

Reidentificação Veicular em Ambientes Reais com Rotulação Assistida e Dados de Baixa Qualidade

Artur Henrique do Nascimento Souza^{1,2}, Augusto Abling^{1,2}, Raquel F. Vassallo¹

¹Universidade Federal do Espírito Santo (UFES)
Departamento de Engenharia Elétrica

²Atman Systems

artur.souza@edu.ufes.br, raquel.vassallo@ufes.br
augusto.abling@atmansystems.com

Abstract. This work presents a vehicle re-identification system for ITS, focusing on Brazilian urban contexts. An open-source tool for labeling assisted by neural networks is proposed, facilitating the creation of datasets tailored to local conditions. Additionally, a spatio-temporal re-identification system is implemented, using physical metrics such as distance and speed between cameras to enhance efficiency and accuracy. Experiments show that combining local datasets with international benchmarks, such as VeRi-776, improves model generalization (mAP of 0.867). The results highlight the importance of regional adaptation for robust and effective systems in real-world scenarios.

Resumo. Este trabalho traz um sistema de reidentificação veicular com foco em contextos urbanos brasileiros. Propõe-se uma ferramenta open-source para rotulação assistida por redes neurais, que facilita a criação de datasets adaptados à realidade local. Ademais, um sistema de reidentificação espaço-temporal é implementado, usando métricas físicas como distância e velocidade entre câmeras para melhorar a eficiência e precisão. Os resultados mostram que a combinação de datasets locais com benchmarks internacionais, como VeRi-776, melhora a generalização dos modelos (mAP 0,867), destacando a relevância da adaptação regional para sistemas robustos e eficazes em cenários reais.

1. Introdução

A computação ubíqua, descrita por Weiser [Weiser 1999], refere-se à integração invisível de computadores ao ambiente cotidiano. Nesse contexto, os Sistemas de Transporte Inteligente (ITS, do inglês *Intelligent Transportation Systems*) têm desempenhado um papel fundamental na modernização da mobilidade urbana. O aumento da frota veicular gera congestionamentos e poluição, afetando a qualidade de vida da população. Esses problemas reforçam a necessidade de soluções eficientes, como os ITS, para otimização do tráfego urbano.

Os ITS dependem de diversas variáveis para representar a dinâmica do tráfego, a qual é essencial para um gerenciamento eficiente dos recursos disponíveis. Uma das aplicações é a otimização do tempo semafórico, de acordo com a demanda horária, para maximizar o fluxo das vias, sem a necessidade de grandes investimentos em infraestrutura como construção de novas rodovias ou duplicações de vias.

Nesse contexto, as câmeras de videomonitoramento destacam-se como sensores valiosos para análise do tráfego urbano. Mais do que simplesmente classificar a intensidade do fluxo, câmeras espalhadas em diferentes vias podem ser usadas para identificar por onde um mesmo veículo trafegou ao longo do tempo. Essas informações podem ser aplicadas na estimativa do tempo de deslocamento dentro de uma mesma malha viária e, assim, permitir que diferentes planos semafóricos possam ser avaliados.

Para isso, pode-se empregar a técnica de reidentificação veicular. Esse processo consiste em estabelecer correspondências entre imagens de um mesmo veículo dentro de uma galeria de registros visuais, capturados por diferentes câmeras nas vias públicas onde possivelmente o veículo transitou. A partir dessa análise, é possível comparar, de forma prática e baseada em dados reais, como cada configuração semafórica impacta a fluidez do tráfego. Isso permite embasar a escolha do plano mais adequado, contribuindo para intervenções mais eficazes na mobilidade urbana.

Especificamente, a reidentificação veicular tem atraído a atenção da comunidade científica e da indústria, impulsionada pelos rápidos avanços nas técnicas de aprendizado de máquina. Contudo, ainda existem desafios a serem superados, entre eles as discrepâncias intra-classe, semelhanças inter-classes, variabilidade de pontos de vista, disponibilidade limitada de dados anotados, processamento em tempo real e questões de privacidade [Amiri et al. 2024].

Os trabalhos em *deep learning* têm impulsionado o desenvolvimento de soluções para reidentificação veicular, sendo esta a principal abordagem atualmente. Modelos baseados em redes neurais profundas são treinados e testados em diversos conjuntos de dados, como o VeRi-776 [Liu et al. 2016]. Entretanto, o desempenho desses modelos está diretamente relacionado à qualidade e representatividade dos dados utilizados no treinamento. Um desafio relevante nesse cenário é a capacidade de generalização dos modelos, especialmente em ambientes com condições de captura distintas. Por exemplo, modelos treinados em imagens de alta resolução podem apresentar desempenho insatisfatório quando aplicados a câmeras de baixa qualidade, uma situação comum em países como o Brasil, onde a infraestrutura de videomonitoramento apresenta grande heterogeneidade.

Desta forma, este trabalho tem como objetivo desenvolver um sistema de reidentificação veicular adaptado a condições reais de videomonitoramento urbano, considerando cenários com câmeras heterogêneas e qualidade de imagem variável, como é comum em cidades brasileiras. Como contribuições principais, propõe-se uma abordagem espaço-temporal que combina similaridade visual entre imagens e restrições físicas baseadas em distância entre câmeras e limites de velocidade, com o intuito de reduzir o número de falsos positivos e o custo computacional do processo de reidentificação. Além disso, apresenta-se uma ferramenta de rotulação assistida, que integra algoritmos de detecção, rastreamento e sugestão automática de correspondência entre veículos, otimizando a criação de conjuntos de dados anotados. Também se destaca a construção de um conjunto de dados próprio, composto por imagens reais de ambientes urbanos brasileiros, o que permitiu a avaliação da generalização dos modelos treinados em contextos não controlados e com dispositivos de baixa qualidade, com implicações práticas para estudos de mobilidade urbana.

2. Trabalhos Relacionados

Desde 2014, a Reidentificação Veicular evoluiu de abordagens manuais para métodos baseados em Redes Neurais Convolucionais (CNNs) [Wang et al. 2024], com destaque para arquiteturas *multi-branch* aplicadas ao dataset VeRi-776 [Liu et al. 2016].

Um exemplo é o modelo MBR (*Multi-Branch Representation*), que representa o melhor resultado de mAP em código aberto para reprodução [Almeida et al. 2023]. O modelo combina uma arquitetura leve com ramos especializados: um ramo baseado em ResNet50-IBN para características globais, como cor, e outro em BotNet (transformers com auto-atenção) para detalhes locais discriminativos, como adesivos ou danos. A inovação está na estratégia *Loss-Branch-Split* (LBS), que treina cada ramo com funções de perda distintas (*Triplet Loss* para métricas de similaridade e *Cross-Entropy* para classificação), garantindo diversidade nas representações extraídas e reduzindo o risco de *overfitting*.

Entre os métodos consolidados, tem-se o VANet [Zhou and Shao 2018], que utiliza um aprendizado métrico sensível ao ponto de vista, por meio de duas redes convolucionais paralelas, para aprender métricas específicas para pontos de vista semelhantes e diferentes. Já o PCRNet (*Parsing-guided Cross-part Reasoning Network*) [Liu et al. 2020], combina CNNs com Redes Convolucionais em Grafos (GCNs) explorando correlações entre partes segmentadas semanticamente do veículo.

Entretanto, o MBR oferece desempenho competitivo no dataset VeRi-776, sem depender de anotações adicionais complexas como segmentação. Nesse contexto, a abordagem MBR destaca-se por sua eficiência computacional e escalabilidade prática, oferecendo uma solução robusta e aplicável a cenários reais, como monitoramento urbano e sistemas inteligentes de transporte.

Outro desafio é a criação de datasets anotados, devido ao alto custo e tempo exigidos pela rotulação manual. Abordagens assistidas por redes neurais, têm surgido para mitigar essa carga. Em [Liao et al. 2022] foi proposto um sistema semiautomático para anotação de trajetórias veiculares multi-câmera, utilizando-se redes neurais. Embora eficiente, a ferramenta não é *open-source*, limitando sua reutilização e personalização.

Por outro lado, ferramentas *open-source* como o MOT_Annot, usada na criação do dataset CityFlow [Tang et al. 2019], oferecem suporte à detecção e rastreamento automático, mas carecem de uma interface gráfica intuitiva e recomendações inteligentes, o que impacta a eficiência prática da sua utilização.

Apesar das iniciativas, ainda há lacunas no desenvolvimento de ferramentas *open-source* para reidentificação veicular. Soluções utilizadas para anotação de dados, como LabelImg [Tzutalin 2015] e CVAT [Sekachev and al. 2020], não são adequadas para atribuição de identidades de veículos em múltiplas câmeras, dificultando a criação de datasets para cenários reais. O presente trabalho visa preencher essa lacuna ao oferecer uma solução *open-source*, com uma interface gráfica intuitiva e assistida por redes neurais.

De volta ao problema de reidentificação, abordagens que integram informações espaço-temporais podem aprimorar a precisão de métodos baseados exclusivamente em aparência. Em [Kim et al. 2023], os autores usam uma janela de Parzen adaptativa para estimar a probabilidade temporal de transição entre câmeras, que, junto com a simila-

ridade visual entre imagens, gera uma pontuação de similaridade final. Tal abordagem superou métodos baseados em características visuais, com uma acurácia de Rank-1 de 99,64% no dataset VeVi-776. Contudo, esse ganho de precisão implica em um aumento no custo computacional, devido ao processamento adicional das informações temporais.

Entretanto, diferentemente dos seres humanos, os veículos seguem padrões de movimento mais rígidos, respeitando faixas e mantendo velocidades próximas aos limites estabelecidos para a via. Com base nessa característica, os autores de [Lv et al. 2019] propõem um mecanismo para restringir a busca no processo de reidentificação, removendo candidatos irrelevantes que não obedecem às regras de deslocamento temporal e espacial. Essa estratégia reduz o tempo de processamento e minimiza a ocorrência de falsos positivos ao restringir as correspondências plausíveis entre câmeras.

Desta forma, o presente trabalho propõe uma abordagem que incorpora métricas de trânsito para modelar a movimentação dos veículos de forma expansível para um número variável de câmeras. Para isso, são aplicados critérios de tempo mínimo e máximo de transferência entre câmeras, com o objetivo de reduzir o espaço de busca e as associações inviáveis. Isso visa equilibrar a precisão do sistema com a eficiência computacional, ao mesmo tempo em que se adapta a cenários urbanos complexos.

3. Metodologia

Esta seção apresenta o método e a estratégia adotados para a realização da reidentificação de veículos em ambientes reais, considerando as particularidades do videomonitoramento no contexto brasileiro. A abordagem proposta integra uma ferramenta de rotulação assistida baseada no modelo MBR4G e um sistema de reidentificação espaço-temporal, visando otimizar tanto a criação de datasets adaptados ao contexto brasileiro quanto a eficiência e precisão do processo de reidentificação em ambientes urbanos reais.

3.1. Sistema de Rotulação Assistida

A ferramenta de rotulação assistida visa facilitar e semi-automatizar a criação de um dataset para reidentificação de veículos, reduzindo o tempo e o esforço manual necessários. O sistema desenvolvido (**Label BR ReID**) utiliza o modelo MBR4G, treinado no dataset VeVi-776, para auxiliar na correspondência automática entre veículos em diferentes câmeras. Além disso, incorpora o algoritmo de rastreamento SORT [Bewley et al. 2016] e diversas ferramentas em Python para a preparação e organização dos dados.

A validação dos pares de veículos é realizada por meio de uma interface gráfica, que permite a confirmação manual das correspondências sugeridas pelo modelo. Após a etapa de validação, o dataset gerado é adequado ao formato do VeVi-776, possibilitando a sua utilização em modelos de reidentificação baseados em aprendizado profundo.

A Figura 1 ilustra o fluxo geral do sistema, em que cada bloco representa uma etapa do processo. Todas as ferramentas desenvolvidas estão disponíveis em um repositório público no GitHub¹. A seguir, cada uma dessas etapas será descrita e detalhada.

Coleta de Vídeos: O sistema coleta vídeos de câmeras via protocolo RTSP e os armazena localmente para processamento posterior. Após a coleta, realiza-se um pré-

¹Github Ferramenta - <https://github.com/Arturhen/br-label-reid>

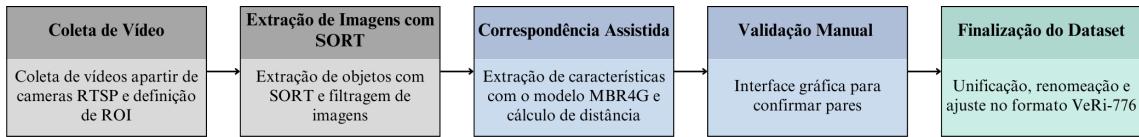


Figura 1. Fluxo do Sistema de Rotulação Assistida

processamento para definir a região de interesse (ROI), excluindo áreas irrelevantes e possibilitando a seleção de vias específicas em câmeras que monitoram múltiplas faixas.

Extração de Imagens com SORT: O sistema utiliza o algoritmo SORT para rastrear veículos, mantendo apenas três imagens representativas (primeira, intermediária e última) e removendo aquelas com menos de 100 pixels de altura. A estrutura resultante organiza os veículos por câmera da seguinte forma:

```

Result/
+-- Cam1/
|   +-- Car1/
|   +-- Car2/
+-- Cam2/
|   +-- Car1/
|   +-- Car2/

```

Correspondência Assistida: Após a extração e organização das imagens, é realizada uma correspondência assistida empregando-se uma rede neural. O modelo MBR4G pré-treinado é utilizado para realizar a correspondência inicial entre veículos em diferentes câmeras. O modelo extrai as características (*features*) de cada veículo e calcula a similaridade entre eles utilizando a distância euclidiana. O resultado é um arquivo CSV, contendo as distâncias entre todos os pares de veículos possíveis, ordenado do par mais semelhante ao menos semelhante.

Validação Manual: Para garantir a precisão, os pares sugeridos pela correspondência automática são revisados manualmente por meio de uma interface gráfica, mostrada na Figura 2. O sistema apresenta ao usuário os pares de veículos mais prováveis, permitindo a confirmação ou rejeição de cada correspondência. O progresso da validação é salvo em um arquivo CSV, permitindo que o processo seja interrompido e retomado posteriormente sem perda de informações. Ao final, gera-se um arquivo contendo as correspondências confirmadas entre veículos, detectados e rastreados em câmeras distintas.

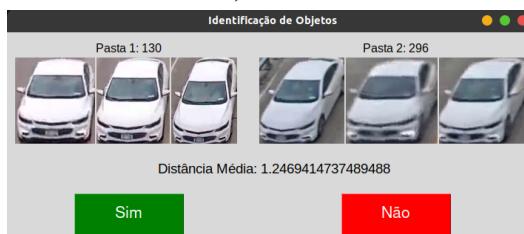


Figura 2. Interface para Validação Manual - BR LABEL REID

Finalização do Dataset: A partir das correspondências validadas, as imagens dos veículos são unificadas em pastas conjuntas, onde cada pasta representa um único veículo identificado em múltiplas câmeras. As imagens são renomeadas para associar cada veículo a um identificador único e registrar a câmera de origem. Em seguida, o dataset é dividido em conjuntos de treinamento e teste. O conjunto de teste ainda é dividido entre *query* e *gallery* garantindo que todas as imagens de uma mesma câmera sejam alojadas em apenas um dos conjuntos. A estrutura segue o mesmo formato do VeVi-776, o que facilita o uso do dataset em outros modelos de rede já treinados nesse padrão.

3.2. Dataset de Videomonitoramento (DatasetTest)

O DatasetTest foi criado a partir de duas câmeras heterogêneas em área urbana brasileira, apresentando desafios como efeito de ghosting e artefatos (Figura 3). Para facilitar a integração com ferramentas já consolidadas, o DatasetTest foi estruturado de forma semelhante ao VeRi-776, um dos principais *benchmarks* para reidentificação de veículos. A Tabela 1 apresenta uma comparação entre os dois conjuntos de dados, destacando suas principais características. Por questões de privacidade e restrições de liberação, o DatasetTest ainda não está disponível publicamente.

Característica	DatasetTest	VeRi-776
Número de veículos	553	776
Número de câmeras	2	20
Total de imagens	3.318	49.357
Finalidade	Validação em cenário local	Benchmark geral de ReID

Tabela 1. Comparação entre o DatasetTest e o VeRi-776.

A Figura 3 apresenta exemplos de imagens de ambos os conjuntos de dados, evidenciando as diferenças de qualidade e os desafios específicos do DatasetTest, como os efeitos de ghosting e a presença de artefatos visuais.



Figura 3. Exemplos de imagens de cada DataSet

3.3. Treinamento e Avaliação dos Modelos

Para avaliar a importância de se complementar o conjunto de dados e aumentar a robustez do modelo quando a intenção é realizar reidentificação veicular em ambientes mais desafiadores, foram realizados três experimentos de treinamento:

- Treinamento apenas com o DatasetTest
- Treinamento apenas com o VeRi-776
- Treinamento combinado (VeRi-776 + DatasetTest)

Para a análise quantitativa, foram utilizadas as métricas mAP (mean Average Precision), CMC@1 e CMC@5, que são amplamente empregadas para avaliar o desempenho em reidentificação de veículos. Uma análise qualitativa foi feita por inspeção no formato Rank-5, utilizando as imagens de *query* e *gallery* do DatasetTest. Essa abordagem permitiu comparar o comportamento dos modelos em diferentes condições e avaliar visualmente as diferenças entre os resultados obtidos. Os detalhes dessas avaliações serão apresentados e discutidos na Seção 4 .

3.4. Sistema de Reidentificação Espaço-Temporal

O modelo treinado foi integrado a um sistema de reidentificação de veículos para extrair métricas de tráfego em um ambiente urbano. O objetivo é garantir que os veículos

reidentificados apresentem alta confiabilidade, focando nas métricas das vias, e não dos veículos individualmente. O sistema usa informações espaço-temporais para correlacionar veículos detectados em diferentes câmeras, considerando a viabilidade do deslocamento entre elas dentro de um intervalo de tempo plausível.

Para cada veículo, o sistema verifica se ele poderia ter se deslocado de uma câmera a outra, considerando o tempo necessário para percorrer a distância entre elas. O tempo mínimo e máximo de deslocamento são calculados pelas seguintes fórmulas:

$$T_{\min} = \frac{D}{V_{\text{MaxVia}} \times 1.5} \quad \text{e} \quad T_{\max} = \frac{D}{V_{\text{congest}} \times 0.8}$$

onde: D é a distância entre as câmeras; V_{MaxVia} é a velocidade máxima da via; e V_{congest} é a velocidade de congestionamento.

O sistema elimina candidatos improváveis, considerando apenas veículos que poderiam ter transitado entre as câmeras dentro do intervalo de tempo calculado, otimizando o processamento distribuído. Para reduzir falsos positivos, duas estratégias são usadas:

1. **Distância mínima entre features:** Apenas candidatos com alta similaridade.
2. **Desambiguação por proximidade:** Quando dois candidatos de reidentificação estão suficientemente próximos ao *query*, ambos são descartados.

Essa abordagem garante que apenas veículos com alta probabilidade de reidentificação sejam considerados na análise entre câmeras. Veículos não reidentificados são ignorados, pois suas métricas podem ser calculadas isoladamente para cada câmera.

4. Experimentos Resultados e Discussões

4.1. Configuração Experimental

Os experimentos utilizaram o modelo MBR4G, treinado em um computador com Intel Core i7 e NVIDIA GeForce RTX 4070. A ferramenta de rotulação assistida foi implementada em Python, garantindo compatibilidade com o formato VeRi-776. Para o treinamento do modelo, foram utilizados os mesmos hiperparâmetros descritos no artigo original do MBR4G [Almeida et al. 2023].

4.2. Resultados Quantitativos

Para avaliar o desempenho dos modelos em diferentes configurações de treinamento, foram calculadas as métricas mAP (*Mean Average Precision*), CMC@1 e CMC@5. A Tabela 2 apresenta os resultados obtidos com os três modelos treinados:

- Modelo treinado apenas com o VeRi-776
- Modelo treinado apenas com o DatasetTest
- Modelo treinado com a combinação do VeRi-776 e DatasetTest

Os resultados revelam que o modelo treinado apenas com o DatasetTest apresenta um desempenho elevado ao avaliar o próprio conjunto de dados (**mAP = 0.8793**), mas tem dificuldades em generalizar para o VeRi-776 (**mAP = 0.1699**). Isso indica que o DatasetTest, embora representativo do cenário local, possui uma diversidade limitada em relação ao conjunto mais amplo.

Treinamento	Avaliação	mAP	CMC@1 / CMC@5
DatasetTest	DatasetTest	0.8793	0.8609 / 0.9544
	VeRi-776	0.1699	0.5196 / 0.6412
VeRi-776	DatasetTest	0.4122	0.3741 / 0.6570
	VeRi-776	0.8178	0.9684 / 0.9868
DatasetTest + VeRi	DatasetTest	0.8669	0.8854 / 0.9869
	VeRi-776	0.8095	0.9684 / 0.9874

Tabela 2. Desempenho do modelo em diferentes cenários de treinamento

Por outro lado, o modelo treinado apenas com o VeRi-776 mostra uma performance robusta no próprio conjunto (**mAP = 0.8178**), mas tem desempenho reduzido ao avaliar o DatasetTest (**mAP = 0.4122**). Esse resultado reforça a necessidade de adaptação do modelo ao ambiente específico de videomonitoramento local.

A combinação dos dois datasets no treinamento proporciona um equilíbrio entre generalização e adaptação ao cenário local. O modelo híbrido alcança (**mAP = 0.8669**) no DatasetTest e (**mAP = 0.8095**) no VeRi-776, demonstrando maior robustez e melhor desempenho em ambos os contextos. Esses resultados destacam a importância de incorporar dados locais no treinamento para melhorar a precisão em ambientes específicos.

4.3. Análise Qualitativa

Foi ainda realizada uma análise qualitativa utilizando as imagens de consulta (*query*) do DatasetTest. Foram geradas visualizações do Rank-5, que são as cinco principais correspondências retornadas pelo modelo em relação à imagem de consulta. Essa análise permitiu identificar as diferenças de comportamento entre o modelo treinado exclusivamente com o VeRi-776 e o modelo treinado com a combinação do VeRi-776 + DatasetTest.

A Figura 4 exibe um exemplo do Rank gerado pelo modelo treinado apenas com o VeRi-776. Nesse caso, o modelo prioriza veículos provenientes da mesma câmera, mas apresenta dificuldades em identificar corretamente o veículo-alvo quando há variações causadas por artefatos como *ghosting* e fragmentos na imagem. Observa-se que apenas o segundo candidato é a correspondência correta, enquanto as demais posições apresentam veículos visualmente semelhantes, porém incorretos.

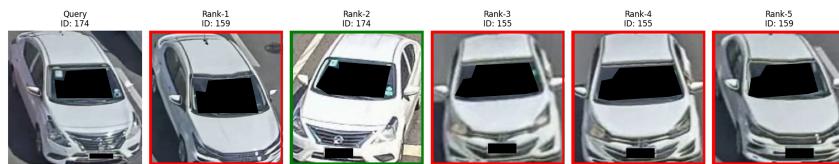


Figura 4. Rank-5 - Modelo treinado apenas com VeRi-776.

Em contraste, a Figura 5 apresenta o Rank-5 gerado pelo modelo treinado com a combinação do VeRi-776 + DatasetTest. Esse modelo demonstra uma capacidade aprimorada de identificar corretamente o veículo-alvo, mesmo em condições adversas. Nesse exemplo, as três primeiras correspondências são precisas, indicando que o modelo aprendeu a considerar características mais específicas e robustas, como adesivos no vidro, em vez de depender apenas de elementos suscetíveis a ruídos, como o contorno do veículo.

A análise qualitativa reforça os achados quantitativos: o modelo híbrido é mais eficaz em cenários reais com variações nas condições de captura. Além disso, essa abor-



Figura 5. Rank-5 - Modelo treinado com VeVi-776 + DatasetTest.

dagem reduz a incidência de falsos positivos, uma vez que o modelo passa a considerar características mais distintivas e menos afetadas por artefatos na imagem.

Esse comportamento é essencial para sistemas de reidentificação em ambientes urbanos brasileiros, onde a diversidade de câmeras e as condições variáveis de iluminação e captura podem impactar significativamente o desempenho.

5. Conclusões e Trabalhos Futuros

Este trabalho buscou trazer duas contribuições para a área de reidentificação veicular, com foco na adaptação a cenários reais e na superação dos desafios impostos por dados de baixa qualidade e infraestrutura heterogênea.

Uma nova ferramenta de rotulação assistida, **Label BR ReID**, foi desenvolvida e disponibilizada. De código aberto, com interface gráfica intuitiva e utilizando redes neurais para auxílio na correspondência entre veículos, esta ferramenta está disponível no GitHub². Sua aplicação na criação de datasets locais, como o DatasetTest, e o treinamento combinado destes com benchmarks internacionais como o VeVi-776, demonstraram o impacto positivo de dados representativos da realidade local. Especificamente, o modelo treinado com a combinação dos conjuntos de dados alcançou um mAP de 0.8669 no DatasetTest e 0.8095 no VeVi-776, evidenciando uma melhoria na capacidade de generalização e reforçando a importância de datasets alinhados ao ambiente de videomonitoramento para a robustez dos sistemas.

Além disso, um sistema de reidentificação que incorpora um filtro espaço-temporal foi implementado. Ao utilizar informações como a distância entre câmeras e os limites de velocidade, o sistema restringe as correspondências plausíveis de forma eficaz, reduzindo o número de comparações necessárias e o custo computacional. Com isso, a solução torna-se viável para execução em dispositivos embarcados, ao mesmo tempo que mantém alta precisão e diminui a ocorrência de falsos positivos.

Os resultados preliminares mostram o potencial da solução proposta. Espera-se que as duas abordagens, em conjunto, contribuam para o desenvolvimento de sistemas de reidentificação veicular mais adaptados e eficientes para os desafios encontrados em infraestruturas de videomonitoramento heterogêneas, comuns no Brasil, e ajudem na obtenção de métricas de tráfego mais confiáveis.

Como trabalhos futuros, considera-se a expansão do DatasetTest com um maior número de câmeras e locais distintos, aumentando sua diversidade e representatividade. A ferramenta Label BR ReID poderá ser utilizada na criação de outros datasets, tanto

²Github Ferramenta - <https://github.com/Arturhen/br-label-reid>

no Brasil quanto em outros países com características semelhantes. Sugere-se, ainda, a implementação de algoritmos de rastreamento mais avançados e a exploração de abordagens que usem informações geoespaciais, melhorando a precisão e a robustez do sistema.

6. Agradecimentos

Este trabalho contou com a empresa Atman Systems LTDA, que emprega os autores Artur Henrique do Nascimento Souza e Augusto Abling como forma de apoio ao projeto.

Referências

- Almeida, E., Silva, B., and Batista, J. (2023). Strength in diversity: Multi-branch representation learning for vehicle re-identification. In *2023 IEEE 26th International Conference on Intelligent Transportation Systems (ITSC)*, pages 4690–4696.
- Amiri, A., Kaya, A., and Keceli, A. S. (2024). A comprehensive survey on deep-learning-based vehicle re-identification: Models, data sets and challenges.
- Bewley, A., Ge, Z., Ott, L., Ramos, F., and Upcroft, B. (2016). Simple online and realtime tracking. In *2016 IEEE International Conference on Image Processing (ICIP)*.
- Kim, H.-G., Na, Y., Joe, H.-W., Moon, Y.-H., and Cho, Y.-J. (2023). Vehicle re-identification with spatio-temporal information. In *2023 14th International Conference on Information and Communication Technology Convergence (ICTC)*.
- Liao, H., Zheng, S., Shen, X., Li, M. J., and Wang, X. (2022). Semi-automatic data annotation system for multi-target multi-camera vehicle tracking. In *2022 IEEE 9th International Conference on Data Science and Advanced Analytics (DSAA)*.
- Liu, X., Liu, W., Ma, H., and Fu, H. (2016). Large-scale vehicle re-identification in urban surveillance videos.
- Liu, X., Liu, W., Zheng, J., Yan, C., and Mei, T. (2020). Beyond the parts: Learning multi-view cross-part correlation for vehicle re-identification. In *Proceedings of the 28th ACM International Conference on Multimedia*. Association for Computing Machinery.
- Lv, K., Du, H., Hou, Y., Deng, W., Sheng, H., Jiao, J., and Zheng, L. (2019). Vehicle re-identification with location and time stamps. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*.
- Sekachev, B. and al., E. (2020). opencv/cvat: v1.1.0.
- Tang, Z., Naphade, M., Liu, M.-Y., Yang, X., Birchfield, S., Wang, S., Kumar, R., Anastasiu, D., and Hwang, J.-N. (2019). Cityflow: A city-scale benchmark for multi-target multi-camera vehicle tracking and re-identification. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 8789–8798.
- Tzutalin (2015). Labelimg. Free Software: MIT License.
- Wang, Z., Wang, L., Shi, Z., Zhang, M., Geng, Q., and Jiang, N. (2024). A survey on person and vehicle re-identification. *IET Computer Vision*, 18(8):1235–1268.
- Weiser, M. (1999). The computer for the 21st century. *SIGMOBILE Mob. Comput. Commun. Rev.*, 3(3):3–11.
- Zhou, Y. and Shao, L. (2018). Viewpoint-aware attentive multi-view inference for vehicle re-identification. In *IEEE Conference on Computer Vision and Pattern Recognition*.