/27/sao-paulo-avanca-em-integracao-de-sistemas-de-hipervigilancia.htm. Acesso em: 05 jan. 2026.

[4] HYME, Shawn; BANBURY, Colby R.; SITUNAYAKE, Daniel; ELIUM, Alexander; WARD, Carl; KELCEY, Matthew; BAAIJENS, Mathijs; MAJCHRZYCKI, Mateusz; PLUNKETT, Jenny; TISCHLER, David; GRANDE, Alessandro; MOREAU, Louis; MASLOV, Dmitry; BEAVIS, Artie; JONGBOOM, Jan; REDDI, Vijay Janapa. Edge Impulse: an MLOps platform for Tiny Machine Learning. *Proceedings of Machine Learning and Systems*, v. 5, p. 254–268, 2023.

[5] PAULO JUNIOR, F. S.; ALVES, B. S.; DA CRUZ, D. O.; DE CARVALHO JUNIOR, A.; VARELLA, W. A.; DA SILVA FILHO, J. I. Sistema especialista por aprendizado de máquina para análise de tuberculose. *Unisanta Science & Technology*, v. 14, n. 1, p. 39–45, 2025. DOI: https://doi.org/10.5281/zenodo.17518594.

[6] EAILAB. Roteiro para criação de dataset de imagens para modelos de aprendizagem profunda. 2024. Disponível em: https://eialab.labmax.org/2024/09/24/roteiro-para-criacao-de-dataset-de-imagens-para-modelos-de-aprendizagem-profunda/. Acesso em: 12 jan. 2026.

