

Um Sistema Móvel Inteligente com Computação na Borda para Atendimento de Ocorrências de Segurança Pública

Stefano Loss, Karine Costa, Alison Hedigliranes, Pedro Varela,
João Pedro Limão, Thais Batista, Aluizio Rocha Neto, Daniel Sabino,
Nélio Cacho, Frederico Lopes, Everton Cavalcante

Universidade Federal do Rio Grande do Norte, Natal-RN, Brasil

{momoloss10, alison.hedigliranes, pedro.varela1289}@gmail.com,
{jplima012, thaisbatista, aluiziorocha}@gmail.com,
{karine.piacentini, daniel, fred}@imd.ufrn.br,
neliocacho@dimap.ufrn.br, everton.cavalcante@ufrn.br

Resumo. Este artigo apresenta um sistema móvel que integra funcionalidades para o gerenciamento de ocorrências de segurança pública, desde o envio de notificações de incidentes para viaturas policiais até o processamento inteligente, em tempo real, de vídeos captados pelos dispositivos móveis dos policiais em atendimento. Esse processamento é realizado na borda da rede e gera alertas para melhorar a consciência situacional dos policiais em campo. O artigo também avalia o desempenho de um algoritmo de reconhecimento de faces em condições de conectividade limitadas, visando encontrar um balanço entre resolução de vídeo, precisão do reconhecimento e consumo de energia e de banda de rede em dispositivo móvel.

Abstract. This paper presents a mobile system that integrates functionalities for managing public safety incidents, from sending incident notifications to police vehicles to intelligent, real-time processing of videos captured by the mobile devices of police officers on duty. This processing is performed at the network's edge and generates alerts to improve situational awareness of officers in the field. The paper also evaluates the performance of a face recognition algorithm under limited connectivity conditions, aiming to find a balance between video resolution, recognition accuracy, and energy and network bandwidth consumption on a mobile device.

1. Introdução

A segurança pública é uma grande preocupação dos brasileiros diante dos altos índices de violência [Fórum Brasileiro de Segurança Pública 2024]. No entanto, muitas vezes ela esbarra na ineficiência do atendimento de ocorrências devido a problemas como recursos limitados, uso de tecnologias defasadas e insuficientes para gerenciar a complexidade dos incidentes, comprometendo a atuação policial em campo. As ocorrências de segurança são recebidas por uma Central de Atendimento e Despacho (CAD) que as despacha às viaturas policiais. Por outro lado, os policiais em campo lidam com múltiplos incidentes que podem exigir avaliação rápida e adaptação em situações inesperadas, a exemplo de reféns ou indivíduos armados. Diante dessas complexidades, o acesso imediato a informações críticas, como o histórico criminal de um suspeito, pode influenciar significativamente as estratégias de resposta e aprimorar a *consciência situacional* (CS), fundamental para que

policiais ajam com precisão e segurança frente a incidentes. A CS refere-se à capacidade de perceber fatores ambientais, compreender seu significado e antecipar situações futuras, visando apoiar a tomada de decisão em cenários dinâmicos [Endsley 1995].

Com o avanço das tecnologias móveis, da computação na borda (*edge computing*), da inteligência artificial (IA) e das redes 5G, surge uma oportunidade para transformar o atendimento de ocorrências, permitindo uma resposta mais ágil, precisa e com melhor aproveitamento dos recursos disponíveis e maior segurança para policiais e cidadãos. A computação na borda permite que os dados coletados por dispositivos móveis sejam processados localmente, em vez de serem enviados para um servidor central, reduzindo a latência. Aliada à IA, a computação na borda torna possível o processamento inteligente de vídeos em tempo real que pode ser usado para detectar e classificar automaticamente objetos de interesse nos vídeos, incluindo reconhecimento de face [Vieira et al. 2024], o que proporciona CS aos policiais em campo. Além disso, a computação na borda garante que, mesmo em ambientes com conectividade de rede instável ou indisponível, o processamento de dados e a tomada de decisões em tempo real continuem sendo possíveis, permitindo a continuidade de operações essenciais e a resposta rápida a eventos urgentes.

Nesse cenário, este trabalho apresenta o CAD-Viatura, um sistema móvel que visa otimizar o atendimento de ocorrências e aprimorar a CS dos policiais em campo. Para isso, o CAD-Viatura explora a sinergia entre a computação na borda, comunicação 5G e os avanços da IA para análise inteligente de vídeo em tempo real. O sistema é composto: (i) pela Central de Atendimento e Despacho (CAD) utilizada no Estado do Rio Grande do Norte desde 2020, interligando 14 agências de segurança pública a nível estadual e de alguns municípios; (ii) o *framework* SAALSA (*Stream AI Analytics for Live Situational Awareness*) [Loss et al. 2025], que fornece processamento inteligente de vídeo na borda da rede, e; (iii) elementos da plataforma FIWARE¹, um *middleware* de código aberto para desenvolvimento de aplicações de cidades inteligentes. O CAD-Viatura facilita a coleta e processamento inteligente de dados de múltiplas fontes e o registro operacional das ações, proporcionando uma resposta rápida a situações de risco, mesmo em cenários sem conectividade de rede. O sistema permite despacho silencioso, acesso a informações do incidente, monitoramento remoto e reconhecimento de elementos de interesse (por exemplo, pessoas) usando técnicas de IA.

Este artigo apresenta ainda uma avaliação do impacto da resolução na precisão do algoritmo de reconhecimento de face empregado no CAD-Viatura, de forma a avaliar se o reconhecimento se mantém preciso mesmo em condições de conectividade variadas. Essa análise é relevante uma vez que a transmissão de vídeo em tempo real utiliza o protocolo WebRTC [Mahmoud and Abozariba 2025] e implementa um controle de congestionamento², resultando em uma variação de resolução de acordo com a largura de banda disponível. Também foram mensurados o consumo energético e o total de dados trafegados para diferentes resoluções visando encontrar o balanço ideal entre precisão, consumo de banda de rede e de energia, o que é relevante no contexto do CAD-Viatura, que precisa operar eficientemente nos dispositivos durante longos períodos. Altas resoluções de vídeo tendem a melhorar a precisão do reconhecimento, porém aumentam o consumo de banda e de energia dos dispositivos [Zhang et al. 2016].

¹<https://www.fiware.org/catalogue>

²https://doc-kurento.readthedocs.io/en/latest/knowledge/congestion_rmcat.html

O restante deste artigo está estruturado da seguinte forma. A Seção 2 apresenta brevemente os elementos básicos que compõem o CAD-Viatura. A Seção 3 detalha a solução proposta em termos de arquitetura e de implementação. A Seção 4 discute resultados experimentais referentes à configuração ideal da resolução do vídeo para processamento na borda, buscando equilibrar precisão no reconhecimento facial com consumo de energia e uso de banda. A Seção 5 aborda trabalhos relacionados. A Seção 6 traz as considerações finais.

2. Elementos Básicos

2.1. FIWARE: Orion Context Broker e Kurento

O *Orion Context Broker* (ou simplesmente Orion) é o componente central da FIWARE, atuando como *broker* de contexto responsável por gerenciar e distribuir em tempo real informações sobre entidades através do modelo *publish-subscribe*. O Orion permite comunicação assíncrona e desacoplada entre componentes sem dependência direta entre si, além de fazer uso de requisições HTTP/REST com dados em formato JSON segundo a especificação NGSI.

O *Kurento* é outro componente da FIWARE que oferece um servidor de mídia e APIs para criação de aplicações avançadas de vídeo, com suporte a funcionalidades como comunicação em grupo, transcodificação, gravação, mixagem, transmissão e roteamento de fluxos audiovisuais, além de ser compatível com protocolos de tempo real como o WebRTC. Com isso, o Kurento torna-se adequado para aplicações que exigem o monitoramento, processamento ou distribuição de vídeo em tempo real.

2.2. O Framework SAALSA

O principal objetivo do SAALSA é permitir a análise inteligente, em tempo real, de fluxos de dados provenientes de diversas fontes. Sua arquitetura (Figura 1) é estruturada em três camadas, cada uma desempenhando um papel distinto no processamento e análise de dados. A *Camada de Fontes de Dados* captura dados brutos de diversos formatos como vídeo, áudio e geolocalização provenientes de câmeras fixas, drones ou dispositivos móveis. A *Camada de Borda* processa os dados próximos à fonte para reduzir a latência e aliviar a carga de transmissão para a nuvem. Essa camada inclui componentes para: acesso aos dados de diversos dispositivos (*Gerenciador de Dispositivos*); gerenciamento dos fluxos multimídia (*Serviço de Streaming*); processamento inteligente para tomada de decisões (*Serviço de IA*); armazenamento local de dados multimídia e registros de *logs* (*Serviço de Armazenamento*); supervisão da execução das tarefas inteligentes nos nós de borda sincronizado com a nuvem (*Gerenciador de Alocação de Recursos*); monitoramento do uso de recursos nos nós de borda para garantir eficiência operacional, e; integração contínua com serviços em nuvem (*Broker de Comunicação*).

A *Camada de Nuvem* gerencia o processamento distribuído entre os nós de borda, fornece serviços analíticos avançados, armazenamento permanente e suporte à visualização global de operações. O *Orquestrador de Processamento de Streaming* coordena a execução de tarefas otimizando a alocação de recursos. O *Gerente de Serviço de IA* acompanha os serviços de inteligência em execução nos nós. O *Gerenciador de Nós de Processamento* monitora a localização dos nós e seus recursos de *hardware* para facilitar

a atribuição eficiente de tarefas. O *Gerenciador de Repositório de Tarefas de Processamento* armazena descrições de tarefas de processamento, possibilitando sua implantação e reutilização eficiente. Vale destacar que os dispositivos como câmeras e *smartphones* têm duplo papel no SAALSA, podendo atuar tanto como fontes de dados quanto como nós de processamento de IA em borda, maximizando a eficiência dos recursos.

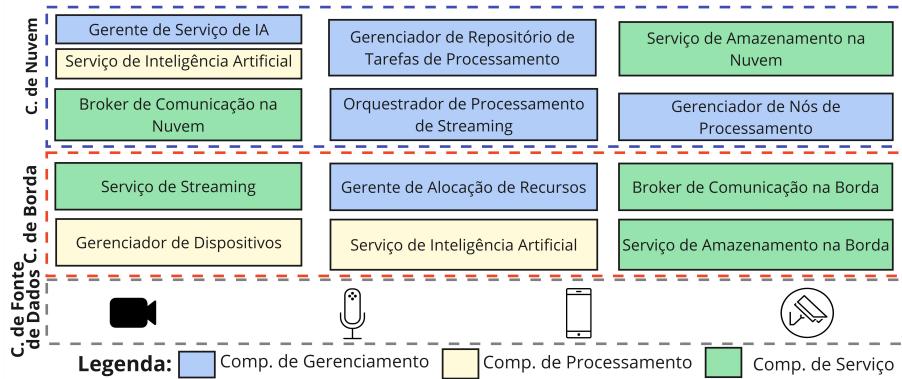


Figura 1. Arquitetura do SAALSA.

3. O CAD-Viatura

O CAD-Viatura é um sistema para dispositivos móveis que visa otimizar o atendimento policial, integrando o gerenciamento de ocorrências e análise de vídeos de operações em tempo real com IA e na borda da rede. A solução visa reduzir a dependência da comunicação tradicional com viaturas via rádio, promove um fluxo mais rápido e organizado das informações entre as equipes envolvidas e permite melhorar a CS por meio do reconhecimento automático de objetos de interesse nas transmissões de vídeo.

Através do CAD-Viatura, o policial recebe informações do incidente registradas em um sistema CAD pelo centro operacional (CO) via ligação para o número 190. Durante o atendimento, o policial pode atualizar o status da ocorrência. No local do incidente, o sistema possibilita a transmissão de *streamings* em tempo real de vídeos captados pelo *smartphone* acoplado ao uniforme do policial. Desta forma, a equipe de apoio e o CO pode realizar o acompanhamento remoto da ocorrência quando houver comunicação de rede disponível. Isto é útil em situações como cercos a residências ou intervenções táticas, onde policiais posicionados em diferentes pontos podem compartilhar suas perspectivas visuais, melhorando a CS e coordenação da equipe.

As imagens captadas pelos *smartphones* dos policiais podem ser processadas por IA diretamente na unidade embarcada presente na viatura, aproveitando a computação na borda, ou, alternativamente, enviadas para a nuvem. Esse processamento possibilita a geração de alertas automáticos sobre objetos de interesse relevantes para a operação, como a identificação de indivíduos procurados ou de veículos irregulares. A emissão de alertas em tempo real permite a adaptação da abordagem da equipe policial com base nas informações identificadas, ou seja, potencializando a CS.

3.1. Arquitetura e Implementação

A Figura 2 apresenta os componentes da arquitetura do CAD-Viatura. Essa arquitetura integra uma instanciação do SAALSA, um sistema CAD e o Orion da FIWARE.

O Orion é responsável pela comunicação assíncrona e descentralizada entre os componentes internos do SAALSA, além da comunicação entre o CAD-Viatura com o sistema CAD, garantindo um fluxo contínuo e eficiente de dados críticos para o atendimento de ocorrências. O Orion foi escolhido por viabilizar uma comunicação robusta e permitir rápida disseminação de eventos, como notificações de reconhecimento de objetos de interesse, favorecendo decisões ágeis baseadas em dados atualizados.

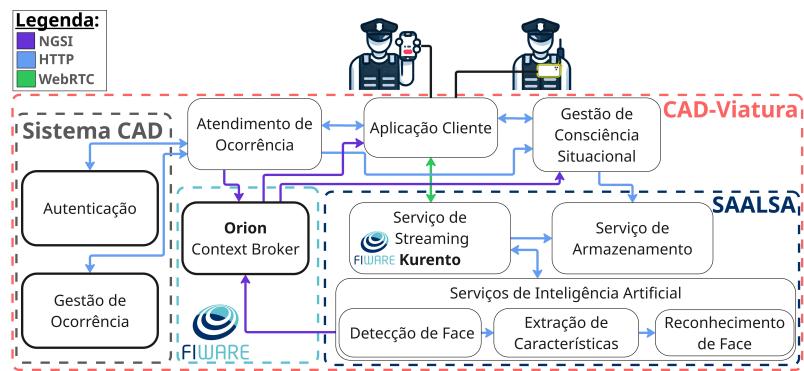


Figura 2. Arquitetura de implementação do CAD-Viatura.

Os policiais utilizam a *Aplicação Cliente* (AC) em um dispositivo móvel, que interage com os componentes de *Gestão de Consciência Situacional* (GCS), *Atendimento de Ocorrência* (AO), com o *Serviço de Streaming* (SS) do SAALSA e com o Orion. Os componentes de AO e o de GCS são responsáveis pela interação com o CAD e com o SAALSA, respectivamente. A AC tem uma comunicação direta com o SS para a transmissão de vídeo em tempo real, usando o protocolo WebRTC, enquanto outras comunicações síncronas são feitas via protocolo HTTP. O componente de AO é responsável pelo gerenciamento do atendimento de ocorrências através da interação com o CAD, em particular com o componente de *Gestão da Ocorrência* (GO). O componente de GO realiza funções como o despacho de ocorrências para viaturas e permite ao componente de AO consultar e atualizar o progresso e o encerramento das ocorrências associadas à viatura do policial.

O componente de GCS é responsável por ampliar a CS dos policiais em campo, fornecendo informações relevantes à tomada de decisões durante operações. Para isso, esse componente busca ativamente dados de interesse a cada operação, considerando sua localização (via GPS), seus objetivos e informações detectadas no entorno. Esse componente integra dados provenientes do *Serviço de IA* (SIA) e do histórico do atendimento da ocorrência, notificando o policial com atualizações que impactam sua atuação. Além disso, o componente de GCS mantém comunicação com o Orion para processar eventos da *pipeline* de reconhecimento de face em vídeos, gerenciando todo o fluxo, desde a transmissão até o retorno das informações ao policial, permitindo que tenham acesso a *insights* relevantes em tempo real, otimizando a tomada de decisão e aumentando a segurança e eficácia das operações. Por sua vez, o *Serviço de Armazenamento* (SA) registra a cronologia dos eventos gerados pelo componente de GCS, bem como os vídeos capturados durante a operação.

A computação na borda é especialmente importante no contexto do CAD-Viatura, pois permite que o processamento de vídeo ocorra próximo à fonte de dados, ou seja, diretamente nos dispositivos embarcados nas viaturas, reduzindo significativamente a

latência e garantindo respostas mais rápidas em situações críticas. Esse processamento local também diminui a dependência de conectividade com a nuvem, assegurando que funcionalidades essenciais, como a detecção e o reconhecimento de face, continuem operacionais mesmo em ambientes com conexão instável ou inexistente. Além disso, a execução das tarefas computacionais na borda contribui para a economia de largura de banda e melhora a escalabilidade do sistema, evitando a sobrecarga de redes e servidores remotos.

A *pipeline* de reconhecimento de faces implementado pelo SIA é composta por três etapas: detecção, extração de características e reconhecimento de face. O CAD-Viatura utiliza modelos de alto desempenho que atendem a restrições de latência e de recursos, mantendo robustez sob condições ambientais variáveis, típicas de cenários de segurança pública. A detecção de faces realiza a extração de faces em uma *bounding box* usando o modelo NVIDIA FaceNet³ [Redmon et al. 2016]. A extração de características das faces detectadas é feita por meio da rede neural convolucional FaceNet [Schroff et al. 2015], formando um vetor de características conhecido como *embedding*. Esse *embedding* mapeia imagens de face para um espaço euclidiano compacto no qual as distâncias entre os vetores correspondem diretamente a uma medida de similaridade facial. Por fim, o reconhecimento de face compara o *embedding* extraído com outros armazenados em um banco de dados vetorial. O indivíduo é identificado como aquele cujos *embeddings* armazenados têm a menor distância euclidiana para o *embedding* da face detectada em execução.

Fluxo de Atendimento de uma Ocorrência. A Figura 3 destaca as interações entre os componentes envolvidos no atendimento da ocorrência. Para viabilizar o atendimento de ocorrências, primeiramente é necessário que uma ocorrência seja cadastrada no componente de GO do CAD. Assim, o fluxo de atendimento de ocorrência é iniciado pela GO, que publica a ocorrência em um canal pré-configurado no Orion correspondente à viatura designada (1). A AC recebe a notificação de ocorrência (2) e apresenta a chamada da ocorrência ao usuário (2.1). Após o aceite pelo usuário (3), a AC solicita a associação do usuário à ocorrência através do componente de AO (4), que encaminha essa solicitação ao componente de GO (4.1), que deve efetivar a alteração no CAD. Na sequência, a ocorrência é mostrada para o usuário (5) e enviada para o componente de GCS (6), que garante o gerenciamento da CS daquela ocorrência para aquele usuário. Com a ocorrência associada a ele, o usuário pode iniciar uma transmissão de vídeo (7), cuja configuração tem início no componente de AO (7.1) e é concluída pelo SS, detalhado na Figura 4. Por fim, o usuário solicita a alteração do status da ocorrência (8), ação refletida no componente de AO (8.1) que solicita a atualização ao componente de GO (8.2), levando à atualização do status da ocorrência no CAD.

Fluxo de processamento de vídeo na borda. A Figura 4 ilustra o fluxo de processamento de vídeo na computação na borda, com suporte do SAALSA, com destaque para as interações entre os componentes do *framework* que possibilitam o processamento de fluxos de dados com baixa latência, o que é essencial para a detecção e reconhecimento de objetos de interesse em tempo real. Para maior clareza, o diagrama representa os serviços de inteligência sendo processados em um único nó na borda, embora, na prática, possam ser distribuídos entre múltiplos nós do SIA. O processo inicia com a AC capturando vídeo e enviando o fluxo para o SS (1), que dispara duas tarefas paralelas, uma tarefa dearma-

³<https://catalog.ngc.nvidia.com/orgs/nvidia/teams/tao/models/facenet>

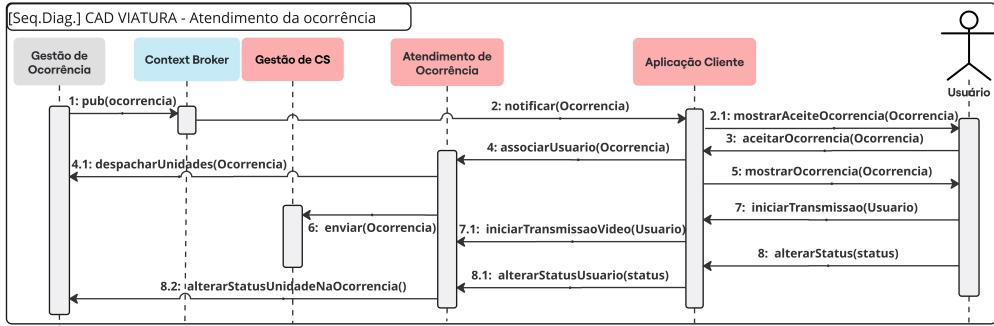


Figura 3. Sequência do atendimento de uma ocorrência.

zenamento (2.1) para salvar o fluxo no SA e o envio do fluxo para o SIA executar uma tarefa de detecção de face (2.2). Uma vez que uma face é detectada (3), o Orion aciona a tarefa de extração de características (4). Após a extração de *embeddings* (5), o Orion é acionado novamente para iniciar o reconhecimento da face (6). Após o reconhecimento, esse resultado é publicado no Orion (7), que notifica o componente de GCS (7.1). Após processar a notificação, o componente de GCS realiza um trabalho de decisão (7.2) buscando em suas bases de dados se aquele indivíduo possui antecedentes criminais, porte de arma ou restrições judiciais. Assim, é verificado se o que foi processado pelo SIA é relevante para ajudar na CS. Caso positivo, a AC é notificada com a decisão tomada (9). Por fim, após o SS receber a face reconhecida e processada (8), ele o encaminha para a AC para permitir o monitoramento de vídeo (8.1) e solicita o registro no SA (8.2). Em alguns cenários, computação na nuvem pode auxiliar no armazenamento de longo prazo e no processamento de algoritmos de IA mais complexos.

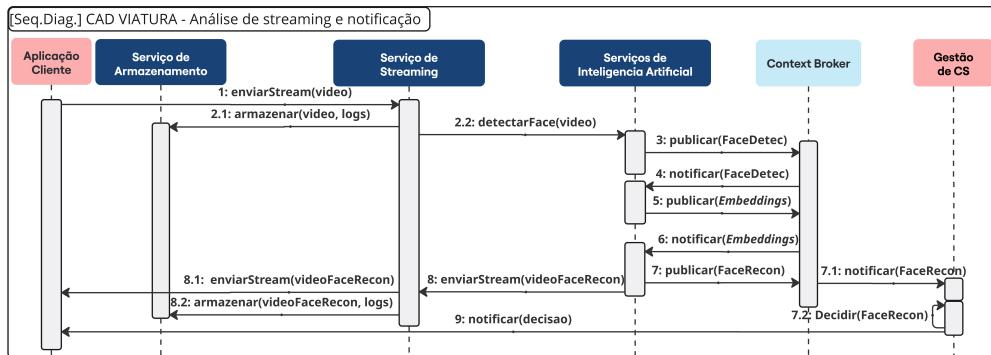


Figura 4. Sequência do processamento do streaming de vídeo de uma operação.

3.2. Interface para Atendimento de Ocorrências com o CAD-Viatura

O fluxo principal de atendimento e acompanhamento de ocorrências por meio do CAD-Viatura encontra-se na Figura 5. Esse fluxo inicia com a notificação de uma nova ocorrência, destacando-a com um alerta. Após a ocorrência ser aceita, a aplicação apresenta informações resumidas da ocorrência, bem como botões que alteram o status da ocorrência com a atividade em andamento (Figura 5a). No centro da tela de atendimento são apresentadas apenas as informações mais recentes do histórico da ocorrência, além de ser possível acessar o histórico completo do atendimento. No canto superior direito da tela, há ainda botões que dão acesso rápido às funcionalidades de mapa e opções de vídeo.

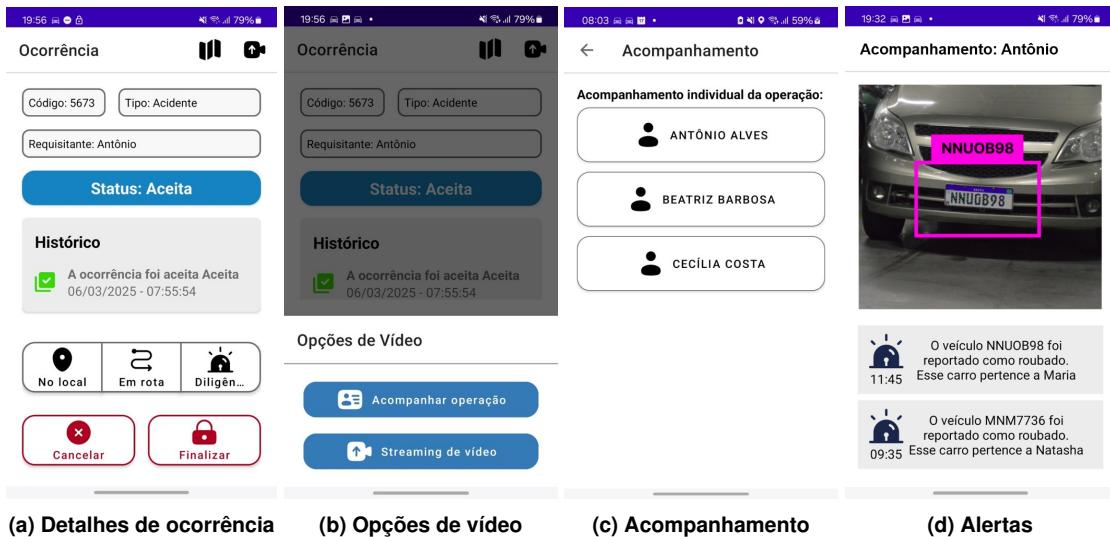


Figura 5. Streaming com alerta de reconhecimento.

Na tela de opções de vídeo (Figura 5b), acionada através do ícone de vídeo, são mostradas as opções referentes a vídeo oferecidas pelo CAD-Viatura. O botão *Streaming de Vídeo* permite aos policiais realizar a transmissão de vídeo em tempo real, enquanto o botão *Acompanhar Operação* oferece acesso à lista de usuários que estão transmitindo seus vídeos (Figura 5c). Ao selecionar um usuário, é possível visualizar o vídeo da sua transmissão, processado por IA (Figura 5d). É possível constatar que o monitoramento remoto de uma operação permite o envio de informações visuais imediatas, reduzindo ambiguidades e incertezas na interpretação do cenário operacional. Além disso, o processamento por IA dos *streamings* de vídeo da operação ocorre em todas as transmissões ativas, possibilitando a troca do agente acompanhado a partir da notificação de melhor ângulo de visão.

4. Avaliação

A avaliação do sistema de reconhecimento de face descrito na Seção 3.1 teve por objetivo analisar seu desempenho considerando diferentes resoluções de vídeo (1080p, 720p, 480p, 360p). Além disso, foi estimado o impacto do *streaming* em diversas resoluções no consumo de dados trafegados e bateria do dispositivo móvel, buscando encontrar um equilíbrio entre consumo de banda e energia, e desempenho do reconhecimento.

Conjunto de dados. O conjunto de dados utilizado nos experimentos foi construído a partir de dois conjuntos de dados públicos disponíveis para reconhecimento de face, o CelebA [Liu et al. 2015] e o *Labeled Faces in the Wild* (LFW) [Huang et al. 2007]. O CelebA contém 10.177 indivíduos, enquanto o LFW inclui 5.749 indivíduos. Foram selecionados aleatoriamente 5.000 indivíduos do CelebA e 800 indivíduos do LFW. Além disso, quatro voluntários do grupo de pesquisa dos autores contribuíram com três imagens cada, capturadas de diferentes ângulos faciais (frontal e ambos os perfis laterais). Essa seleção resultou em um total de 5.804 indivíduos e 9.211 imagens.

Descrição do experimento. O experimento de avaliação do reconhecimento de face envolveu um dispositivo simulado representando um *smartphone* e um dispositivo de borda, responsável por rodar a *pipeline* de reconhecimento de face. A avaliação foi condu-

zida por meio do processamento, no dispositivo de borda, de um fluxo de vídeo enviado pelo dispositivo simulado. O dispositivo simulado usa o GStreamer⁴ para codificar um vídeo gravado para H264 e enviá-lo como um fluxo de dados para o dispositivo de borda via protocolo RTSP⁵ em 30 quadros por segundo (fps). O vídeo usado para a avaliação de desempenho tem duração de 3min10s e mostra quatro indivíduos do grupo de voluntários movendo-se em um ambiente externo, com variações de postura, movimentações faciais e mudança de iluminação para testar robustez dos modelos de redes neurais utilizados. No dispositivo de borda, a aplicação de reconhecimento de face captura e processa esse fluxo, e, durante a fase de reconhecimento, usa o banco de dados vetorial Qdrant⁶ tanto para hospedagem das faces quanto para busca eficiente do indivíduo no banco. O dispositivo de borda foi um NVIDIA Jetson Nano 4GB com uma CPU Quad-core ARM Cortex-A57 e uma GPU NVIDIA Maxwell com 128 núcleos CUDA, capaz de atingir 0,5 TFLOPs (FP16). O Qdrant foi executado localmente, um serviço intermediário recebendo os *embeddings* da face e realizando a consulta de busca.

Para avaliação do impacto do *streaming* em diversas resoluções no consumo de banda e gasto energético, foram considerados os valores de resolução 360p, 480p, 720p e 1080p⁷, com taxas de transmissão de dados (*bitrates*)⁸ de 1700 kbps, 2000 kbps e 2500 kbps, respectivamente. Foi utilizado ainda um *smartphone* Samsung Galaxy S23 com processador Qualcomm Snapdragon 8 Gen 2 e câmera principal de 50 MP gravando vídeos a 30 fps e bateria com capacidade de 3900 mAh. A transmissão do vídeo foi realizada para um servidor executando uma aplicação Kurento na versão 7.2.0 utilizando conectividade 5G comercial móvel. Para maior controle, o *streaming* capturou a imagem do mesmo vídeo utilizado na avaliação de reconhecimento sendo reproduzido continuamente por 30 minutos. As medidas foram extraídas utilizando a ferramenta Battery Historian⁹, que afere o consumo de bateria de cada aplicação separadamente.

Resultados e discussão. Avaliou-se o desempenho do algoritmo de reconhecimento usando métricas de precisão e *recall*. Essas métricas foram analisadas para diferentes valores de resolução do vídeo de entrada (1080p, 720p, 480p e 360p) e de *threshold* de distância, ou seja, a distância euclidiana máxima permitida entre o *embedding* a ser reconhecido e o indivíduo da base de dados para que um reconhecimento seja considerado correto. A precisão é calculada usando a Equação 1 e o *recall* usando a Equação 2:

$$\text{Precisão} = \frac{VP}{VP + FP} \quad (1) \quad \text{Recall} = \frac{VP}{VP + FN} \quad (2)$$

sendo:

- *VP* (Verdadeiros Positivos) o número de reconhecimentos corretos, nos quais a distância média está abaixo do valor de *threshold* e a pessoa reconhecida corresponde à identidade real;
- *FP* (Falsos Positivos) o número de reconhecimentos incorretos, nos quais a distância média está abaixo do valor de *threshold*, mas a pessoa reconhecida não corresponde à identidade real, e;

⁴<https://gstreamer.freedesktop.org/>

⁵<https://www.ietf.org/rfc/rfc2326.txt>

⁶<https://qdrant.tech/>

⁷O WebRTC Android permite fixar só a resolução da transmissão. Disponível em <https://bit.ly/4cnrM3h>

⁸Recomendações do recomendados pelo Kurento conforme documentação em <https://bit.ly/42gnhD6>

⁹<https://developer.android.com/topic/performance/power/setup-battery-historian>

- *FN* (Falsos Negativos) o número de reconhecimentos perdidos que deveriam ter sido corretos, nos quais a distância média está acima do valor de *threshold*, mas a pessoa corresponde à identidade real.

O valor de *threshold* pode ser ajustado conforme a necessidade da aplicação, permitindo um controle sobre o equilíbrio entre precisão e sensibilidade. Para alcançar uma precisão maior, o valor de *threshold* pode ser reduzido, reduzindo a taxa de FPs, mas ao custo de uma menor sensibilidade, ou seja, uma menor capacidade de identificar corretamente todos os indivíduos positivos. Por outro lado, diminuir o valor de *threshold* aumenta a sensibilidade, mas pode resultar em uma maior taxa de FPs, impactando a precisão.

A Figura 6 mostra como a precisão e o *recall* variam em função da resolução do vídeo e do valor de *threshold* de distância euclidiana. Para valores de *threshold* mais baixos há uma alta precisão e baixo *recall*, já que a estratégia está sendo mais seletiva no que está sendo reconhecido. Além disso, a resolução mais baixa de 360p é notavelmente inferior em precisão, mas essa diferença não é observada significativamente para outras resoluções. Considerando a criticalidade de operações na segurança pública, é importante reconhecer indivíduos com a maior confiabilidade possível e não necessariamente ter uma grande quantidade de reconhecimentos. Desta forma, com um valor de *threshold* igual a 0,8 resulta em uma precisão acima de 93%, sendo indiferente usar imagens em 1080p, 720p ou 480p, além de um *recall* acima de 50% para 1080p e 720p. Sendo assim, uma variação na resolução da transmissão devido ao algoritmo de controle de congestionamento da banda usado pelo protocolo WebRTC resulta em desempenho semelhante desse reconhecimento quando resoluções estão acima de 480p.

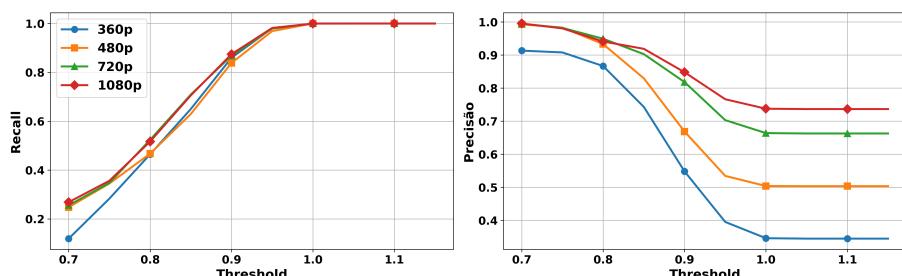


Figura 6. Precisão e *recall* para diferentes valores de *threshold* e resolução de imagem.

Os experimentos de consumo de banda e energia durante o *streaming* de vídeo para a resolução de 1080p registraram um tráfego médio de dados transmitidos de 579,23 MB e um consumo médio de bateria de 200,44 mAh, correspondendo aproximadamente a 5,41% da carga total. Para 720p, os resultados foram semelhantes, com 576,06 MB de dados trafegados na média e 201,06 mAh de uso médio de bateria, já que a taxa de transmissão recomendada para as resoluções 1080p e 720p é o mesmo. No entanto, na resolução de 480p, o tráfego médio de dados foi reduzido para 461,46 MB e a utilização média da bateria foi de 197,48 mAh (5,33%). Na resolução 360p, foram mensurados um uso médio de 197,39 mAh (5,33%) de bateria e 395,27 MB de dados transferidos. Portanto, utilizar a resolução de 480p diminui cerca de 1,5% do gasto de bateria e 20% dos dados transmitidos em comparação com a resolução de 1080p, mantendo a precisão acima de 93%.

5. Trabalhos Relacionados

Atualmente, soluções de processamento de vídeos por IA vêm sendo testadas pela Polícia Militar (PM) em vários locais do Brasil [Vieira et al. 2024]. O *PMSC Mobile* [PMSC 2017] é uma solução desenvolvida pela PM do Estado de Santa Catarina para otimizar a gestão de atendimentos policiais por meio de celulares. Esse aplicativo permite o registro de ocorrências, consulta de informações sobre veículos e indivíduos, despacho automatizado de viaturas, integração com sistemas de segurança pública e visualização de ocorrências por GPS. No entanto, não processa vídeos por IA, aspecto central ao CAD-Viatura. O *HxGN OnCall Dispatch*¹⁰, adotado pelas PMs dos Estados da Bahia e do Amazonas, realiza o despacho de ocorrências, auxiliado por IA, para os policiais em operação por meio de dispositivos móveis. A solução oferece processamento de vídeo com IA utilizando câmeras instaladas nas cidades, porém, além de ser uma solução proprietária, ela não oferece suporte para processamento na borda de *streaming* de vídeo em tempo real durante as operações, como é feito no CAD-Viatura.

Nakashima [Nakashima 2024] aborda o uso da IA pela PM do Estado do Paraná, destacando seu potencial e desafios para a predição e prevenção de crimes. São apresentadas tecnologias como o aplicativo 190, sistemas de análise comportamental e o Sistema de Atendimento e Despacho de Emergências usado para despacho automático de viaturas com base na geolocalização. O trabalho enfatiza o uso de câmeras com reconhecimento de face instaladas nas viaturas policiais, auxiliando na identificação de indivíduos com mandado de prisão ou outras irregularidades. Além disso, menciona iniciativas como o Projeto Falcão, que usa helicópteros e drones para reforçar a segurança pública, e o Projeto Olho Vivo, que promove a instalação de câmeras de monitoramento em diferentes regiões do Paraná. Diferentemente do CAD-Viatura, essas soluções exigem infraestrutura de nuvem, servidores de alto desempenho e conectividade contínua com a internet. Ademais, o Projeto Falcão envolve o aluguel de equipamentos com IA embarcada, sugerindo o uso de *software* de reconhecimento proprietário, ao contrário do CAD-Viatura, que permite a implantação de algoritmos personalizados.

6. Conclusão

O CAD-Viatura é uma solução integrada para melhorar o atendimento de ocorrências policiais, combinando computação móvel, IA, computação na borda e processamento de vídeo em tempo real. O sistema é um aliado estratégico na otimização da resposta policial, proporcionando uma melhora significativa na CS dos policiais, mesmo em contextos com baixa conectividade. A proposta também se destaca por sua arquitetura modular que possibilita a integração com sistemas CAD existentes em outros centros operacionais.

Os experimentos conduzidos confirmam a robustez do reconhecimento de face em resoluções acima de 480p, mantendo precisão acima de 93%, fundamental para a operação relacionada à segurança pública. No tocante ao consumo de energia durante o *streaming* de vídeo, a aplicação apresentou pouca variação entre as resoluções, enquanto o consumo de banda apresentou economia de cerca de 20% em comparação entre resoluções de 480p e 1080p. Futuras pesquisas deverão ampliar o desenvolvimento do CAD-Viatura e avaliar a escalabilidade e eficiência da solução.

¹⁰<https://hxgnpublicsafety.com/pt-br/seguranca-publica-integrada/>

Agradecimentos

Esta pesquisa é apoiada pela FINEP (Ref. nº 2827/22) e pelo INCT Redes de Comunicação e Internet das Coisas Inteligentes – ICoNIoT, financiado pelo CNPq (processo nº 405940/2022-0) e pela CAPES (processo nº 88887.954253/2024-00).

Referências

- Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors*.
- Fórum Brasileiro de Segurança Pública (2024). 18º Anuário Brasileiro de Segurança Pública: 2024.
- Huang, G. B. et al. (2007). Labeled faces in the wild: A database for studying face recognition in unconstrained environments. Technical Report 07-49, University of Massachusetts, Amherst.
- Liu, Z. et al. (2015). Deep learning face attributes in the wild. *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, pages 3730–3738.
- Loss, S. et al. (2025). Framework for live situational awareness in stream-based 5g applications. In *Proceedings of the 1st IEEE Latin American Conference on Internet of Things (IEEE LACIoT)*. IEEE.
- Mahmoud, H. and Abozariba, R. (2025). A systematic review on webrtc for potential applications and challenges beyond audio video streaming. *Multimedia Tools and Applications*.
- Nakashima, M. (2024). Desvendando o potencial e os desafios da inteligência artificial na polícia militar do paraná: Estratégias para predição e prevenção de crimes. *REASE*, 10(1):1321–1336.
- PMSC (2017). PMSC mobile – tecnologia móvel para gestão de atendimentos policiais. *Gestão Eletrônica de Trânsito – Talonário Eletrônico*.
- Redmon, J. et al. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.
- Schroff, F. et al. (2015). Facenet: A unified embedding for face recognition and clustering. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.
- Vieira, F. et al. (2024). Aplicabilidade da inteligência artificial nas polícias militares brasileiras: Vantagens, desvantagens e/ou limitações. In *Anais do Congresso Internacional de Conhecimento e Inovação (ciKi)*.
- Zhang, J. et al. (2016). Profiling energy consumption of dash video streaming over 4g lte networks. In *Proceedings of the 8th International Workshop on Mobile Video, MoVid '16*, NY, USA. ACM.