Anais do Congresso Brasileiro de Iniciação Científica
(https://periodicos.unisanta.br/COB)

ISSN 2317-1316

RASTREIO DE INDIVÍDUOS EM MÚLTIPLAS CÂMERAS UTILIZANDO SOFT BIOMETRICS

PERSON TRACKING ON MULTIPLE CAMERAS USING SOFT BIOMETRICS

Francesco Zangrandi Coppola, André Franco Raineri, Theodoro Priosti de Almeida, Bruno da Silva Rodrigues, Fabio Kawaoka Takase

Universidade Presbiteriana Mackenzie, Faculdade de Computação e Informática (FCI), curso de Ciência da Computação
E-mail para contato: Francescozcoppola@outlook.com

RESUMO – Sistemas de vigilância modernos geram um grande volume de vídeo, tornando desafiadora a tarefa de encontrar manualmente uma pessoa específica em meio a milhares de horas de gravações. Um cenário típico envolve a busca por uma pessoa de interesse, como por exemplo uma criança que se perdeu, com base em uma descrição. Atualmente, essa busca é realizada de forma majoritariamente manual, com operadores analisando múltiplas câmeras por horas. Além de ser um processo lento e sujeito a falhas humanas, a busca manual pode não abranger toda a vasta quantidade de dados disponível. Surge, portanto, a necessidade de automatizar o rastreio de indivíduos em diferentes câmeras momentos, utilizando características gerais para localizar ocorrências de uma pessoa de interesse em uma base de dados, dentro de uma área específica e em um determinado intervalo de tempo.

Palavras-chave: Soft biometrics. Reidentificação de pessoas. Vigilância por vídeo. Visão computacional. YOLOv5. HSV. MFA-ViT. LGPD.

ABSTRACT – Modern surveillance systems use a large volume of video, making it challenging to manually find a specific person among thousands of hours of recordings. A typical scenario is a search for a person of interest, such as lost child, based on a description. Currently, this search is mostly performed manually, with operators reviewing multiple cameras for hours. In addition to being a slow process and prone to human error, manual search may not cover the entire vast amount of available data. Therefore, there is a need to automate the re-identification of individuals across different cameras and times, using a generic description to locate occurrences of a person of interest within a database, in a specific area and within a given time interval.

Keywords: Soft biometrics. Person re-identification. Video surveillance. Computer vision. YOLOv5.HSV. MFA-ViT. LGPD**.**

ISSN 2317-1316

1 INTRODUÇÃO

1.1 Contexto e Motivação

Sistemas de vigilância estão presentes em ambientes urbanos, comerciais e institucionais em escala crescente, gerando diariamente volumes massivos de gravações que tornam inviável a análise manual por operadores humanos. Um cenário típico e recorrente é a busca por uma pessoa de interesse com base em descrição verbal, como por exemplo uma criança desaparecida ou um suspeito com características físicas específicas, tarefa realizada de forma majoritariamente manual, com operadores analisando múltiplas câmeras por horas ou dias, processo lento, sujeito a falhas e que frequentemente não abrange toda a base de dados disponíveis. O problema se agrava em ambientes com câmeras não sobrepostas, nosquais um indivíduo que transita entre diferentes pontos de um edifício ou área urbana aparece em câmeras distintas com ângulos, iluminações e resoluções variadas, tornando inviável qualquer abordagem de comparação direta por frame ou pixel. Surge, portanto, a necessidade de automatizar a reidentificação de indivíduos em diferentes câmeras e momentos, utilizando descrições gerais de atributos para localizar ocorrências de uma pessoa de interesse em uma base de dados, dentro de uma área específica e em um determinado intervalo de tempo.

A definição do problema abrange a reidentificação em câmeras não sobrepostas sem depender de identificadores biométricos rígidos, como reconhecimento facial; em vez disso, o foco recai sobre os soft biometrics, atributos físicos e comportamentais que não identificam de forma única um indivíduo como vestuário; gênero e cor do cabelo (DANTCHEVA; ELIA; ROSS, 2016). Embora ambíguos e não permanentes, esses atributos servem como evidências valiosas para reduzir a dimensão da busca e, ao evitar traços que possibilitam a identificação direta e inequívoca de um indivíduo, a abordagem minimiza os riscos associados ao tratamento de dados pessoais e sensíveis previstos na Lei Geral de Proteção de Dados Pessoais (LGPD) (BRASIL, 2018).

1.2 Objetivos específicos

1º - Implementar um módulo de detecção de pessoas em vídeo utilizando o detector YOLOv5s, capaz de identificar e recortar automaticamente frames contendo pessoas a partir de gravações de câmeras de vigilância.

ISSN 2317-1316

2º - Desenvolver um módulo de identificação de atributos soft biométricos, para classificação de gênero e detecção de cor de vestimenta, bem como o mecanismo de reidentificação por similaridade.

3º - Integrar os módulos de detecção e identificação em um sistema capaz de realizar busca por atributos visuais determinados como gênero aparente e cor de roupa com filtros de intervalo temporal e espaço.

4º - Adicionar validação manual das imagens resultantes da busca com fila dinâmica baseada em similaridade de resultados positivos.

5º - Desenvolver uma interface para simular a integração da aplicação a um sistema de vigilância com funcionalidades comums como perfil com níveis de acesso diferente (administrador e operador), o gerenciamento de câmeras, ingestão de vídeos, execução de buscas, validação e exportação dos achados.

6 º - Avaliar o desempenho do sistema em cenários simulados.

*Cabe ressaltar que a busca não é baseada em texto livre: o operador seleciona atributos visuais diretamente na interface gráfica, sem consulta por linguagem natural.

1.3 Referencial teórico

1.3.1 Soft Biometrics

O termo soft biometrics refere-se a atributos pessoais e físicos que, diferentemente de identificadores biométricos rígidos hard biometrics como impressões digitais, íris ou face, não identificam univocamente um indivíduo, mas fornecem informações complementares sobre sua aparência. Exemplos incluem gênero, faixa etária, altura, peso, cor dos olhos, cor do cabelo, tons de pele, estilo e cor das roupas, entre outros (DANTCHEVA; ELIA; ROSS, 2016).

Esses atributos podem ser extraídos a partir de dados biométricos primários como por exemplo, estimar gênero e faixa etária a partir de uma imagem facial, e têm aplicações importantes em vigilância e na indexação de bases de dados biométricas. Embora não sejam permanentes nem suficientemente distintos para autenticar alguém com alta confiabilidade isoladamente (JAIN; DASS; NANDAKUMAR, 2004) os soft biometrics se mostram úteis quando combinados a identificadores primários, para filtrar candidatos em uma busca ou aumentar a acurácia de reconhecimento quando utilizados com algoritmos tradicionais.

ISSN 2317-1316

Dantcheva, Elia e Ross (2016) destacam que esses atributos fornecem uma forma de descrição qualitativa de indivíduos (e.g., "mulher com cabelos castanhos") que ajuda a transpor a lacuna semântica entre descrições humanas e a interpretação automática de dados biométricos. Ou seja, soft biometrics permitem que sistemas computacionais lidem com descrições em linguagem natural ou atributos semânticos da mesma forma que um humano descreveria uma pessoa, tornando-os particularmente valiosos para vigilância.

1.3.2 Reidentificação de pessoas

A reidentificação de pessoas é um problema clássico de visão computacional que busca reconhecer a mesma pessoa através de diferentes câmeras ou em diferentes momentos, mesmo quando não há relação temporal direta. Em ambientes reais de CCTV(*Closed Circuit Television*), câmeras distintas capturam ângulos e iluminações diversos, e uma pessoa pode aparecer em resoluções baixas ou parcialmente oclusa por outros objetos. Essas variações de pose, iluminação, ponto de vista e qualidade das imagens tornam o person re-id um grande desafio técnico.

Métodos modernos baseados em aprendizado profundo treinam redes neurais ou modelos de classificação com milhares de fotos para aprender padrões robustos de aparência, capazes de agrupar fotos da mesma pessoa mesmo sob mudanças de câmera. Contudo, esses métodos geralmente exigem que as pessoas a serem reconhecidas apareçam previamente em um conjunto de treino ou sejam cadastradas previamente para atingir alta acurácia. Na prática, isso significa que se uma pessoa nunca foi vista ou cadastrada no sistema, os métodos tradicionais de reidentificação falham em encontrá-lo apenas a partir de características visuais

Em cenários reais de segurança, muitas vezes não se dispõe de imagens prévias do indivíduo de interesse, tem-se apenas a descrição de familiares ou testemunhas. Nesses casos, buscas baseadas em atributos semânticos emergem como solução viável. Halstead et al. (2018) enfatizam que técnicas de busca semântica por descrição textual eliminam a dependência de uma imagem de referência pré-cadastrada, permitindo procurar uma pessoa no acervo de vídeos a partir apenas de informações como vestuário e características físicas relatadas. Esse paradigma levou ao surgimento de competições e benchmarks específicos.

Em 2018, a IEEE AVSS lançou um desafio internacional de *Semantic Person Retrieval* com duas tarefas:

ISSN 2317-1316

(1) reidentificação de pessoa em uma galeria de imagens a partir de uma consulta semântica textual, e

(2) busca por uma pessoa específica dentro de um vídeo, também guiada por uma descrição textual.

Para suportar essas tarefas, os organizadores disponibilizaram uma base de dados anotada com atributos semânticos soft *biometrics* de indivíduos, bem como protocolos de avaliação padronizados. A criação desse tipo de benchmark mostra tanto a importância prática quanto a dificuldade do problema.

1.3.3 Detecção de Pessoas (Modelos Utilizados)

Detectores YOLO (*You Only Look Once*) (REDMON et al., 2016) estabelecem uma solução para detecção em tempo real por meio de uma única passagem pela rede neural. o YOLO trata a detecção como um problema único de regressão, processando a imagem inteira em uma única execução do modelo para detectar simultaneamente as *bounding boxes* e suas respectivas classes. A versão YOLOv5s (JOCHER, 2026) é utilizada neste trabalho para detectar indivíduos em filmagens de vigilância para posterior extração de atributos.

OpenCV DNN (Caffe) proposto por (LEVI; HASSNCER, 2015) é uma CNN (*Convolutional Neural Network*) que aplica filtros convolucionais hierárquicos sobre recortes faciais para retornar probabilidades para as classes Male e Female. A detecção prévia da face é realizada por uma rede SSD (*Single Shot Multibox Detector*) com backbone ResNet-10, operando via OpenCV DNN (*Deep neural network*). Neste trabalho, o modelo é empregado como filtro primário de busca, descartando candidatos cujo gênero aparente não corresponde ao informado pelo operador.

Extended U-Net é uma CNN em formato de "U" composta por um codificador que extrai características semânticas e um decodificador que reconstrói a segmentação pixel a pixel, conectados por skip connections que preservam informações espaciais de alta resolução. Neste trabalho, o modelo gera uma máscara binária que isola pixels de vestimenta sobre a qual a análise de cor é aplicada.

O MFA-ViT (*Multimodal Fusion Attention Vision Transformer*) (TIONG et al.,2024) é um modelo baseado na arquitetura ViT (Vision Transformer), que divide a imagem em patches processados por camadas de atenção multi-cabeça. O modelo integra face, região periocular e

ISSN 2317-1316

atributos soft biométricos em um espaço de embedding comum por meio do módulo MFA (*Multimodal Fusion Attention*) e do mecanismo MPT (*Multimodal Prompt Tuning*), gerando *embeddings* de alta dimensionalidade comparáveis por similaridade de cosseno. Neste trabalho, o modelo é empregado como mecanismo opcional de reordenação da fila de candidatos durante a validação humana.

1.3.4 Lei Geral de Proteção de Dados (LGPD)

A LGPD (BRASIL, 2018) regula o tratamento de dados pessoais no Brasil e estabelece categorias especiais de dados sensíveis que incluem dados biométricos que permitam a identificação direta de pessoas naturais.

O uso de reconhecimento facial como mecanismo primário de vigilância levanta questões legais e éticas sob essa legislação, especialmente em ambientes públicos e sem o consentimento explícito dos titulares.

Este trabalho foi desenvolvido minimizando riscos regulatórios, adotando salvaguardas alinhadas aos princípios de finalidade, necessidade e prevenção da LGPD:

Soft biometrics como mecanismo primário: gênero aparente e cor de roupa são atributos descritivos que não permitem identificação inequívoca de um indivíduo isoladamente, configurando-se como dados pessoais não sensíveis nessa modalidade de uso.

2 MATERIAIS E MÉTODOS

Para implementar as diferentes abordagens, diferentes tecnologias e bibliotecas foram selecionados. As principais bibliotecas e versões utilizadas são apresentadas na Tabela 1.

Tabela 1: Tecnologias

Biblioteca	**Versão**	**Finalidade**
Python	3.10	Implementação da solução
PyTorch	1.13.1+cu117	Inferência YOLOv5, U-Net e MFA-ViT
Torchvision	0.14.1+cu117	Transformações para MFA-ViT
Ultralytics	≥ 8.2.64	Suporte ao YOLOv5
OpenCV	≥ 4.x	Leitura de vídeo e processamento HSV
Pillow	≥ 9.0	Carregamento e exibição de imagens
NumPy	≥ 1.23	Operações matriciais e vetoriais
Tkinter	(stdlib)	Interface gráfica
Bcrypt	latest	*Hashing* de senhas
Sqlite3	(stdlib)	Persistência de dados de usuários
Concurrent.futures	(stdlib)	Uso de *threads*
Rtdl	0.0.13	Dependência do MFA-ViT
Seaborn	latest	Auxiliar utilizada no YOLOv5

ISSN 2317-1316

Neste trabalho, através de uma revisão bibliográfica diferentes abordagens para tratar o caso de reidentificação de indivíduos foram estudadas. Na Tabela 2 os aspectos são relacionados, indicando o racional para a decisão sobre a implementação e a fonte de informação utilizada em sua implementação:

Tabela 2: Aspectos implementados

Aspecto	Racional	Fonte
Detecção de pessoas via YOLOv5s	Detector em tempo real, já utilizado em vários setores.	[redmon_you_2016; jocher_yolov5_2026]
Classificação de gênero por OpenCV+ Caffe	O modelo de Levi e Hassner (2015) é referência já utilizada em condições reais.	[levi_age_2015; dantcheva_what_2016]
Segmentação de roupas por Extended U-Net	Gera uma máscara sobre os pixels da imagem pertencentes à roupa, eliminando ruído de fundo e cor de pele.	[vozarikova_clothing_2021]
Análise de cor no espaço HSV (em vez de RGB)	O RGB mistura cromaticidade e brilho. O HSV separa matiz, saturação e valor, permitindo análises robustas.	[sural_segmentation_2002]
Reidentificador facial MFA-ViT	Arquitetura já projetada para integrar atributos soft. Usado como refinamento da verificação de resultados.	[tiong_flexible_2024]
Busca por múltiplas cores simultâneas e filtros temporais e de câmera	Soft biometrics são inerentemente ambíguos; uma testemunha pode hesitar entre tonalidades. Filtros contextuais reduzem o conjunto de revisão.	[dantcheva_what_2016]
InterfaceTkinter	Biblioteca modular, de fácil entendimento e baixo custo técnico.	Conhecimentos prévios

2.1 Metodologia dos testes

Os testes foram conduzidos sobre o dataset público Market-1501, amplamente empregado em pesquisas de re-identificação de pessoas. O conjunto contém imagens de pedestres capturadas por múltiplas câmeras fixas em um campus universitário.

Foram realizadas duas baterias de testes independentes: uma para avaliar o módulo de busca por soft biometrics, e outra para avaliar o módulo de similariedade facial. Ambas as baterias contaram com dez sessões cada, utilizando identidades distintas e sorteadas aleatoriamente do dataset.

Scripts dedicados foram desenvolvidos para automatizar a preparação do ambiente de cada sessão (seleção de identidade, montagem de galeria com proporção 1 pessoa alvo com 9 impostores), bem como para a execução e coleta de métricas.

Várias métricas foram coletadas a cada busca sendo as mais relevantes:

1. Tempo de execução (segundos);

ISSN 2317-1316

2. Imagens da pessoa presentes no resultado retornado (verdadeiro positivo);
3. Imagens da pessoa ausentes do resultado(falso negativo);
4. Recall = VP / X, que expressa a fração das imagens da pessoa recuperada;
5. Redução, fração da galeria removida durante a busca.
6. Recall Top X, que expressa a fração das imagens VP estão no top X posições, sendo X a quantidade de imagens VP.

Vale ressaltar que os testes foram divididos em sub-buscas, de modo a avaliar separadamente a contribuição de cada atributo de soft biometrics e o efeito do número de referências utilizadas no módulo de similaridade.

3 RESULTADOS E DISCUSSÃO

O protótipo implementado encontra-se funcional e integra todos os módulos previs-tos no objetivo geral. Especificamente, foram concluídos:

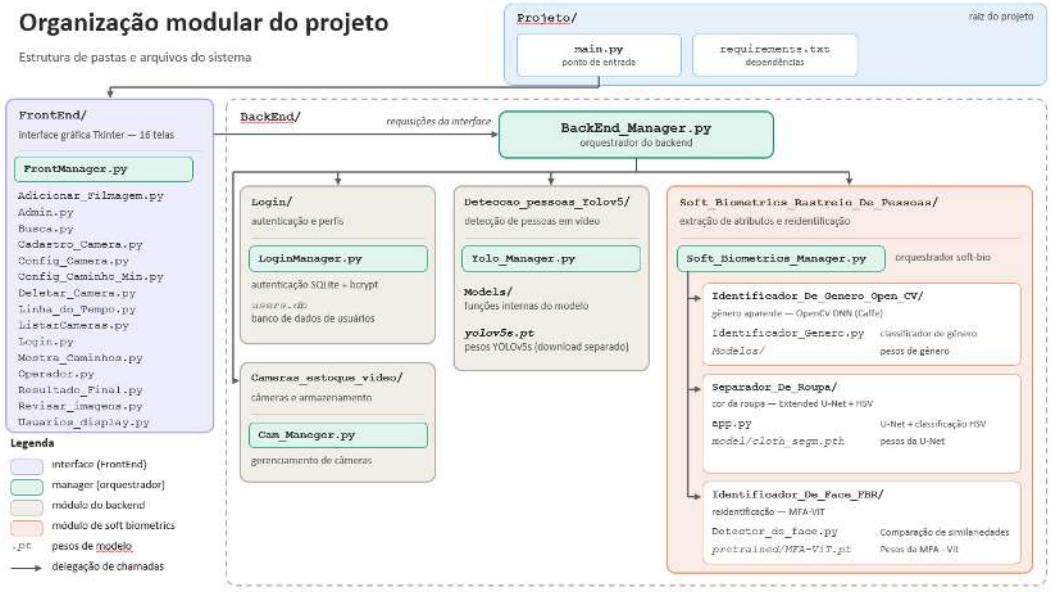
1. Pipeline de detecção: YOLOv5s opera sobre filmagens gerando recortes individuais.
2. Classificação de gênero: implementado de forma paralela, retornando (classe, confiança) ou “None” quando o modelo não consegue identificar.
3. Análise de cor: implementado de forma paralela, classificação em intervalos de cores 6 cores e ponderação por saturação com possibilidade de utilização do modelo Extended U-Net para extração de pixels de roupa.
4. Reordenador por similaridade (MFA-ViT): Durante a fase de revisão manual, se utiliza o modelo MFA-ViT juntamente com soluções heurísticas para a reordenação dos resultados baseado na similaridade das imagens positivamente revisadas.
5. Interface gráfica: 16 telas Tkinter funcionando, com fluxo completo Login →Busca →Revisão →Linha do Tempo →Resultado. Telas de gerenciamento (câmeras e usuários) operacionais.
6. Testes: Aplicados a uma base publica Market-1501 e teste em cenário simulado em um sistema real de segurança.

3.1 Visão Geral da Arquitetura Implementada

Esta seção detalha a abordagem de arquitetura e organização adotada para alcançar os objetivos delineados.

ISSN 2317-1316

Figura 1. Estrutura de pastas e arquivos



Esses componentes são orquestrados em duas camadas Backend e Frontend. Divididos em diversos níveis de abstração existindo um manager para integrar diferentes componentes de cada nível como representado na Figura.

As requisições do operador na camada de Frontend (interface Tkinter) são tra-duzidas em chamadas direcionadas ao BackEnd_Manager. Ao receber uma requisição, o BackEnd_Manager analisa o contexto da operação e delega a execução ao manager especializado correspondente: o Cam_Manager para operações de câmeras e armazenamento, o Yolo_Manager para input e detecção de pessoas, e o Soft_Biometrics_Manager para classificação de atributos e reidentificação.

Essa arquitetura em camadas, elimina dependências diretas entre o Frontend e os módulos de processamento, conferindo ao sistema alto grau de modularidade. Como consequência, novos classificadores, como um identificador de idade ou de postura, podem ser integrados ao Soft_Biometrics_Manager sem qualquer alteração, assim como módulos existentes podem ser refatorados ou substituídos de forma isolada, sem impacto nas demais camadas do sistema.

ISSN 2317-1316

Figura 2: principais telas da aplicação



3.2 Interface Gráfica

A interface gráfica foi construída de forma modular utilizando a biblioteca Tkinter de tal forma que é possível adicionar novas telas de forma simples. Telas que recebem dados dinâmicos são destruídas e recriadas a cada navegação. As principais telas construídas são apresentadas na Figura 2.

3.3 Testes de busca por Softbiometrics e Similaridade de face

Os resultados das dos testes de busca por soft biometrics foram consolidados na Tabela 3 como valores agregados para cada uma das quatro modalidades de busca.

ISSN 2317-1316

Tabela 3. Resultados agregados das 10 sessões de teste de busca por soft biometrics.

Métrica	Busca 1 (Cor+Gênero)	Busca 2 (Cor s/ mod.)	Busca 3 (Cor c/ mod.)	Busca 4 (Gênero)
Tempo médio	4.83	1.88	55.89	3.01
Imagens na entrada	2,093	2,093	2,093	2,093
Imagens retornadas	791	664	664	197
Redução (%)	62.21%	68.28%	68.28%	90.59%
Imgs pessoa disp. (X)	177	177	177	177
Verdadeiro Positivo (VP)	160	152	152	25
Falso Negativo (FN)	17	25	25	152
Recall = VP/X	90.40%	85.88%	85.88%	14.12%
VP / Imgs retornadas	20.23%	22.89%	22.89%	12.69%

O filtro por cor de roupa mostrou-se o principal da aplicação: sozinho (Busca 2), recuperou 85,88% das imagens disponíveis da pessoa buscada e reduziu em 68,28% o conjunto de imagens de entrada. O filtro de gênero, por sua vez, teve papel secundário, pois não atingiu resultados expressivos por si só (Busca 4), com recall de apenas 14,12%; porém, combinado ao filtro de cor (Busca 1), elevou o recall para 90,40%, ao custo de uma redução menor (62,21%, contra 68,28%). Já a Busca 3 (cor com o modelo U-Net) apresentou os mesmos valores de VP, FN e recall que a Busca 2, indicando que o modelo de segmentação de roupas não alterou os resultados nesta amostra; no entanto, seu tempo médio de execução foi de 55,89 s, cerca de 30 vezes superior ao da Busca 2 (1,88 s), sem ganho correspondente nas métricas avaliadas.

Tabela 4. Resultados agregados das 10 sessões do modulo de similaridade

ISSN 2317-1316

Métrica	Busca 1	Busca 2	Busca 3
Tempo médio (s)	3.35	2.09	2.09
Total imgs na busca	1,290	1,290	1,290
Imgs pessoa disp. (X)	129	129	129
VP (Top X)	35	31	40
FN (Top X)	94	98	89
Recall Top X	31.39%	30.33%	31.45%
Pessoa na 1ª metade	89	79	94
Recall 1ª Metade	74.47%	60.13%	74.52%
Média score pessoa	0.6596	0.661	0.6605
Média score total	0.6122	0.6246	0.6095

O Recall Top X médio de 31,06% indica que, dos primeiros X candidatos retornados no ranking (onde X é o número de imagens da pessoa na galeria), aproximadamente 1 em cada 3 corresponde ao alvo. Considerando que a galeria contém 10% de imagens da pessoa (proporção 1:9) e que um ranqueamento aleatório produziria um recall top X de 10% nessa proporção, o modelo MFA-ViT demonstra uma capacidade discriminativa real, triplicando o recall esperado por acaso. O Recall da Primeira Metade (69,71% em média) revela que o modelo é eficaz como mecanismo de ordenação: cerca de 70% das imagens da pessoa aparecem na metade superior do ranking, o que significa que o operador revisaria a maioria das imagens relevantes ao analisar os primeiros 50% dos resultados. Esse comportamento é

A análise individual das buscas mostra que a busca com uma referência (Busca 1) obteve resultados melhores do que a busca com duas referências (Busca 2), mas foi superada pela busca com três referências (Busca 3). Esse comportamento sugere que os resultados tendem a melhorar à medida que novas instâncias da pessoa buscada são encontradas e passam a ser usadas como referência.

4 CONCLUSÃO

Este trabalho apresentou o projeto, a implementação e a avaliação experimental de um sistema de reidentificação de pessoas em câmeras não sobrepostas, baseado em soft biometrics e refinado por similaridade via MFA-ViT. O sistema foi estruturado em uma arquitetura modular alinhada aos princípios de finalidade, necessidade e prevenção da LGPD. O objetivo

ISSN 2317-1316

geral foi atingido na forma de um protótipo funcional avaliado quantitativamente no dataset Market-1501. A busca combinada por cor e gênero obteve o maior recall (90,40%, ou 160 de 177 imagens-alvo) com redução de 62,21% do volume entregue ao operador, o que confirma a funcionalidade da aplicação. O modulo de reordenação por similaridade alcançou recall Top X médio de 31,06%, cerca do triplo do esperado ao acaso, e concentrou 69,71% das imagens-alvo na primeira metade do ranking, cumprindo sua função de ordenar a fila de revisão humana. O enriquecimento incremental do pool também se confirmou. Em conjunto, os resultados mostram que o sistema cumpre seu papel de ferramenta de apoio ao operador, com a validação humana mantida como decisão final. As limitações identificadas orientam os trabalhos futuros, facilitados pela arquitetura modular. Os trabalhos futuros incluem um classificador de gênero de corpo inteiro, alternativas à Extended U-Net mais adequadas ao contexto de busca, modelos de reidentificação de corpo inteiro complementares ao MFA-ViT e o processamento em tempo real de fluxos contínuos de CCTV.

5 DECLARAÇÃO DE USO DE IA

Não foram utilizadas ferramentas de IA no desenvolvimento desse artigo.

6 REFERÊNCIAS

Aggarwal S, Babu RV, Chakraborty A. Text-based person search via attribute-aided matching. In: IEEE Winter Conf Appl Comput Vis (WACV). Snowmass Village: IEEE; 2020. p. 2606-14.

Brasil. Lei nº 13.709, de 14 de agosto de 2018. Lei Geral de Proteção de Dados Pessoais (LGPD). Brasília; 2018.

Dantcheva A, Elia P, Ross A. What else does your biometric data reveal? A survey on soft biometrics. IEEE Trans Inf Forensics Secur. 2016 Mar;11(3):441-67.

Galiyawala H, Raval MS. Person retrieval in surveillance using textual query: a review. Multimed Tools Appl. 2021;80(18):27343-83.

Halstead M, Denman S, Fookes C, Tian Y, Nixon MS. Semantic person retrieval in surveillance using soft biometrics: AVSS 2018 Challenge II. In: 15th IEEE Int Conf Adv Video Signal Based Surveill (AVSS). Auckland: IEEE; 2018. p. 1-6.

Jain AK, Dass SC, Nandakumar K. Soft biometric traits for personal recognition systems. In: Zhang D, Jain AK, editores. Biometric Authentication. Berlin: Springer; 2004. p. 731-8. (LNCS; v. 3072).

ISSN 2317-1316

Jocher G. YOLOv5 [software]. Ultralytics; 2026. Disponível em: github.com/ultralytics/yolov5.

Levi G, Hassner T. Age and gender classification using convolutional neural networks. In: IEEE Conf Comput Vis Pattern Recognit Workshops (CVPRW). Boston: IEEE; 2015. p. 34-42.

Redmon J, Divvala S, Girshick R, Farhadi A. You only look once: unified, real-time object detection. In: IEEE Conf Comput Vis Pattern Recognit (CVPR). Las Vegas: IEEE; 2016. p. 779-88.

Smahesh29. Gender-and-Age-Detection [software]. GitHub; 2020. Disponível em: github.com/smahesh29/Gender-and-Age-Detection.

Sural S, Qian G, Pramanik S. Segmentation and histogram generation using the HSV color space for image retrieval. In: Proc Int Conf Image Process. Rochester: IEEE; 2002. v. 2, p. II-589-92.

Tiong LCO, Sigmund D, Chan CH, Teoh ABJ. Flexible biometrics recognition: bridging the multimodality gap through attention, alignment and prompt tuning. In: IEEE/CVF Conf Comput Vis Pattern Recognit (CVPR). Seattle: IEEE; 2024. p. 267-76.

Vozáriková G, Staňa R, Semanišin G. Clothing parsing using extended U-Net. In: Proc 16th Int Jt Conf Comput Vis Imaging Comput Graph Theory Appl (VISIGRAPP). SCITEPRESS; 2021. p. 15-24.