



INSTITUTO
UNIVERSITÁRIO
DE LISBOA

Reidentificação de pessoas em ambientes comerciais multicâmara com base no seu percurso

Diogo Amaro Mendes

Mestrado em Engenharia Informática,

Orientador:

PhD Luis Miguel Martins Nunes, Professor Associado,
ISCTE-IUL

Co-Orientadora:

PhD Patrícia Paula Lourenço Arriaga Ferreira, Professora Associada,
ISCTE-IUL

Setembro, 2023



TECNOLOGIAS
E ARQUITETURA

Departamento de Ciências e Tecnologias da Informação

Reidentificação de pessoas em ambientes comerciais multicâmara com base no seu percurso

Diogo Amaro Mendes

Mestrado em Engenharia Informática,

Orientador:

PhD Luis Miguel Martins Nunes, Professor Associado,
ISCTE-IUL

Co-Orientadora:

PhD Patrícia Paula Lourenço Arriaga Ferreira, Professora Associada,
ISCTE-IUL

Setembro, 2023

Agradecimentos

Ao longo do mestrado e da realização da dissertação, contei com o apoio de várias pessoas que tornaram todo este percurso mais simples e agradável e às quais gostaria de agradecer.

Em especial aos meus orientadores, professores Luís Nunes e Patrícia Arriaga, e aos restantes membros da equipa de investigação, professor Tomás Brandão, Simão Correia e Pedro Jorge, pelo acompanhamento e orientação, pela disponibilidade incondicional e pela boa disposição ao longo de todo o processo.

À minha família por toda a preocupação, pelo suporte, pela confiança que têm em mim e por estarem cá sempre para me apoiar.

À Marta por todo o apoio incondicional, pela positividade e encorajamento mesmo nos momentos mais difíceis, pelos conselhos e pelo exemplo de pessoa que representa para mim.

Aos meus amigos pela motivação e incentivo, pela amizade e preocupação.

Por fim, quero agradecer a mim próprio pelo compromisso, pela dedicação, pela disciplina e por ser o primeiro a acreditar que conseguia superar este desafio.

Financiamento

A realização desta dissertação foi parcialmente financiada por fundos nacionais através da FCT - Fundação para a Ciência e Tecnologia, I.P. no âmbito do projeto ECI4.0 LISBOA-01-0247-FEDER-047155 e através do Financiamento Plurianual das Unidades de I&D UIDP/04466/2020 (Information Sciences, Technologies and Architecture Research Center, ISTAR) e UID/PSI/03125/2019 (Centro de Investigação e Intervenção Social, Cis-Iscte).

Resumo

A presente dissertação propôs um algoritmo de reidentificação de pessoas baseado em trajetórias, inserido numa ferramenta para detetar e seguir clientes presentes numa grande superfície comercial, num ambiente constituído por múltiplas câmaras. Os dados da trajetória dos clientes foram obtidos a partir de imagens de videovigilância, captadas por várias câmaras, sendo os clientes localizados e seguidos ao longo dos *frames* que compunham os vídeos. Devido às características de um ambiente multicâmara e à ocorrência de oclusões causadas por objetos (e.g., prateleiras, balcões), são atribuídos vários identificadores a cada pessoa, quando na verdade deveria ser identificada com um identificador único. Desta forma, a ferramenta desenvolvida procurou resolver este problema, no âmbito de um cenário onde existem limitações na utilização de imagens de pessoas, devido a questões relacionadas com a privacidade de dados. O melhor método testado demonstrou ter uma fiabilidade aceitável (82%) ainda que, devido à falta de informação nas trajetórias extraídas, a qualidade da trajetória após o processo de reidentificação pudesse ser melhorada. Assim, este método deverá ser usado em duas situações: para informar e ser combinado com métodos baseados em imagem, reduzindo o peso do processamento de imagem quando as reidentificações podem ser obtidas com um elevado grau de confiança, e em situações onde as restrições impostas pela privacidade de dados impedem o processamento direto das imagens.

Keywords: reidentificação de pessoas; trajetória; multicâmara; deteção de objetos; tracking de objetos;

Abstract

This study presented a trajectory-based person re-identification algorithm, embedded in a tool to detect and track customers present in a large retail store, in a multi-camera environment. The customer trajectory data was obtained from video surveillance images captured by multiple cameras, and customers detected and tracked along the frames that compose the videos. Due to the characteristics of a multi-camera environment and the occurrence of occlusions, caused by objects (e.g., shelves, counters), different identifiers tend to be assigned to each person when in fact they must be identified with a unique identifier. Thus, our proposed tool aimed at solving this problem, in a scenario where there are constraints in using images of people, due to data privacy concerns. The best tested method proved to be reliable (82%) although, due to the lack of information in the extracted trajectories, the quality of the trajectory after the re-identification process could be improved. Thus, this method should be used in two situations: to inform and be combined with image-based methods, reducing the load of image processing when re-identifications can be obtained with a high confidence score, and in situations where the restrictions imposed by data privacy prevent direct image processing.

Keywords: person re-identification; trajectory; multi-camera; object detection; object tracking;

Índice

Agradecimentos	iii
Financiamento	iv
Resumo	v
Abstract	vii
Capítulo 1. Introdução	1
1.1. Motivação	1
1.2. Enquadramento e Objetivos	1
1.3. Metodologia	2
Capítulo 2. Revisão da Literatura	5
2.1. Técnicas de extração de trajetórias	5
2.2. Comportamentos extraídos através de vídeo	8
2.3. Detecção e <i>Tracking</i> de Objetos	9
2.4. Reidentificação de Pessoas	10
2.5. Síntese	11
Capítulo 3. Método de Reidentificação	14
3.1. Esquema Geral	14
3.2. Descrição do Conjunto de Dados	15
3.3. Preparação de Dados	16
3.4. Extração de Trajetórias	21
3.4.1. Detecção de Objetos	22
3.4.2. <i>Tracking</i> de Objetos	23
3.4.3. Estrutura de Dados e Homografia	24
3.5. Processo de Reidentificação	24
Capítulo 4. Experiências e Resultados	30
4.1. Experiência 1 – Reidentificação com primeiros e últimos pontos	30
4.2. Experiência 2 – Reidentificação com distância fixa	32
4.3. Experiência 3 – Reidentificação com distância dinâmica	36
4.4. Comparação dos Resultados	37
Capítulo 5. Conclusão	40

Referências Bibliográficas	42
----------------------------	----

Anexos	48
--------	----

Introdução

Este capítulo inicia com a motivação associada ao tema da dissertação, bem como o contexto em que a mesma se encontra inserida, seguida pela apresentação dos objetivos e questões de investigação às quais a presente dissertação pretende responder. Por fim, é descrita a metodologia utilizada fazendo um enquadramento com a estrutura da dissertação.

1.1. Motivação

As empresas de retalho têm vindo a passar por grandes processos de adaptação devido ao avanço da tecnologia no domínio da digitalização nos últimos anos, associada à crescente procura por parte dos consumidores [1]. Desta forma, as grandes superfícies comerciais procuram cada vez mais compreender o comportamento dos clientes de forma a conseguirem prever e satisfazer as suas necessidades e, consequentemente, aumentarem o número de vendas.

Com o aumento do número de dados possíveis de serem recolhidos neste tipo de estabelecimentos (por exemplo através de imagens provenientes de câmaras de videovigilância) [2], e graças aos avanços científicos realizados no campo da visão por computador, atualmente é possível identificar as pessoas presentes num conjunto de várias imagens consecutivas e extrair a trajetória efetuada por essas pessoas ao longo do seu percurso na loja.

Este tipo informação, mais concretamente a trajetória percorrida por cada cliente num estabelecimento comercial, aliado a técnicas de Inteligência Artificial (IA) permite diversos tipos de aplicações. A previsão da trajetória efetuada por um cliente e a análise do comportamento do mesmo através de características como o tempo que permaneceram na loja, as zonas abrangidas, ou a velocidade de caminhada [3] são algumas das finalidades para dados de trajetórias de clientes. Para o efeito é necessário que os dados recolhidos se aproximem o máximo possível da trajetória real do cliente, o que por vezes não é possível de obter devido à presença de objetos que causam oclusões no campo de visão da câmara, a variações nas condições de iluminação ou à incapacidade de reidentificação de uma pessoa em ambientes multicâmara. A presente dissertação procurou encontrar soluções para alguns desses problemas.

1.2. Enquadramento e Objetivos

A presente dissertação tem como foco o desenvolvimento de uma ferramenta capaz de efetuar a reidentificação de clientes em ambientes multicâmara, através de dados da trajetória desempenhada pelos mesmos ao longo do seu percurso na loja. Enquadra-se no projeto “ECI4.0 – Espaços Comerciais

Inteligentes”¹, que tem como objetivo geral o desenvolvimento de um sistema que permita analisar padrões e características do comportamento humano em espaços comerciais de retalho, recorrendo a técnicas de inteligência artificial e visão por computador. O projeto contou com a colaboração das empresas AXIANSEU Digital Solutions, S.A. e SONAE MC - SERVIÇOS PARTILHADOS, S.A. e da entidade do Sistema I&D ISCTE – Instituto Universitário de Lisboa.

Os dados fornecidos foram obtidos a partir de vídeos de câmaras de videovigilância recorrendo a algoritmos de visão por computador, que conseguem extrair trajetórias a partir de vídeos, estando suscetíveis a erros devido a diversos fatores, como a variação nas condições de luminosidade e/ou a presença de oclusões. Deste modo, será necessário que a ferramenta seja capaz de superar problemas causados por partes da trajetória que possam estar erradas ou não ter sido registadas. A ferramenta terá ainda de ser capaz de agrupar a informação das trajetórias provenientes de câmaras diferentes, (re)construindo a trajetória de cada cliente ao longo das diferentes câmaras por onde este passou, para que no final seja identificada como a trajetória de um único cliente.

Neste contexto, a dissertação pretende responder às seguintes questões de investigação:

- Através de dados de trajetória, como pode ser feita a reidentificação de pessoas presentes num ambiente comercial composto por várias câmaras?
- Quanto pode ser melhorada a identificação de pessoas resultante da aplicação de algoritmos de *tracking* num ambiente composto por várias câmaras, usando apenas os dados extraídos relativos às trajetórias, sem recurso às imagens originais?

Desta forma, o foco da dissertação incide na reidentificação de pessoas através do uso exclusivo de dados de trajetória. Isto permite que a ferramenta criada possa usar dados extraídos a partir de diferentes fontes, sendo possível tirar partido de outras tecnologias usadas para obter trajetórias, para além das trajetórias extraídas a partir de imagens.

1.3. Metodologia

Com o intuito de atingir os objetivos, optou-se por utilizar a metodologia *Cross Industry Standart Process for Data Mining* (CRISP-DM) para guiar a elaboração da dissertação. A escolha deste processo deve-se ao facto de o mesmo providenciar uma estrutura bem estabelecida para projetos na área de análise e exploração de dados.

¹ <https://ciencia.iscte-iul.pt/projects/espacos-comerciais-inteligentes/1736>

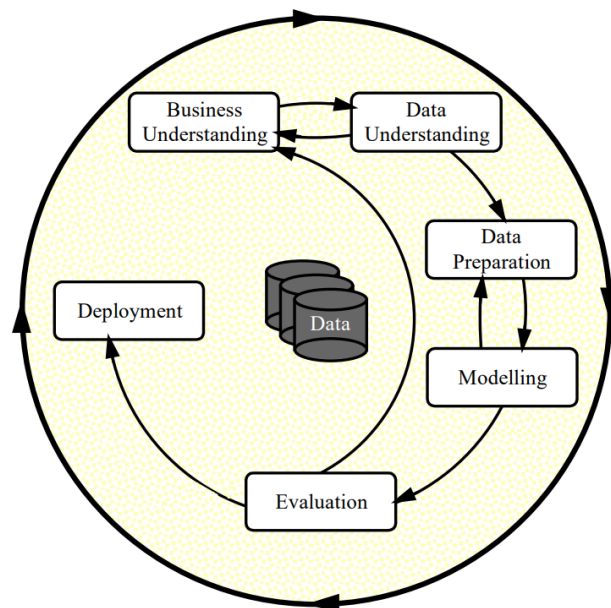


Fig. 1.1 – Modelo CRISP-DM [4]

O modelo CRISP-DM, encontra-se representado na Figura 1.1 e é composto por uma sequência de 6 fases:

- *Business Understanding*: fase focada em perceber e definir os objetivos do projeto, requisitos disponíveis, definidas as tecnologias a ser utilizadas e o planeamento para as restantes fases. Esta fase encontra-se apresentada no Capítulo 2, no qual se apresenta o enquadramento do estudo e revisão de literatura sobre técnicas de extração de trajetórias, deteção e *tracking* de objetos e reidentificação de pessoas.

- *Data Understanding*: fase que consiste em recolher, analisar e descrever os dados a utilizar para atingir os objetivos e verificar a qualidade dos mesmos. Esta parte é descrita na Secção 3.1, onde são apresentados o conjunto de dados e as suas características.

- *Data Preparation*: fase na qual é feita a seleção dos dados que irão ser utilizados, a limpeza e tratamento dos dados devido a questões como a formatação ou a inexistência dos mesmos e a construção de novos atributos tendo como base atributos já existentes. Este processo de tratamento de dados é explicado nas Secções 3.2 e 3.3.

- *Modeling*: fase em que são desenvolvidos e aplicados os algoritmos de reidentificação de pessoas com base na trajetória das mesmas, onde são descritas as experiências realizadas e comparados os resultados obtidos entre as mesmas. A exposição do processo de reidentificação e a apresentação dos resultados é feita nas Secções 3.4 e 4, respetivamente.

- *Evaluation*: fase onde é determinado se os objetivos foram atingidos e se é necessária uma nova iteração pelas fases anteriores ou se se procede para a fase seguinte.

- *Deployment*: última fase do processo, onde é desenvolvido um plano para efetuar a implementação do modelo, onde são apresentados os resultados e é feita uma retrospectiva do que correu bem e do que poderia ser melhorado num trabalho futuro.

As duas últimas fases não se encontram diretamente inseridas na dissertação, uma vez que não fazem parte das tarefas da equipa no projeto. No entanto, no Capítulo 5 são apresentadas as conclusões obtidas, as contribuições do trabalho e indica alguns tópicos considerados interessantes explorar num trabalho futuro.

Revisão da Literatura

Este capítulo apresenta trabalhos relacionados com o estudo que se pretende realizar e encontra-se dividido por secções, sendo as quatro primeiras sobre temas enquadrados no âmbito da dissertação e a última uma síntese do que foi abordado.

2.1 Técnicas de extração de trajetórias

Nos dias de hoje, existem várias técnicas diferentes para obter dados de trajetória de objetos [3]. Para além de pessoas, estes objetos podem ser animais, veículos ou até mesmo fenómenos naturais como furacões ou incêndios. No contexto deste trabalho, os dados de trajetória consistem em coordenadas espaciais que correspondem a um ponto 2D, registadas num determinado instante de tempo, aos quais é associado um identificador do objeto a que correspondem. Ao unir estes pontos cronologicamente é obtida uma aproximação do caminho efetuado pelo objeto.

Os pontos gerados, antes de serem utilizados por algoritmos de IA, devem ser tratados e processados de acordo com as necessidades. Desta forma, são aplicadas técnicas de pré-processamento como o tratamento de pontos cuja localização não faça sentido para o contexto do conjunto de dados (*outliers*) ou a inferência de pontos que não foram registados (*missing values*) de forma a completar a trajetória [3]. Este tipo de informação pode ser aplicado em diferentes contextos de várias formas diferentes, entre as quais se destaca a análise de trajetórias, a descoberta de conhecimento através do caminho percorrido e de características extraídas a partir do mesmo, ou a previsão da posição futura de um objeto.

A trajetória percorrida por um cliente é um tipo de dados que tem vindo a ganhar cada vez mais atenção, devido à sua potencial capacidade explicativa do comportamento das pessoas presentes num espaço comercial. Para tal, existem várias tecnologias que são utilizadas para conseguir obter dados de trajetórias em espaços interiores.

Radio Frequency Identification (RFID) é uma das tecnologias utilizadas e consiste num conjunto de *tags* que emitem frequências rádio com informação sobre a sua localização para o leitor mais próximo. Para além da localização, a informação contém características como o identificador único de cada *tag* ou o instante de tempo em que foi recolhida [5]. J. S. Larson et al. [6] foi um dos primeiros a utilizar dados de trajetórias de pessoas num supermercado, recolhidos através de *tags* RFID. O estudo consistiu numa análise exploratória deste tipo de dados, recorrendo a algoritmos de *clustering*, os quais permitiram analisar aspetos como o tempo que o cliente esteve na loja, o comprimento do caminho percorrido e os corredores que foram visitados.

M. Kholod et al. [7] utilizaram dados recolhidos através de *tags* RFID, presentes nos carros de compras de um supermercado no Japão, e combinaram-nos com os dados de transação de produtos que o cliente comprou para tentarem perceber de que maneira a distância percorrida pelo cliente influencia o seu comportamento de compra. Outros autores [8]–[13] combinam informação dos produtos adquiridos pelo cliente com o caminho percorrido de modo a tentarem prever e analisar o seu comportamento.

Em [5] é proposto um modelo de predição baseado numa rede neuronal *Long Short-Term Memory* (LSTM) O modelo procura prever a localização do próximo ponto da trajetória do cliente, com base em dados históricos dos movimentos do cliente na loja. Foram testados outros modelos (RNN e BPNN) mas o LSTM foi o que obteve melhores resultados. M. Liu et al. [14] recolhem dados de trajetória através de RFID e tentam classificar o comportamento dos clientes com base em técnicas de *Process Mining*. O estudo [15] recorre a dados recolhidos via RFID e a dados respetivos ao comportamento de compra dos clientes para estudar o comportamento dos clientes num espaço comercial de retalho. Numa primeira fase é feito um mapeamento dos dados recolhidos, de seguida são aplicadas técnicas de pré-processamento, de forma a eliminar trajetos redundantes, para depois tentarem extrair padrões dos caminhos percorridos pelos clientes.

Ao contrário dos sistemas de recomendação mais comuns, que se baseiam na frequência de compra de um determinado produto, os autores do estudo [16] propõem um sistema de recomendação que tem como base o comportamento dos clientes. Este utiliza dados recolhidos via RFID, que são posteriormente combinados com informação dos produtos comprados por uma determinada pessoa, para treinar um sistema de recomendação que tem como base a similaridade do tempo que os clientes despendem na loja. Já no estudo [17], os autores avaliaram a influência do tempo que um cliente se encontra parado tem no seu comportamento de compra. Concluíram que tem uma influência positiva forte quando a pessoa se encontra presente em determinadas áreas do supermercado, como a zona da carne ou a zona de produtos frescos.

Outra solução para extrair informação acerca do caminho percorrido por um cliente é o Bluetooth. A maioria das arquiteturas consiste num conjunto de *beacons* Bluetooth, presentes por exemplo numa loja, que emitem sinais de baixa energia para os smartphones dos clientes e devolvem informações como o ID associado ao dispositivo em questão, a distância aproximada do *beacon* e a data e hora correspondentes ao registo [18]. Trabalhos recentes [19], [20] utilizaram dados obtidos a partir de tecnologias baseadas em Bluetooth para tentar prever o género das pessoas presentes num centro comercial, através de informação relativa ao caminho percorrido pelas mesmas. D. Zaim et al. [18] analisam as trajetórias dos clientes através da técnica de *clustering Longest Common Subsequence* (LCSS). Tal como nos trabalhos anteriores, estes autores utilizam *beacons* Bluetooth para recolher os dados e formar as trajetórias percorridas pelos clientes. De seguida, aplicam o algoritmo de

clustering LCSS para agrupar as trajetórias e, por fim, procuram extrair informações acerca do comportamento dos clientes, designadamente as áreas mais visitadas da loja ou as trajetórias principais dos clientes. M. Oziom et al. [21] criaram uma aplicação de telemóvel que indica a localização de um determinado produto na loja, assim como a distância do cliente a esse produto. Para tal, recorreram a *beacons* Bluetooth de forma a obter a posição do cliente na loja. Através dos registos do caminho efetuado pelos clientes no processo de teste da aplicação, foi permitido analisar alguns comportamentos das pessoas que testavam a aplicação, como, por exemplo, se o cliente passava por um produto sem se aperceber, ou se o trajeto tinha sido influenciado pela proximidade da pessoa aos produtos desejados.

O Wi-Fi é uma opção alternativa às técnicas referidas anteriormente. H. Hwangbo et al. [22] tentaram encontrar os caminhos percorridos pelos clientes que geraram um maior número de vendas, de forma a melhorar o layout da loja. Para tal, combinaram dados da posição do cliente na loja, obtidos através de Wi-Fi, com dados de transação de produtos associados ao mesmo. Também extraindo dados da trajetória através de Wi-Fi, os autores [23] tentam prever o género e a idade das pessoas presentes num centro comercial, com base nos dados recolhidos e no tipo de lojas visitadas. No estudo [24] foi desenvolvido um novo método de *clustering* designado E-DBSCAN que tem como objetivo analisar o comportamento dos clientes em espaços interiores, através de dados de trajetória obtidos via Wi-Fi.

Existem ainda estudos [25]–[27] que utilizam outras abordagens para a recolha de dados, como tecnologias baseadas em *ultrawide-band* ou sensores de profundidade. Em [26], os autores recorreram à tecnologia *Real Time Locating System* (RTLS) para encontrar as zonas da loja mais atrativas, com base no caminho percorrido pelos clientes e no tempo de permanência dos mesmos numa determinada área.

Y. Kamiya et al. [28] propõem um novo sistema de posicionamento interior baseado em *map matching*. Usam uma rede LSTM para reconhecer algumas atividades importantes para a localização das pessoas e, em conjunto com a sua trajetória inicial, são integradas num modelo chamado *Hidden Markov Model* (HMM) de forma a obter uma posição mais precisa de cada pessoa. A solução proposta apresentou um erro médio de aproximadamente 2 metros.

Para além das tecnologias mencionadas anteriormente, a extração de trajetórias de clientes através de vídeo tem sido cada vez mais utilizada. Isto deve-se em parte à evolução recente dos algoritmos de deteção e *tracking* de pessoas, permitindo obter uma trajetória cada vez mais fiável, mas também ao facto de ser uma solução fácil de implementar em espaços comerciais devido à presença de câmaras de videovigilância para fins de segurança. Os autores do estudo [29] recorreram a trajetórias obtidas a partir de vídeo, em conjunto com características extraídas das mesmas como a

velocidade e a aceleração, para detetar segmentos da trajetória onde os clientes revelem um comportamento de compra ou de interesse num determinado produto.

Outra vantagem da extração de trajetórias através de vídeo é a possibilidade de ser facilmente complementada com outras informações extraídas através do vídeo, como por exemplo as ações que o cliente está a desempenhar num determinado momento, que não seriam possíveis de obter através de meios como RFID, Bluetooth ou Wi-Fi. M. C. Popa et al. [30] recorrem a câmaras de forma a extrair características da trajetória e das ações desempenhadas pelas pessoas, a fim de tentar classificar o seu comportamento. Para tal, os autores dividem a loja em várias zonas: entrada, zona de descanso, zona de produtos, zona de passagem, provadores, zona com espelho e a zona de pagamento. Com base nesta informação, nas características extraídas da trajetória e nas ações desempenhadas pela pessoa, criam um modelo comportamental que permite dividir os clientes em quatro categorias: *Disoriented*, quando o cliente não sabe o que procura ou o onde encontrar, movendo-se de um sítio para outro sem um plano aparente; *Looking for Support*, quando o cliente vai de encontro ao assistente de loja para pedir ajuda; *Goal Oriented*, quando o cliente sabe o que procura e vai de encontro ao produto; *Looking Around*, quando o cliente vê os produtos mas não procura nada em especial. O estudo apresentou resultados acima dos 93% para classificação de trajetórias e reconhecimento de comportamentos com base no modelo proposto.

2.2 Comportamentos extraídos através de vídeo

A análise de vídeo recorrendo a técnicas de IA, mais concretamente algoritmos de visão por computador, tem sido uma ferramenta utilizada para analisar o comportamento das pessoas em espaços comerciais. Para além do caminho percorrido pelo cliente, é possível extrair informações como as ações desempenhadas pelo mesmo, se se encontra sozinho ou em grupo ou para onde está a olhar.

O artigo [2] explora o uso de tecnologias de observação em ambientes de retalho, como por exemplo câmaras de videovigilância, e em como a sua utilização pode ter impacto na compreensão do comportamento dos consumidores. Os autores analisam os benefícios e desvantagens deste tipo de tecnologias e concluem que estas podem auxiliar na tomada de decisão e na organização dos produtos das superfícies comerciais, desde que sejam respeitadas as questões éticas associadas à sua utilização. Analisam ainda o comportamento de compra do consumidor com base na escolha do tipo de equipamento que o mesmo usa para transportar os produtos (carro, cesto, nenhum). Para além deste artigo, o estudo [31] analisa o impacto que o equipamento que os clientes usam para transportar os produtos tem no seu comportamento. Os autores afirmam que é uma boa variável para segmentar os consumidores pois explica de uma forma consistente alguns dos seus comportamentos, como a

velocidade a que se deslocam, o tamanho do caminho percorrido, o tempo despendido na loja ou o número de produtos adquiridos.

M. Popa et al. [32] apresentam um sistema de classificação do comportamento de clientes numa loja, através de imagens de vídeo. O protótipo desenvolvido consegue detetar movimento, fazer a extração da trajetória com base no *tracking* e localizar a cara dos clientes. A classificação é feita com base em três modelos comportamentais pré-definidos: *goal oriented shopper*, corresponde ao cliente que tem uma lista de produtos que pretende comprar e vai direto ao sítio dos mesmos, *disoriented shopper*, que não sabe os produtos que procura ou onde os encontrar, *looking for the shop assistant*, que como o nome indica procura um assistente de loja para pedir ajuda, *fun shopper*, que procura demonstrações de produtos e gosta de fazer compras nos sítios mais sobrelotados, e *duo-shopper* quando o cliente se encontra acompanhado pelos filhos. Através da velocidade de andamento conseguiram classificar os clientes em dois modelos comportamentais (*disoriented* e *goal orientend*).

R. Arroyo et al. [33] também tiveram como foco a análise do comportamento dos clientes dentro de uma loja, mas desta vez orientada a detetar comportamentos suspeitos desempenhados pelos mesmos. Para tal, propõem um sistema em tempo real, capaz de capturar a trajetória das pessoas e identificar ações questionáveis num ambiente comercial. Neste sistema tentam detetar pessoas que permanecem muito tempo numa zona de risco ou pessoas que se encontrem próximas de uma zona de pagamento quando não existe um funcionário por perto. Caso uma destas situações aconteça, é emitido um aviso de segurança.

J. Liu et al. [34] tentam classificar o comportamento dos clientes que estejam próximos de uma prateleira, de acordo com seis categorias: sem interesse, ver, virar para a prateleira, tocar, pegar e voltar a colocar na prateleira e pegar e colocar no cesto. De forma a reconhecer qual a classificação mais adequada num determinado momento, são avaliados aspetos como a orientação da cabeça e do corpo, para tentar perceber onde está focada a atenção do cliente, e os movimentos dos braços, para detetar a interação com os produtos nas prateleiras.

2.3 Detecção e *Tracking* de Objetos

A deteção de objetos consiste numa técnica de visão por computador, que tem como objetivo localizar os objetos presentes numa determinada imagem ou vídeo. Após a deteção, os algoritmos de *tracking* são utilizados para seguir os movimentos dos objetos identificados, ao longo de uma sequência de *frames*. Desta forma, é possível associar a deteção de um objeto, num determinado conjunto de *frames*, a um único identificador.

No que diz respeito à tarefa de deteção de objetos, os modelos de aprendizagem profunda baseados em redes neuronais convolucionais (CNN), como Faster R-CNN [35], YOLO [36] e SDD [37],

têm obtido bons resultados em diversas aplicações. Em relação ao *tracking* de objetos, o modelo DeepSORT [38] e mais recentemente o ByteTrack [39] são alguns dos algoritmos que têm apresentado bons desempenhos.

A integração destas soluções nas superfícies comerciais pode ajudar em tarefas como a gestão de inventários [40] ou a otimização do layout da loja [41]. Para além disso, são um elemento necessário para o processo de extração de trajetórias através de vídeo, que é feito com base na informação fornecida por estes algoritmos.

2.4 Reidentificação de Pessoas

No contexto da dissertação, reidentificação de pessoas entende-se pela uniformização de diferentes identificadores, obtidos a partir de câmaras diferentes e correspondentes a uma única pessoa, de forma a obter um identificador único representativo da pessoa em questão. Ao contrário do conceito de reidentificação de pessoas associado a algoritmos de *tracking* de objetos que se baseia em características da imagem, a reidentificação na dissertação é feita unicamente com base em dados de trajetória.

No que diz respeito a estudos relacionados com re-identificação de pessoas através de dados de trajetórias, alguns autores já abordaram temas semelhantes. Em [42], os autores propõem um algoritmo para identificar e seguir objetos capturados em imagens provenientes de várias câmaras. Utilizam polinómios *B-Spline* para calcular e definir a trajetória dos objetos, permitindo uma representação mais suave quando os mesmos se encontram em movimento. Em seguida, recorrem a transformações de homografia para comparar as trajetórias capturadas por diferentes câmaras e reidentificar os objetos correspondentes com base na semelhança de trajetórias. Este método, ao contrário do método descrito em seguida, necessita de encontrar semelhanças nas trajetórias de diferentes câmaras para voltar a identificar uma pessoa.

Em [43], os autores criaram um conjunto de dados para Multi-Target Multi-Camera Tracking (MTMCT) a partir do jogo GTA V. O conjunto de dados intitulado Multi Camera Track Auto (MTA), é composto por um conjunto de vídeos gravados em 6 câmaras diferentes, com mais de 100 minutos de vídeo por câmara e com a presença de mais de 2800 *bots* (identidades de pessoas). Para além da construção do conjunto de dados, propuseram também um sistema de deteção, reidentificação e *tracking* de *bots*, bem como o cálculo da distância percorrida por estes e a associação de trajetórias. O processo de associação de trajetórias utiliza um método de *clustering* para agrupar trajetórias de *bots* com base numa distância. Esta distância é calculada através da agregação ponderada de 5 distâncias individuais ou restrições relacionadas com a presença de trajetos numa câmara e em várias câmaras,

uma distância de correspondência de homografia, um desconto de previsão linear e uma característica de aparência.

Os autores de [44] propõem um método para fazer o *tracking* de pessoas em cenas com muitos indivíduos. Recorrem à utilização de múltiplas câmaras de modo a ultrapassar obstáculos introduzidos por oclusões. O método proposto consiste em detetar as cabeças das pessoas presentes na cena para extrair a trajetória realizada, com o objetivo de tentar superar dificuldades introduzidas por oclusões, uma vez que não é necessário reconhecer o corpo na sua totalidade para detetar a presença de uma pessoa nesse local. Em alguns casos, em que os rostos das pessoas têm de ser anonimizados de forma a proteger a privacidade dos indivíduos envolvidos, pode ser uma desvantagem pois poderia aumentar a dificuldade em detetar a cabeça das pessoas.

No estudo [45], os autores propõem um método que deteta pessoas em vídeos de câmaras de videovigilância e gera um mapa de calor que ilustra os locais mais visitados na loja. O algoritmo de deteção utilizado foi o YOLOv5 e utilizam uma matriz de homografia para encontrar as coordenadas de projeção das pessoas presentes na loja.

2.5 Síntese

Apesar da evidente importância que a análise de trajetórias aparenta ter para a compreensão do comportamento do consumidor e para o desenvolvimento do setor de retalho, a literatura existente referente a este tópico é ainda limitada. O trabalho desenvolvido no âmbito desta Dissertação apresenta um carácter de inovação no sentido em que pretende explorar a reidentificação de pessoas em ambientes comerciais multicâmara, utilizando os dados de trajetória dos clientes para atingir o objetivo proposto. A Tabela 2.1 seleciona e sintetiza alguns dos artigos mais relevantes.

Grande parte dos estudos revistos, que utilizam dados de trajetórias em ambientes comerciais, focam-se na previsão da trajetória que o cliente irá desempenhar, na análise do caminho percorrido para melhorar o layout da loja, na combinação de dados de trajetória com dados de checkout para prever o comportamento de compra, na classificação do género com base na trajetória desempenhada e no *clustering* de trajetórias com o objetivo de achar as zonas e trajetos mais utilizados pelo consumidor. Por outro lado, alguns dos estudos encontrados no domínio da análise do comportamento de clientes em ambientes de retalho centram-se em definir modelos comportamentais de forma a classificar os clientes com base nos modelos definidos. Por fim, relativamente aos estudos que abordam questões relacionadas com a reidentificação de pessoas, estes vão ao encontro do objetivo da dissertação, na medida em que procuram desenvolver técnicas para associar trajetórias pertencentes a um determinado cliente, mediante a sua presença num ambiente composto por várias câmaras.

Neste contexto, a presente dissertação acrescenta valor à investigação de reidentificação de pessoas em ambientes multicâmara, através do desenvolvimento de uma ferramenta que permita reidentificar pessoas neste tipo de ambiente e explora a utilização exclusiva de dados de trajetória, extraídos a partir de vídeo e obtidos num ambiente de retalho real, com condicionantes ao nível da necessidade de processar informação anonimizada.

Tab. 2.1 - Seleção de artigos relacionados

Estudo	Descrição e Conclusões
[5]	Utilizando dados de trajetória recolhidos através de RFID, propõem um modelo que tem como objetivo prever a localização futura do cliente com base em dados passados. O modelo é baseado em redes neuronais LSTM e apresenta resultados superiores aos modelos RNN e BPNN.
[6]	Apresenta uma análise exploratória de dados de trajetória recolhidos através de RFID. Tiram conclusões em relação ao tipo de percurso dos consumidores e a aspetos que devem ter sido em conta na disposição dos produtos na loja.
[9]	Combina dados obtidos através de RFID com informação dos produtos comprados pelo cliente para estudar o impacto que o tamanho da trajetória tem no número de itens comprados. Concluíram que quanto maior o caminho percorrido pelo cliente, maior o número médio de produtos comprados e maior a diversidade desses produtos.
[19]	Através da trajetória efetuada pelos clientes num centro comercial, comparam as diferenças de comportamento entre géneros em aspetos como o tempo despendido, o caminho percorrido e o tipo de lojas visitadas.
[2]	Discute o uso das tecnologias de observação para estudar o comportamento dos consumidores em loja. Exploram o impacto que o equipamento usado para transportar produtos tem no comportamento de compra do consumidor.
[32]	Criam modelos comportamentais de clientes através da identificação de comportamentos em vídeos. Desenham e implementam um protótipo com módulos de deteção de movimento, análise de trajetória e localização e seguimento da cara dos clientes. Através da análise da velocidade dos clientes classificam os mesmos em duas categorias: orientados a um objetivo e desorientados.
[42]	Propõem um algoritmo para identificar e seguir objetos capturados em imagens provenientes de várias câmaras. Utilizam polinómios <i>B-Spline</i> para calcular a trajetória e

	recorrem a transformações de homografia para reidentificar objetos com base na semelhança de trajetórias.
[43]	Criam um processo de associação de trajetórias que utiliza um método de <i>clustering</i> para agrupar trajetórias de <i>bots</i> com base numa distância calculada através da ponderação de distâncias individuais e restrições.
[45]	Utilizam o algoritmo YOLOv5 para fazer a deteção de pessoas numa loja. Através da homografia, fazem a projeção das coordenadas das pessoas e geram um mapa de calor dos sítios mais visitados.

Método de Reidentificação

Neste capítulo encontra-se descrito o esquema geral da ferramenta de reidentificação desenvolvida, descreve-se o conjunto de dados utilizado, bem como o contexto em que o mesmo se encontra inserido. Por fim, é feita uma descrição mais em detalhe de cada fase do *pipeline*, nomeadamente, a preparação de dados, a extração de trajetórias e o processo de reidentificação.

3.1. Esquema Geral

A ferramenta de reidentificação consiste num conjunto de quatro etapas, necessárias para conseguir associar um cliente, presente em vídeos de várias câmaras, a um único identificador. As etapas que constituem o *pipeline* encontram-se ilustradas na Figura 3.1.



Fig. 3.1 – Esquema geral do pipeline das várias partes do processo de reidentificação

Resumidamente, o processo começa por decompor em *frames* cada vídeo capturado pelas câmaras de videovigilância. A cada *frame* foi aplicado um algoritmo de deteção de objetos que procura localizar as pessoas presentes no mesmo. Com base nas *bounding boxes* geradas pelo detetor de objetos, é utilizado um algoritmo de *tracking* de objetos que atribui um identificador numérico (ID) a cada pessoa. De seguida, são extraídos os pontos de trajetória para cada ID e calculada a sua projeção num plano visto de cima, correspondente à planta da loja. Com a informação das trajetórias guardada, é aplicado o tratamento de dados necessário para levar a cabo o processo de reidentificação. Por fim, são aplicadas as duas fases que constituem o processo de reidentificação, de forma a obter um ID de cliente único e transversal a todas as câmaras.

A implementação da ferramenta descrita encontra-se disponível no GitHub através do link: <https://github.com/DiogoMendess/master-thesis>

3.2. Descrição do Conjunto de Dados

O conjunto de dados fornecido é composto por diferentes tipos de dados.

Primeiro, a informação das trajetórias efetuadas pelos clientes na loja foi disponibilizada através de ficheiros em formato .CSV, divididos por dias e pelo número da câmara onde foram extraídos. A estrutura de cada ficheiro era composta por 22 características, das quais foram selecionadas as que têm interesse para o problema de investigação. Desta forma, foram consideradas relevantes as seguintes características: i) as coordenadas da *bounding box* utilizada para calcular o ponto de trajetória; ii) as coordenadas do ponto de trajetória após ter sido feita a projeção para o plano 2D; iii) o identificador do cliente a que pertence o ponto de trajetória; iv) o tempo e o número do *frame* do registo.

Segundo, o conjunto de dados incluiu vídeos a partir dos quais foram extraídos os pontos de trajetória. Estes vídeos foram gravados por 5 câmaras de videovigilância presentes numa superfície comercial de retalho, mas apenas abrangem uma pequena parcela da área total. As gravações correspondem a 7 dias consecutivos e foram disponibilizados com uma resolução 1080p e *frame rate* de 20 *fps*, contabilizando um total de 91 horas. Os vídeos encontram-se acelerados, ou seja, cada segundo do vídeo corresponde a 6 segundos de tempo real. Foi aplicado um algoritmo de introdução de desfocagem na zona da face dos clientes, de forma a anonimizar as pessoas presentes na loja. Para o objetivo do trabalho apenas foram selecionados dados de câmaras cujos campos de visão se sobreponham parcialmente. Por este motivo, não foram utilizados dados relativos a uma câmara das câmaras disponíveis. A Figura 3.2 representa um *frame* de um vídeo, capturado a partir de uma das câmaras consideradas elegíveis.



Fig. 3.2 – Exemplo do frame de um vídeo

Terceiro, foram ainda fornecidos: i) um mapa da planta da loja, que inclui a disposição das prateleiras e a localização aproximada de cada câmara; ii) a informação da homografia de cada câmara,

em forma de uma matriz de projeção, que permite a transformação de pontos do vídeo para pontos no plano 2D (plano visto de cima – planta); e iii) um ficheiro que contém informação sobre a hora de início de gravação de cada câmara. A Figura 3.3 ilustra a planta da área considerada da loja, assim como a localização das câmaras consideradas. As zonas pintadas a preto e a cinzento correspondem à delimitação das prateleiras e do mapa, respetivamente, e a azul as Áreas de Saída/Entrada.

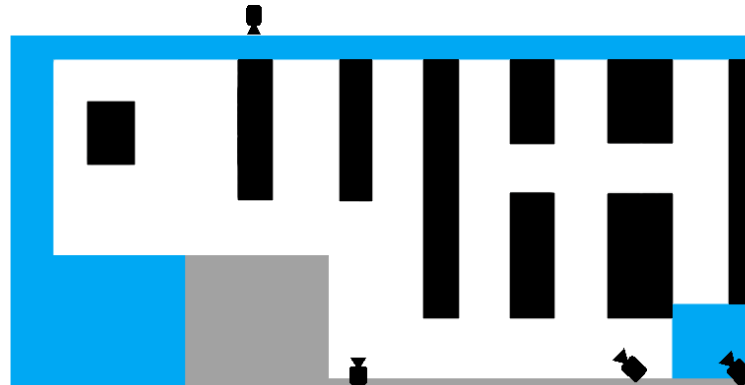


Fig. 3.3 – Planta da área considerada da loja e localização das câmaras

3.3. Preparação de Dados

O processo de preparação de dados começa com a análise da disposição dos pontos de trajetória num eixo cartesiano, para perceber se a forma da mancha gerada por esses pontos corresponde à forma do *layout* apresentado pela planta da loja. A Figura 3.4 é o resultado obtido após a projeção dos pontos de todas as câmaras num eixo cartesiano, onde cada cor corresponde à câmara de onde foram obtidos os mesmos.



Fig. 3.4 – Pontos projetados num eixo cartesiano, coloridos por câmara

Como podemos observar na Figura 3.4, os espaços vazios que correspondem aos locais onde se encontram prateleiras, diferem de sítio de câmara para câmara, ficando preenchidos com pontos de outras câmaras uma vez que não se encontram corretamente posicionados. É então necessário efetuar transformações no valor das coordenadas de cada ponto para tentar ir ao encontro do layout apresentado na planta da loja e conseguir que uma pessoa seja representada no mesmo sítio, independentemente da câmara de onde foi obtida a trajetória. Para tal, começa-se por criar uma *feature* que nos indica o nome da câmara e associa-se a cada ponto de trajetória a câmara utilizada para extrair o mesmo. De seguida, é feita a concatenação de todos os ficheiros .CSV de um determinado dia num único ficheiro. Para que a nova escala de pontos corresponda com os pixéis da planta da loja, é feita uma transformação da escala das coordenadas e altera-se a coordenada da origem do conjunto de pontos correspondente a cada câmara.

Com o objetivo de confirmar as transformações efetuadas, recorreremos aos vídeos de onde foram extraídos os pontos de trajetória para verificar se as projeções iam ao encontro do percurso percorrido pelas pessoas nos vídeos.

Ao tentar fazer estas verificações percebemos que o tempo associado aos pontos de trajetória, não é igual ao tempo apresentado no vídeo. Reparou-se também que o tempo dos vídeos não está sincronizado de câmara para câmara, isto é, nas câmaras 1.03, 1.04 e 1.05 um determinado acontecimento X (por exemplo, uma pessoa a passar num determinado sítio da loja) acontece ao mesmo tempo em todas as câmaras, enquanto na câmara 2.05 o mesmo acontecimento X acontece num tempo diferente (mais 4 minutos e 18 segundos que o tempo das outras câmaras). O segundo problema é relativamente simples de resolver, sendo que apenas seria necessário subtrair 4 minutos e 18 segundos ao tempo dos registos feitos pela câmara 2.05. A questão principal deve-se ao facto do tempo dos registos não corresponder ao tempo apresentado no vídeo.

Posto isto, numa primeira abordagem verificou-se apenas um atraso no tempo de registo das trajetórias. Para calcular o tempo de atraso foi utilizado o primeiro vídeo da câmara 1.03 onde aparecia uma pessoa e comparado com os pontos do cliente da primeira entrada (número 13) do ficheiro .CSV, ordenado cronologicamente pelo campo “time”. Para além do registo do tempo não coincidir com o tempo do vídeo, como foi referido anteriormente, também os pontos que compunham a trajetória não correspondiam à trajetória efetuada pelo cliente no vídeo. Analisou-se outros momentos, e também vídeos e trajetórias de câmaras diferentes, tendo concluído que não existia nenhum atraso constante ou padrão que pudesse justificar o erro no campo do tempo. A Tabela 3.1 demonstra exemplos de resultados das experiências efetuadas. Note-se que a diferença entre os tempos não é constante e isso acontece mesmo dentro do mesmo vídeo.

Tab. 3.1 – Tentativas de sincronização do tempo das coordenadas

Câmara	Tempo no vídeo	Tempo nas coordenadas	Diferença	Vídeo
1.03	09:42:43	09:30:13	00:12:30	1_11
1.03	23:15:28	23:30:32	00:15:04	23_17
1.04	09:33:04	09:28:22	00:04:42	1_52
1.04	08:42:51	08:31:25	00:11:26	1_27
1.05	09:33:09	09:30:15	00:02:54	1_70
1.05	09:36:23	09:37:17	00:00:54	1_71
2.05	07:53:03	07:54:44	00:01:41	1_1
2.05	07:53:35	07:56:56	00:03:21	1_1
2.05	08:15:49	08:03:57	00:11:52	1_12
2.05	08:17:44	08:05:40	00:12:04	1_13
2.05	08:19:26	08:06:32	00:12:54	1_14

Após as experiências mencionadas anteriormente, decidiu-se ordenar o ficheiro pelo “Customer_id”, o que nos levou a conseguir obter a ordem correta dos acontecimentos, ou seja, a ordem dos registos presentes no ficheiro .CSV corresponde agora à ordem pela qual as pessoas aparecem no vídeo. Com isto concluí-se que o campo do tempo estava mal anotado e que ao ordenar pelo tempo, a ordem dos pontos não corresponde à realidade. É então necessário encontrar uma forma de calcular um novo tempo, que possa ser referido da mesma forma independentemente da câmara de onde tenha sido registado o ponto de trajetória, e que vá ao encontro do tempo que aparece no vídeo.

Numa primeira abordagem, foi anotado o primeiro *frame* em que um cliente aparecia e calculado o tempo dos restantes *frames* onde o cliente estava presente com base no número de *frames* por segundo a que os vídeos se encontram gravados. Esta solução, para além de não ser muito precisa, uma vez que podem ser introduzidos erros na escolha do primeiro *frame* em que o cliente aparece, também consumiria muito tempo dado o volume de dados considerado.

Desta forma, uma alternativa para obter os dados de trajetória com o tempo corretamente anotado seria recorrer a algoritmos de deteção e *tracking* de pessoas para implementar uma solução que nos permitisse fazer a extração das trajetórias dos clientes, a partir dos vídeos que nos foram fornecidos. Os algoritmos utilizados bem como outros detalhes de implementação são apresentados na Secção 3.4.

Feita a extração dos pontos de trajetória e a respetiva projeção dos mesmos num plano 2D correspondente à planta da loja, são aplicadas as transformações de coordenadas mencionadas no início desta secção. Como a matriz de projeção utilizada é a mesma que tinha sido utilizada para obter os dados fornecidos nos ficheiros .CSV, continua a ser necessária uma transformação de escala de forma a ser coerente com o *layout* apresentado na planta da loja. As transformações mencionadas

bem como o tratamento de dados apresentado no resto desta secção compõem a fase de preparação de dados destacada na Figura 3.5.



Fig. 3.5 – Esquema do processo de Preparação de Dados

Tendo por base a trajetória de cada pessoa identificada numa câmara (ainda com alguns erros devido a oclusões), é feito o cálculo do tempo respetivo a cada ponto de trajetória. Sendo que as gravações de cada dia se encontram divididas em pequenos segmentos de vídeo, este cálculo foi feito com base em duas *features*, o número do segmento de vídeo e o *frame* a que foi registada a deteção do cliente. Cada segmento tem uma duração (L) de 36 segundos a uma taxa de 20 imagens por segundo (*fps*), o que corresponde a 120 segundos de tempo acelerado. Desta forma, o fator de aceleração (a) associado é calculado por $720/120 = 6$.

$$\frac{(vn-1) \times L \times fps + df}{a} = s \quad (1)$$

O tempo de deteção é calculado com base na fórmula anterior, onde vn é o número do segmento da deteção, L é a duração do segmento, fps é a taxa de *frames* a que o vídeo foi gravado, df é o número do *frame* no segmento onde foi feita a deteção, a é a taxa de aceleração e s corresponde ao número de segundos desde o momento em que a câmara começou a gravar. Posto isto, o número de segundos resultante da aplicação da fórmula apresentada anteriormente é adicionado ao instante de tempo em que a câmara começou a gravar de forma a obter o tempo em que a deteção foi feita.

Considerando o exemplo em que uma determinada pessoa, detetada no segmento número 4, no *frame* número 300, e cuja câmara que capturou esta pessoa começou a gravar às 09:00:00. Então, o tempo associado a esta deteção é calculado como se segue:

$$\frac{(4-1) \times 36 \times 20 + 300}{6} = 410 \text{ segundos} \quad (2)$$

$$09:00:00 + 410 \text{ segundos} \rightarrow 09:06:50 \quad (3)$$

Como referido anteriormente, no caso dos vídeos da câmara 2.05, estes apresentam um erro de sincronização. Desta forma, o mesmo tem de ser tido em consideração no cálculo do tempo de coordenadas extraídas a partir de vídeos da câmara em questão, subtraindo ao tempo final o acerto necessário (4 minutos e 18 segundos).

A atribuição do ID de cliente é feita pelo algoritmo de *tracking*, que associa um identificador a um conjunto de deteções. Uma vez que o *tracker* foi aplicado individualmente aos vídeos de cada câmara, grande parte dos identificadores criados eram duplicados (porque os IDs de cada câmara começavam no número 1). De forma a garantir que cada ID é único entre diferentes câmaras, transformámos o ID de cliente adicionando o último ID de uma câmara aos IDs da câmara seguinte.

Tendo os dados projetados no mapa da loja, tornou-se óbvio que pontos muito afastados da posição da câmara não eram precisos quando comparados com a trajetória realizada pelo cliente. Estes pontos podem resultar de erros no cálculo da matriz de homografia ou devido a serem consideradas deteções de baixa confiança. Para além disso, estes pontos não são relevantes para os objetivos propostos, pelo que decidimos limitar a área de interesse para reidentificação.

Para além da limitação mencionada, existe uma área da loja que o resultado da projeção é diferente do esperado, devido à presença de objetos (prateleiras) que causam a oclusão de clientes no campo de visão de uma das câmaras. Uma vez que a zona se encontra coberta por duas câmaras diferentes e que o campo de visão de uma das câmaras não é afetado por este tipo de problemas devido à inexistência de oclusões, decidimos remover os pontos da câmara afetada, mantendo apenas os pontos da câmara que tem boa visibilidade das pessoas que circulam nessa área. A área de pontos afetada por oclusões foi definida através da análise dos vídeos de onde foram extraídas as trajetórias e das coordenadas que apresentavam translações no plano da planta da loja. Esta medida tem como objetivo obter uma trajetória mais fidedigna e aproximada da realidade. As Figuras 3.6 e 3.7 ilustram o mesmo exemplo após aplicar o algoritmo de reidentificação, sendo que na primeira figura não foram removidos os pontos afetados por oclusões. Cada cor representa um ID de cliente diferente e são aleatoriamente geradas para cada exemplo.



Fig. 3.6 – Exemplo antes de serem removidos os pontos afetados por oclusões



Fig. 3.7 – Exemplo depois de serem removidos os pontos afetados por oclusões

Por último, com o intuito de melhorar a qualidade dos dados disponíveis, decidimos remover os pontos identificados com o mesmo ID e cuja posição não sofre qualquer tipo de alterações, ou seja, remover os pontos de um cliente que se encontre na mesma posição em todas as detecções associadas ao mesmo. Isto visa resolver falsos positivos originados pelo detetor de objetos, onde por exemplo uma cadeira é indevidamente identificada como se fosse uma pessoa. A posição dos pontos pertencentes a esse ID não irá ser alterada ao longo do tempo, sendo considerado um bom critério de exclusão. Para além disso, o contrário é sempre verificado caso a deteção corresponda a uma pessoa pois esta irá mover-se ao entrar ou sair do plano da câmara. A Figura 3.8 demonstra um exemplo onde existe um conjunto de pontos no mesmo sítio (pontos marcados a vermelho), que são removidos ao aplicar o tratamento de dados.

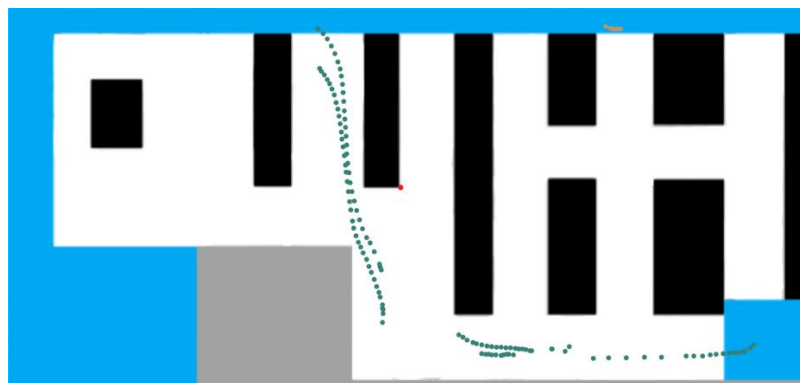


Fig. 3.8 – Exemplo com pontos falsos positivos (marcados a vermelho)

3.4. Extração de Trajetórias

Nesta secção são descritas com mais detalhe as fases que compõem o processo de extração dos dados de trajetória, sendo estas a decomposição dos vídeos em *frames*, a deteção de pessoas presentes em cada *frame*, o *tracking* das mesmas ao longo do vídeo e o cálculo da trajetória e correspondente projeção recorrendo à homografia de cada câmara. Andámos para trás na ordem de eventos apresentada no esquema geral porque foi necessário refazer o processo de aquisição de trajetórias devido a incoerências presentes nos dados fornecidos, como foi referido na Secção 3.3. A Figura 3.9 ilustra as fases do processo de extração de trajetórias e o seu enquadramento no esquema geral.



Fig. 3.9 – Esquema do processo de Extração de Trajetórias

3.4.1 Deteção de Objetos

Para fazer a deteção de cada cliente presente nos *frames* dos vídeos, decidimos utilizar um detetor de objetos amplamente utilizado chamado You Only Look Once (YOLO), proposto pela primeira vez em [36]. A arquitetura deste modelo baseia-se numa CNN e é capaz de prever simultaneamente várias

bounding boxes, bem como as probabilidades de classe associadas a cada uma delas. O nome deste algoritmo está relacionado com o facto de apenas precisar de analisar cada imagem uma vez para inferir os objetos presentes e a localização dos mesmos. As principais vantagens em relação a outros detetores de objetos, como o R-CNN, são o facto de ser mais rápido, poder ter em conta o contexto em que os objetos aparecem na imagem e ser muito generalizável na forma como aprende as representações de cada objeto.

Foram propostas diferentes variações do algoritmo, sendo a versão 8 a sua versão mais recente. Para o trabalho proposto, decidimos utilizar a versão 5 (YOLOv5) dado que se encontra estável, apresenta bom desempenho, é adequada ao poder computacional que temos disponível e pode ser facilmente acedida através do repositório de modelos pré-treinados PyTorch Hub.

3.4.2 *Tracking de Objetos*

O ByteTrack foi proposto em [39] e representa o estado da arte atual no que diz respeito a algoritmos de *tracking* de objetos. Segue o paradigma *tracking-by-detection*, que consiste em utilizar *bounding boxes* geradas por algoritmos de deteção de objetos para criar uma associação entre cada uma delas ao longo do tempo, através de IDs numéricos.

Para tirar partido de todas as deteções geradas pelos detetores de objetos, o ByteTrack separa-as com base em níveis de confiança (baixo ou alto). A decisão de incluir as deteções com níveis de confiança baixos deve-se ao facto de, segundo os autores, muitas delas serem resultado da presença de oclusões, o que não invalida a sua utilidade. Tendo isto em conta, as *bounding boxes* com confiança superior a um determinado limiar são associadas às previsões efetuadas pelo Filtro de Kalman [46], com base na semelhança de movimento (IoU) ou de aparência. De seguida, o Método Húngaro [47] é utilizado para atribuir os IDs, com base na informação de semelhança recebida. Se as associações não forem bem-sucedidas, o processo repete-se através de IoU, utilizando agora as deteções com confiança abaixo do limiar, de modo a tentar resolver casos que possam ter sido originados por oclusões ou por alterações no plano de fundo.

Relativamente aos parâmetros do ByteTrack, foram efetuadas várias experiências com diferentes configurações, e os valores que obtiveram melhores resultados são os que se apresentam na Tabela 3.2.

Tab. 3.2 – Parâmetros ByteTrack

track_buffer	fps
match_thresh	0.8
track_thresh	0.5
aspect_ratio_thresh	1.6
min_box_area	10

É importante mencionar que tentámos alterar os pesos de posição e velocidade do Kalman Filter, mas estas modificações não produziram melhorias consistentes, pelo que decidimos manter os valores originais. A implementação do algoritmo ByteTrack encontra-se disponível no repositório do autor, no Github [39].

Em termos de *trackers* de objetos, também foi considerado o algoritmo DeepSORT [38], devido à eficácia apresentada em estudos anteriores [48], [49]. Quando comparado com o ByteTrack, na maioria dos casos este supera o desempenho apresentado pelo DeepSORT, sendo essa a razão de escolha do ByteTrack como algoritmo de *tracking* a utilizar no desenvolvimento da dissertação.

3.4.3 Estrutura de Dados e Homografia

Após cada deteção ter um identificador associado, guardamos os resultados obtidos em ficheiros .CSV, indo ao encontro da estrutura dos ficheiros que nos tinham sido fornecidos inicialmente. Desta forma, cada ficheiro tem informação correspondente a um vídeo, onde são guardadas as coordenadas de cada deteção, o ID de cliente resultante da aplicação do algoritmo de *tracking*, o *frame* em que foi feito o registo, o nome do vídeo, a câmara e as coordenadas dos pontos de trajetória respetivas a um plano 2D visto de cima (planta).

Para calcular os pontos de trajetória é feita uma transformação de coordenadas utilizando a homografia de cada cena. Em primeiro lugar, com base nas coordenadas das *bounding boxes* geradas pelo detetor de objetos, calculamos o ponto central da aresta inferior das mesmas (localização dos pés dos clientes). No entanto, devido à maneira como as pessoas andam, as *bounding boxes* acabam por oscilar de *frame* para *frame* (devido ao movimento dos braços e das pernas), o que leva as trajetórias a apresentarem algumas irregularidades. De forma a reduzir este impacto, aplicamos um método de *smoothing*, que consiste na média deslizante dos 3 pontos mais recentes, sendo que se o número de pontos for inferior a 3, apenas esses pontos são considerados. Por fim, recorrendo à informação da homografia de cada cena, são calculadas as coordenadas do ponto correspondente à localização do cliente na planta da loja (vista de cima), através da multiplicação da matriz de projeção fornecida com o ponto central da aresta inferior da *bounding box* obtida anteriormente.

3.5. Processo de Reidentificação



Fig. 3.10 – Esquema do processo de Reidentificação

O processo de reidentificação foi implementado com o objetivo de associar o mesmo identificador a todos os pontos de trajetória que pertençam à mesma pessoa. A mudança de identificação de um cliente pode acontecer devido a oclusões, condições de luminosidade ou questões como a natureza de um ambiente composto por várias câmaras. A dissertação tem como objetivo o uso exclusivo de dados de trajetória para reidentificar os clientes, pelo que por este motivo não foram consideradas abordagens relacionadas com análise e comparação de imagens. Como conseguimos ver na Figura 3.10, o processo de reidentificação é composto por duas partes, a primeira que visa juntar os pontos que se encontrem muito próximos uns dos outros e a segunda que pretende corresponder o último ponto da trajetória de um cliente com o primeiro ponto da trajetória de um cliente identificado com um número diferente.

Desta forma, o processo de reidentificação começa por juntar as identificações de clientes cujos pontos se encontram próximos uns dos outros, mas têm uma identificação diferente devido a questões que o *tracker* não conseguiu superar ou devido à natureza multicâmara do ambiente. Estes pontos são normalmente extraídos da mesma câmara ou de duas câmaras cujo campo de visão abrange uma área comum, mas vista de perspetivas diferentes. Para tal, fazemos uma iteração ponto a ponto por ordem cronológica por todos os pontos do conjunto de dados e, para cada um destes pontos é inferida uma lista de pontos que estejam dentro de um intervalo de tempo (Δt_1), que sejam próximos do ponto a ser iterado e que tenham um ID de cliente diferente. A distância máxima (Δs_1), que os pontos de diferentes clientes podem estar para serem considerados a mesma pessoa, tem de ser um valor inferior à distância média entre os pontos da trajetória de duas pessoas que caminham lado a lado e o intervalo de tempo considerado deve ser pequeno.

Posto isto, iteramos pela lista de pontos criada e o ID de cliente do ponto que for mais próximo do ponto da primeira iteração é o ID de cliente que irá ser reidentificado. Este cliente já não pode ser considerado para reidentificação até que a primeira parte do processo de reidentificação esteja concluída. A Figura 3.11 apresenta um fluxograma da primeira parte do processo de reidentificação, descrita anteriormente.

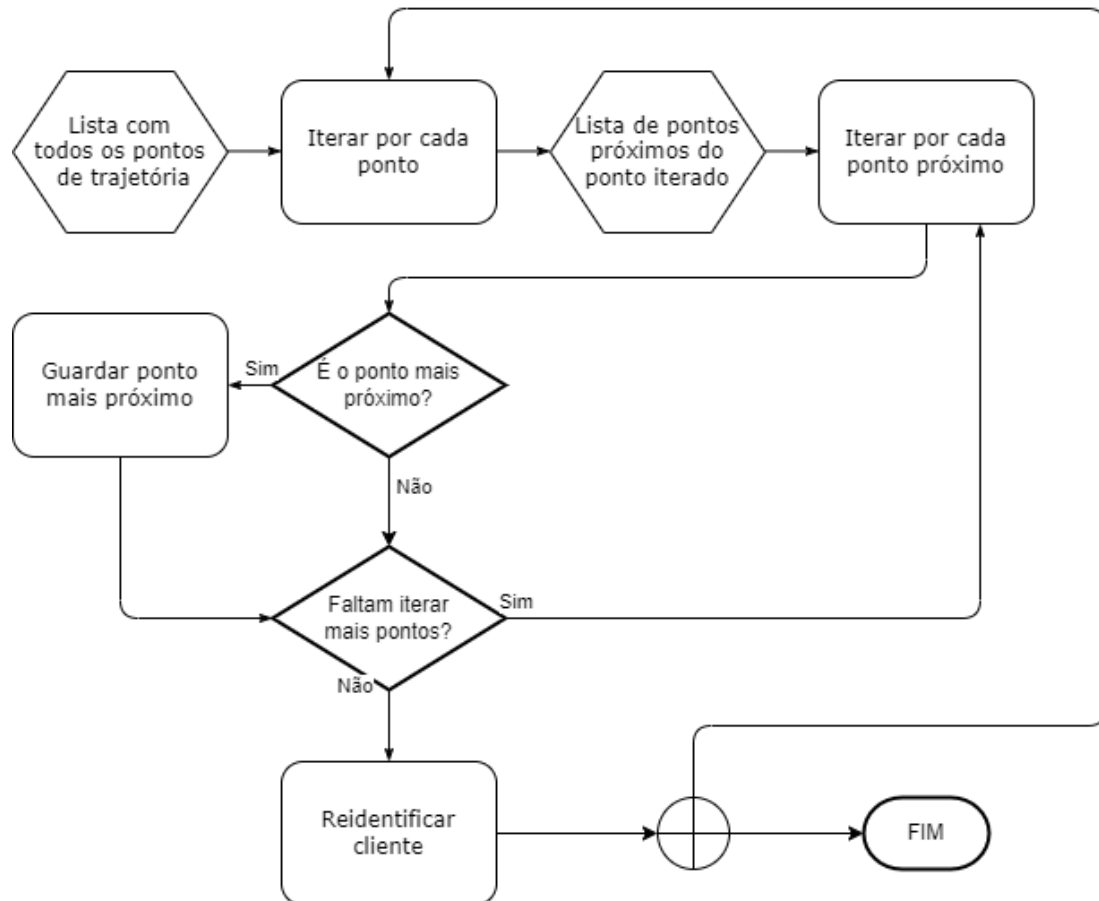


Fig. 3.11 – Fluxograma da primeira parte do processo de reidentificação

Após a primeira iteração, na segunda parte do processo é feita uma distinção entre dois tipos de áreas, as Áreas de Saída/Entrada e a Área de Reidentificação, correspondendo respetivamente às zonas azul e branca ilustradas na Figura 3.3. A Área de Saída/Entrada representa uma área onde os pontos não são considerados para reidentificação, uma vez que é provável que os clientes estejam a sair ou a entrar no local, enquanto a Área de Reidentificação representa a área onde os pontos podem ser reidentificados.

Nesta parte, começamos por fazer a criação de dois grupos. Um grupo com todas as entradas correspondentes às primeiras posições dos clientes e o outro com as entradas respetivas às últimas posições dos clientes, ambos ordenados pelo tempo. A localização dos pontos tem de estar presente na Área de Reidentificação, caso contrário não são considerados para reidentificação. Posto isto,

iteramos sobre o grupo composto pelas últimas posições dos clientes e, para cada posição, tentamos encontrar no grupo das primeiras posições, uma que esteja dentro de um intervalo de tempo (Δt_2) definido e não ultrapasse uma dada distância máxima (Δs_2). O intervalo de tempo (Δt_2) é calculado desde o momento em que o ponto a ser iterado foi registrado até ao momento em que foi registrado o ponto que estamos a considerar.

Em relação à distância máxima (Δs_2), consideramos duas abordagens, ambas descritas no Capítulo 4. A primeira consiste numa distância máxima fixa e definida previamente e a segunda varia de acordo com o tempo decorrido entre os pontos a ser considerados, onde cada segundo corresponde a um incremento de 25 unidades (velocidade média de caminhada) na distância máxima. Nas experiências em que o valor da distância máxima é fixo, a sua escolha tem de ter em conta que se o valor for demasiado pequeno, quando os pontos de trajetória se encontram demasiado afastados (por exemplo devido a oclusões), o algoritmo não conseguirá considerar pontos que deveriam ser considerados. Por outro lado, se escolhermos uma distância máxima demasiado grande, pode levar à reidentificação de IDs que não deveriam de ser reidentificados.

Se mais do que um ponto resultar dos critérios referidos anteriormente, escolhemos o cliente a ser reidentificado com base na distância mais pequena até ao ponto a ser iterado. Para calcular a distância de um ponto a outro, utilizámos a fórmula da distância euclidiana entre dois pontos. De seguida, pegamos em todos os pontos que pertencem ao cliente a ser reidentificado, atribuímos um novo ID que é igual ao ID de cliente do ponto iterado e marcamos o cliente reidentificado para que não possa ser considerado para reidentificação novamente. Se um dos pontos iterados não tiver um ponto que cumpra os critérios necessários, após a conclusão de todo o processo de iteração, assumimos que o cliente iterado não pode ser reidentificado e o seu ID de cliente permanece o mesmo. A Figura 3.12 apresenta de forma resumida um diagrama de fluxo da segunda parte do processo de reidentificação.

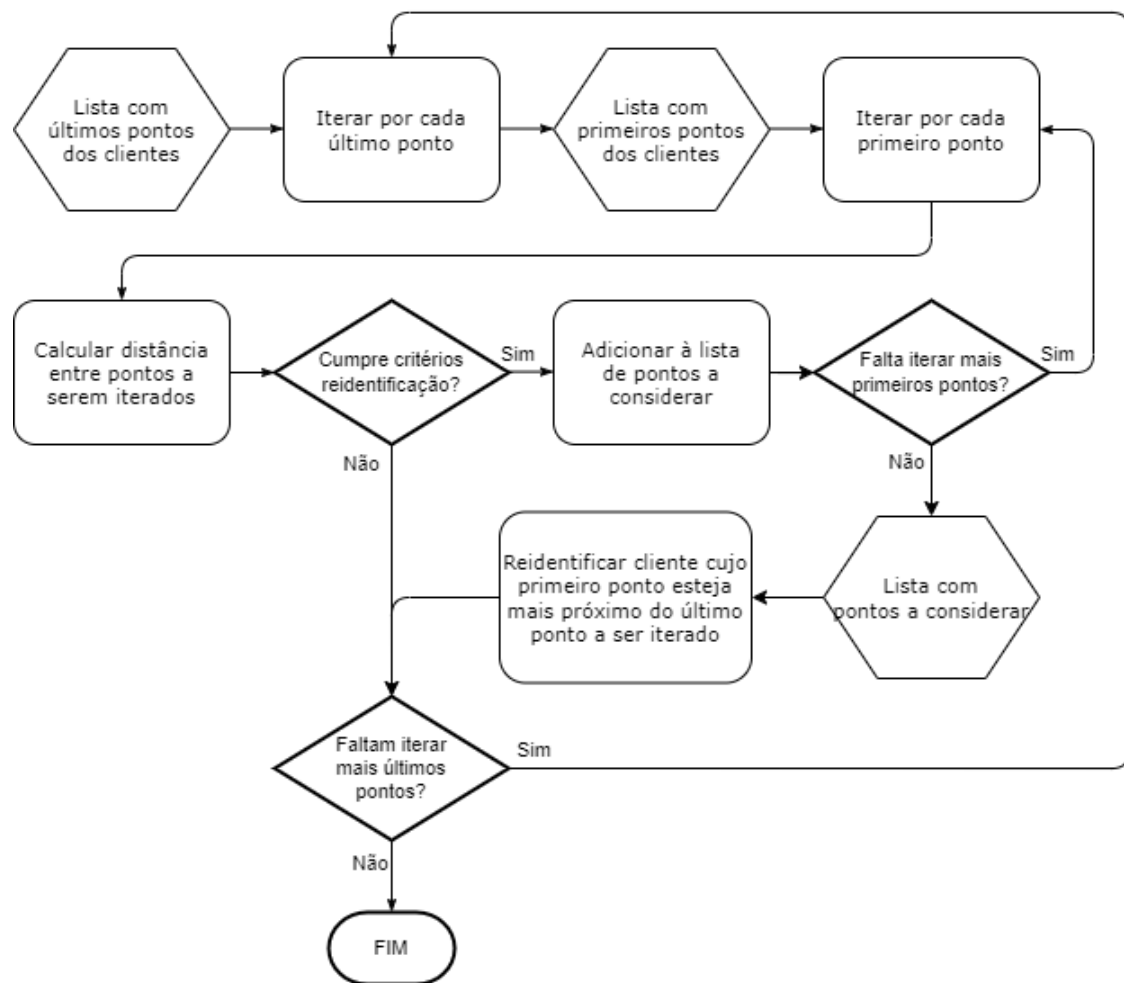


Fig. 3.12 – Fluxograma da segunda parte do processo de reidentificação

Cada ID de cliente está agora associado a um novo ID, chamado “Customer Re-ID”. No entanto, como podemos ver no exemplo da Tabela 3.3, as primeiras quatro linhas deveriam corresponder todas ao mesmo cliente, isto é, ao cliente número 1. Neste contexto, para garantir que um cliente é representado apenas por um único ID, agrupamos os pares (ID do cliente, Re-ID do cliente) que tenham um elemento em comum numa lista. Por fim, fazemos a alteração do Re-ID do cliente para o valor mais baixo da lista. Usando a Tabela 3.3 como exemplo, podemos ver que após aplicar este processo, o Re-ID do cliente nas primeiras quatro linhas será 1 e a quinta linha ficará igual.

Tab. 3.3 – Comparações entre os Customer Re-IDs antes e depois da alteração

Customer ID	Customer Re-ID (pré alteração)	Customer Re-ID (pós alteração)
1	1	1
2	1	1
3	2	1
4	3	1
5	5	5

De forma resumida, o processo de reidentificação procura reidentificar as pessoas presentes no campo de visão de várias câmaras com um identificador único. Para tal, é tomada uma abordagem que tem como objetivo reidentificar clientes com pontos muito próximos uns dos outros, devido a questões como a sobreposição dos campos de visão de duas ou mais câmaras, que originam pontos duplicados correspondentes à mesma pessoa, mas com identificadores diferentes. De seguida, a segunda parte do processo de reidentificação consiste na reidentificação de pedaços de trajetória que um cliente desempenhou, através da procura da melhor solução para unir estes pedaços, associando a cada último ponto de um cliente um primeiro ponto de um cliente identificado com um número diferente.

CAPÍTULO 4

Experiências e Resultados

Neste capítulo encontram-se descritas as experiências realizadas e os resultados obtidos ao longo deste processo. As primeiras secções são compostas por uma descrição da implementação de cada experiência, seguida dos parâmetros utilizados e dos resultados obtidos. Na última secção é feita uma comparação das abordagens adotadas e dos resultados alcançados em cada experiência.

4.1. Experiência 1 – Reidentificação com primeiros e últimos pontos

Esta experiência consiste na primeira tentativa de reidentificação, sendo o ponto de partida para a criação do processo de reidentificação adotado. A implementação desta experiência consiste na recolha de todos os primeiros e últimos pontos de cada cliente e procura fazer uma associação de IDs, fazendo corresponder cada último ponto de um cliente com o primeiro ponto de outro cliente. Para um cliente ser elegível para ser reidentificado necessita de cumprir dois critérios, sendo eles, o seu primeiro ponto ter sido registado dentro de um intervalo de tempo (Δt_2) e dentro de uma distância máxima (Δs_2) pré-definidos, calculados a partir do último ponto do cliente a ser iterado. Se mais do que um cliente cumprir os requisitos necessários, isto é, esteja dentro do tempo e da distância máxima pré-definidos, é escolhido o cliente cuja distância entre pontos seja menor. Esta abordagem de reidentificação consiste na primeira fase do processo de reidentificação final, e encontra-se descrita com maior detalhe na Secção 3.5.

A distância máxima (Δs_2) considerada entre o ponto a ser iterado, que corresponde ao último ponto da trajetória de um cliente, e o primeiro ponto de outro cliente não pode ser superior a 100 unidades, assim como o intervalo de tempo (Δt_2) entre os mesmos não pode ser superior a 15 segundos, para o cliente ser considerado para reidentificação. Estes valores foram determinados de forma experimental e podem variar em ambientes com condições e configurações diferentes.

Uma vez que o conjunto de dados fornecido não contém anotações respetivas à posição real dos clientes, optámos por analisar manualmente um conjunto de 25 exemplos registando o número total de pessoas presente em cada exemplo, de forma a conseguir avaliar a abordagem atual.

Os resultados (Anexo A) indicam que o número de identificadores obtidos após aplicar o algoritmo de *tracking* é de 512. Após o processo de reidentificação, o número total de identificadores foi reduzido para 144, o que representa menos de metade do número de identificadores originais. No entanto, continua a ser perto do dobro do número de clientes efetivamente presentes na área considerada da loja ($N = 76$).

Com base na análise dos valores resultantes desta experiência, e nos mapas anotados com as trajetórias dos clientes após ter sido aplicado o processo de reidentificação, foram examinados com mais detalhe os casos de erros na reidentificação.

Um dos casos de erro mais frequentes acontece quando existem áreas da loja abrangidas pelo campo de visão de duas ou mais câmaras. Desta forma, uma pessoa ao passar numa área da loja abrangida por várias câmaras, será detetada tantas vezes quanto o número de câmaras que abrangem essa zona. Isto resulta na extração de pontos de trajetória extra, com um identificador diferente, para o mesmo cliente, introduzindo escolhas de pontos desnecessários no processo de reidentificação. A Figura 4.1 compara dois *frames* onde aparece a mesma pessoa, num determinado instante de tempo, capturada a partir de duas câmaras diferentes.



Fig. 4.1 – Pessoa capturada num determinado instante de tempo, a partir de duas câmaras diferentes

O segundo problema está relacionado com as Áreas de Saída/Entrada e acontece quando o primeiro e o último ponto de um cliente se encontram numa destas áreas. Quando isto acontece, o processo de reidentificação não pode ser aplicado uma vez que os critérios anteriores são motivos de exclusão dos pontos para reidentificação. Este problema acontece quando uma pessoa entra no campo de visão da câmara e o *tracker* não é capaz de fazer a associação da pessoa entre o primeiro *frame* (onde só se encontra visível a cabeça) e os *frames* seguintes, acabando por gerar um identificador diferente.

As Figuras 4.2 e 4.3 representam um exemplo para as situações descritas anteriormente. Na Secção 4.2 introduzimos uma nova fase no algoritmo de reidentificação, que permite lidar com este tipo de problemas.



Fig. 4.2 – Exemplo de uma pessoa detetada por câmaras diferentes originando IDs extra

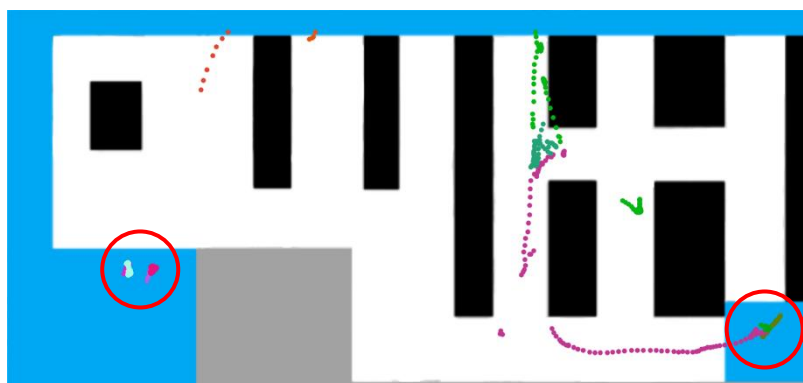


Fig. 4.3 – Exemplo com pontos em áreas de Saída/Entrada, impossíveis de reidentificar

4.2. Experiência 2 – Reidentificação com distância fixa

Com base nos resultados obtidos na experiência anterior e tendo em conta as limitações identificadas no processo de reidentificação, decidimos reutilizar o algoritmo criado e acrescentar uma fase inicial para resolver os problemas identificados.

A implementação deste algoritmo encontra-se dividida em duas partes, uma primeira parte que visa reidentificar clientes cujos pontos se encontrem muito próximos uns dos outros, e uma segunda fase que consiste na adoção do algoritmo implementado na experiência anterior. Uma explicação mais detalhada do algoritmo de reidentificação utilizado, assim como outros detalhes de implementação encontram-se expostos na Secção 3.5.

Face à solução anterior, esta abordagem conseguiu melhorar os problemas identificados. A reidentificação de clientes cujas trajetórias são compostas por pontos muito próximos de pontos de outro cliente permite resolver a questão da perda de identificação dentro de zonas de Saída/Entrada, uma vez que na primeira fase do algoritmo as mesmas não são tidas em conta e os pontos se encontram muito próximos, e permite associar o mesmo identificador a trajetórias efetuadas pela mesma pessoa, mas capturadas a partir de câmaras diferentes.

A primeira fase deste novo processo de reidentificação também tem como parâmetros a distância máxima (Δs_1) e o intervalo de tempo (Δt_1) considerados, ainda que definidos com valores diferentes. De forma a não reidentificar clientes que se encontrem próximos uns dos outros, por caminharem lado a lado, mas conseguir na mesma unir as trajetórias que correspondem à mesma pessoa, a distância máxima (Δs_1) utilizada foi definida de forma a ser inferior ao valor da distância média a que os pontos de duas pessoas que caminham lado a lado apresenta. Neste contexto o valor utilizado são 10 unidades. Em relação ao intervalo de tempo (Δt_1), como apenas pretendemos considerar pontos muito próximos, utilizamos um intervalo de tempo de 5 segundos.

A segunda fase do algoritmo mantém os valores da distância máxima (Δs_2) e do intervalo de tempo (Δt_2) considerados na experiência da Secção 4.1, ou seja, a distância máxima entre o ponto a ser iterado e o primeiro ponto do cliente a ser considerado para reidentificação não pode ser superior a 100 unidades e o intervalo de tempo entre os respetivos pontos não pode ser superior a 15 segundos.

Uma vez que os exemplos utilizados consistem nos mesmos exemplos da experiência anterior, o número total de pessoas presentes nos exemplos bem como o número total de IDs de cliente não sofrem alterações, ou seja, apresentam um total de 76 e 512 identificadores, respetivamente. Após aplicar o processo de reidentificação, agora constituído pelas duas fases referidas anteriormente, os resultados aproximam-se mais do número de pessoas presentes nos exemplos, contabilizando um total de 109 identificadores.

De forma a perceber melhor os resultados obtidos, são feitas comparações entre uma versão com as trajetórias coloridas de acordo com os IDs fornecidos pelo *tracker* e uma versão com as trajetórias coloridas com base nos IDs gerados após ter sido aplicado o processo de reidentificação. As Figuras 4.4 e 4.5 ilustram um exemplo destas comparações.

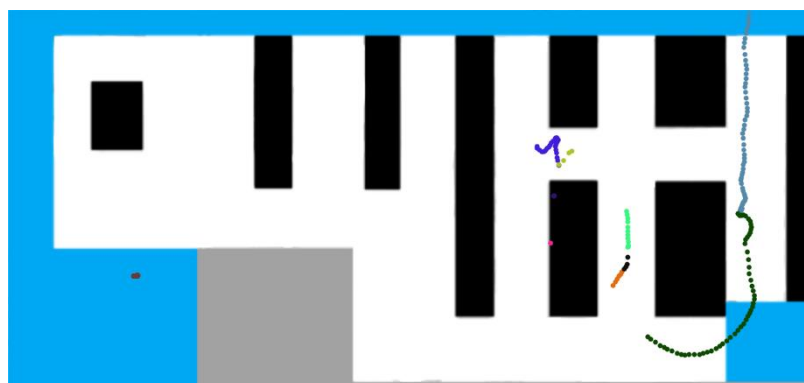


Fig. 4.4 – Exemplo colorido com base nos IDs resultantes do *tracker*



Fig. 4.5 – Exemplo colorido com base nos IDs resultantes do processo de reidentificação

Através da comparação das Figuras 4.4 e 4.5, conseguimos perceber que existem duas pessoas presentes nas gravações, sendo que uma se encontra parada (pontos no canto inferior esquerdo das imagens) e a outra em movimento. A trajetória da pessoa que se encontra em movimento foi identificada pelo *tracker* com 11 IDs diferentes, sendo o nosso algoritmo capaz de unir estes fragmentos de trajetória e associá-los a um único identificador de cliente. Existe ainda um conjunto de pontos coloridos a cor-de-rosa na Figura 4.4, que resultam de um erro de detecção por parte do detetor de objetos e que são removidos através da aplicação de operações de pré processamento.

Considerando os exemplos apresentados nas Figuras 4.2 e 4.3 da subsecção 4.1, onde estavam presentes aspetos que não conseguiam ser corretamente reidentificados com a abordagem anterior, percebemos que a implementação desta experiência consegue superar esses obstáculos. As imagens presentes nas Figuras 4.6 e 4.7 fazem uma comparação dos exemplos apresentados nas Figuras 4.2 e 4.3, sendo que na primeira imagem é aplicado o algoritmo da experiência anterior (Experiência 1) e na segunda imagem é ilustrado o resultado colorido com identificadores da experiência desta secção (Experiência 2).

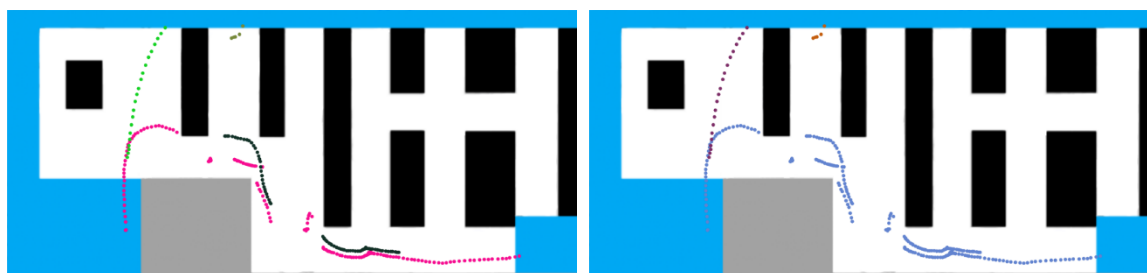


Fig. 4.6 – Exemplo com trajetória corretamente reidentificada nas áreas abrangidas por duas câmaras

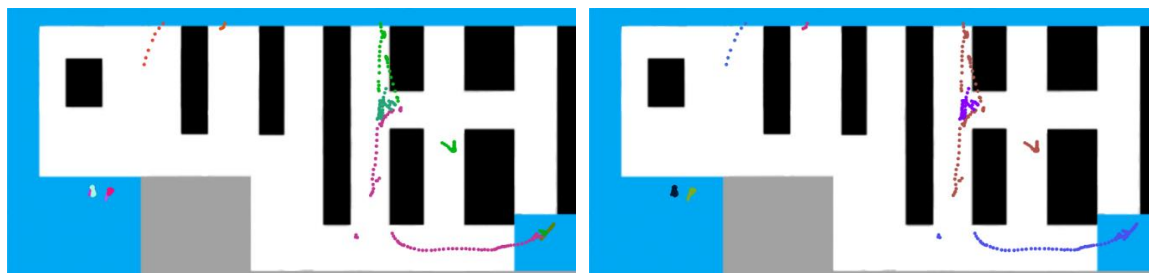


Fig. 4.7 – Exemplo com trajetória corretamente reidentificada nas áreas de Saída/Entrada

Através da comparação das imagens presentes na Figura 4.6, conseguimos perceber que o algoritmo implementado nesta secção consegue resolver o problema apresentado no algoritmo da experiência anterior, relativo a um cliente captado por duas câmaras não conseguir ser reidentificado com um único identificador. Nas imagens da Figura 4.7, comparando novamente os resultados das experiências 1 e 2, é possível observar que os pontos presentes nas áreas de Saída/Entrada que não conseguiam ser reidentificados, devido à totalidade dos pontos do cliente estarem inseridos nessas áreas, deixa de ser um problema, encontrando-se agora todos os pontos identificados com o mesmo ID.

Apesar das melhorias introduzidas neste algoritmo face ao apresentado na subsecção 4.1, as oclusões continuam a representar uma dificuldade para a qualidade de reidentificação de trajetórias, na medida que nos sítios onde os clientes não conseguem ser detetados devido a estarem ocluídos por um objeto (geralmente prateleiras), não é possível extrair a sua trajetória. Isto resulta em falhas na trajetória que causam dificuldades no algoritmo em associar e reidentificar corretamente os clientes.

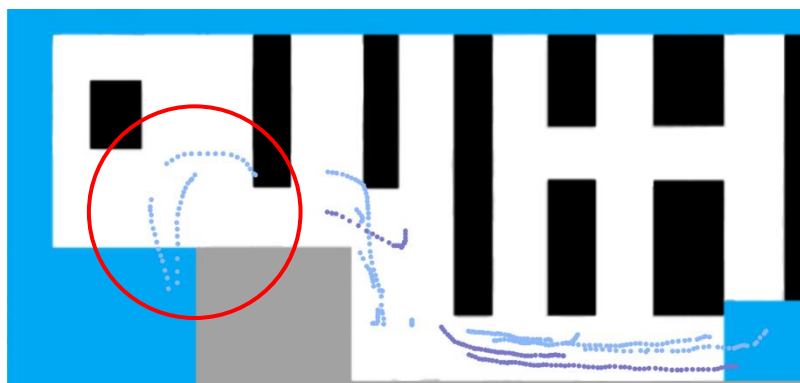


Fig. 4.8 – Exemplo 1, falha na trajetória impossibilitando reidentificar corretamente

Na Figura 4.8, os três conjuntos de pontos de trajetória realçados deveriam ser reidentificados como dois clientes, mas como não existe informação acerca da trajetória de um dos clientes quando o mesmo se encontra ocluído por detrás de uma prateleira, o algoritmo não consegue associar as

trajetórias a dois clientes diferentes, apresentando assim uma limitação no processo de reidentificação.

4.3. Experiência 3 – Reidentificação com distância dinâmica

Com o objetivo de tentar melhorar os resultados obtidos na implementação anterior, nesta experiência decidimos aplicar o algoritmo desenvolvido para a experiência anterior, mas considerando agora o parâmetro de distância máxima (Δs_2) de forma dinâmica, ou seja, fazendo o valor da distância máxima variar consoante o tempo entre os pontos considerados. Este parâmetro apenas é aplicado à segunda parte do processo de reidentificação, mantendo a primeira parte igual à descrita na experiência da subseção 4.2.

A distância dinâmica é calculada com base num valor estimado da velocidade média de caminhada de uma pessoa multiplicado pelo intervalo de tempo entre os pontos a serem considerados, sendo eles o último ponto do cliente a ser iterado e os primeiros pontos dos clientes elegíveis para reidentificação. Desta forma, para cada cliente elegível para reidentificação é calculada uma distância máxima a que o mesmo tem de estar para continuar a ser considerado. Após o cálculo desta distância, à semelhança do que acontecia na experiência anterior, são excluídos os clientes cuja distância seja superior à distância máxima calculada e é selecionado para reidentificação o cliente cujo primeiro ponto esteja mais próximo do último ponto do cliente a ser iterado.

Analisando o número de identificadores respetivo a todos os exemplos (Anexo C), obtido após ter sido aplicado o processo de reidentificação, este contabiliza um total de 105 IDs. O valor apresenta uma diferença de menos 4 identificadores face à abordagem anterior.

As alterações no algoritmo descrito nesta experiência foram introduzidas com o propósito de tentar resolver situações que a versão do algoritmo da experiência anterior tinha algumas dificuldades em superar, nomeadamente questões relacionadas com a dificuldade de reidentificação devido à incapacidade de deteção causada por oclusões, quando apenas existe informação de um cliente, mas na realidade estão presentes dois ou mais clientes.

Comparando o exemplo ilustrado na Figura 4.8 da experiência anterior com o mesmo exemplo aplicando o algoritmo da experiência atual, que utiliza uma distância máxima dinâmica, conseguimos perceber que a introdução desta distância, como critério de exclusão dos clientes considerados para reidentificação, não acrescenta melhorias na qualidade das trajetórias. Pelo contrário, neste exemplo a qualidade da reidentificação piora face ao algoritmo anterior, sendo um pedaço da trajetória de um dos clientes presentes incorretamente identificado, o que dá origem a um identificador extra. A Figura 4.9 apresenta as trajetórias coloridas de acordo com os identificadores gerados por este algoritmo.

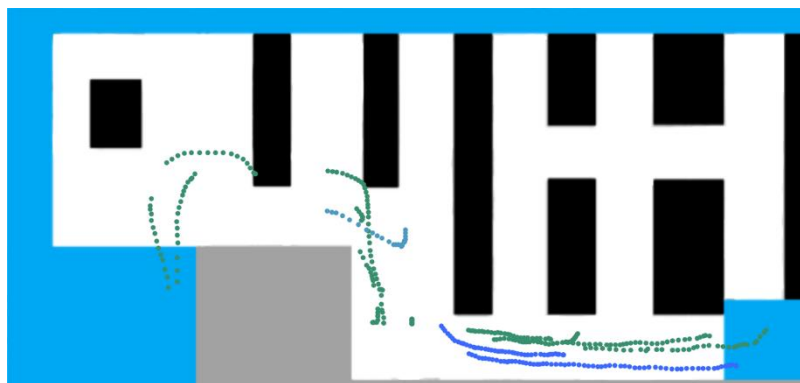


Fig. 4.9 – Exemplo 1 reidentificado com algoritmo da Experiência 3

4.4. Comparação dos Resultados

Com base na informação apresentada nas secções anteriores do Capítulo 4, ao comparar os resultados das experiências efetuadas conseguimos perceber que existe uma grande melhoria da Experiência 1 para as Experiências 2 e 3, passando de um número total de 144 identificadores para 109 e 105 identificadores, respetivamente. Relativamente às duas últimas experiências, ambas obtiveram bons resultados, sendo o número total de identificadores após o processo de reidentificação mais baixo na Experiência 3.

De uma forma geral, o algoritmo desenvolvido na Experiência 1 apesar de conseguir reidentificar alguns clientes não é uma opção suficientemente fiável. No entanto, os algoritmos das experiências 2 e 3 conseguem aproximar o número de identificadores a um estado considerado aceite, sendo em vários exemplos o resultado de reidentificação ideal. Considerando uma boa aproximação quando o número de identificadores após o processo de reidentificação é igual ou, tem apenas uma unidade de diferença do número real de pessoas presentes na loja (uma vez que a incapacidade de reidentificação de apenas um ID não invalida a qualidade das trajetórias reidentificadas), o algoritmo da Experiência 2 consegue obter uma boa aproximação da trajetória desempenhada pelos clientes em 68% dos exemplos, enquanto o algoritmo da Experiência 3 obtém resultados com bom grau de aproximação em 72% dos exemplos. No entanto, existem situações em que o número de Re-IDs é menor, mas que a qualidade das trajetórias obtidas não espelha o que é apresentado nos números.

De acordo com os resultados numéricos obtidos nas Experiências 2 e 3 (Anexos B e C), o número de identificadores é mais próximo do número real de clientes na Experiência 3, face à Experiência 2, em 8 exemplos. No entanto, para dois destes exemplos, apesar de o número de identificadores ser mais próximo do número real de clientes, o algoritmo da Experiência 2 apresenta uma qualidade superior porque os resultados do algoritmo da Experiência 3 contém pedaços de trajetória pertencentes a clientes diferentes, mas que foram reidentificados como um único cliente.

Para 9 exemplos, o número de identificadores é mais próximo do número real de clientes na Experiência 2, quando comparado com os resultados da Experiência 3. Destes exemplos, apenas num deles a qualidade não se traduz pelo número de identificadores, ou seja, apenas num deles a qualidade das trajetórias é melhor se forem geradas pelo algoritmo desenvolvido na Experiência 3.

Os restantes exemplos apresentam o mesmo número de identificadores em ambas as experiências. Após analisar manualmente cada situação percebemos que as situações de ambas as experiências são idênticas em todos os exemplos, à exceção do exemplo 3 no qual a Experiência 2 apresenta uma reidentificação mais próxima da realidade.

A Tabela 4.1 sintetiza a informação descrita.

Tab. 4.1 – Qualidade das reidentificações - Experiência 2 vs Experiência 3

Qualidade	Exemplos
Igual	2, 4, 7, 8, 17, 18, 19, 23, 25
Maior na Experiência 2	1, 3, 6, 12, 13, 14, 15, 16, 20, 21, 24
Maior na Experiência 3	5, 9, 10, 11, 22

Analisando os resultados apresentados, o algoritmo considerado na Experiência 1 obteve resultados iguais ou inferiores em todos os exemplos, quando comparado com os algoritmos das Experiências 2 e 3. Em termos da qualidade das reidentificações geradas, o algoritmo correspondente à Experiência 2 apresenta resultados superiores ao algoritmo utilizado na Experiência 3 em 44% dos exemplos. Por outro lado, o algoritmo da Experiência 3 é melhor em 20% dos exemplos. Os restantes 36% correspondem à percentagem de exemplos em que ambos os algoritmos apresentam reidentificações com qualidades semelhantes.

Com base nas análises efetuadas, a primeira com base no valor do número de identificadores após ser aplicado o processo de reidentificação e a segunda direcionada à qualidade apresentada nas trajetórias após aplicado o algoritmo de reidentificação, conseguimos perceber que em alguns exemplos, apesar do número de identificadores gerados pela Experiência 2 não ser tão próximo do número real de identificadores, a qualidade da trajetória obtida é melhor do que a qualidade das trajetórias geradas pela Experiência 3. Apesar de a diferença de performance entre ambas as abordagens ser pequena, conseguimos concluir que o algoritmo da Experiência 2 obteve melhores resultados nos exemplos testados.

Posto isto, decidimos analisar a taxa de acerto das reidentificações efetuadas pelos algoritmos das Experiências 2 e 3, uma vez que foram os algoritmos que geraram reidentificações com melhor qualidade. A forma como foi calculada a taxa de acerto das reidentificações é apresentada na Fórmula 4.

$$Taxa\ de\ Acerto = 1 - \frac{N^{\circ}\ Reidentifica\c{c}o\ Erradas}{N^{\circ}\ Total\ Reidentifica\c{c}o\ es} \quad (4)$$

No número de reidentificações erradas contam reidentificações por excesso ou por defeito, ou seja, reidentificações que não eram necessárias por juntarem dois clientes diferentes ou reidentificações que eram necessárias, mas não foram efetuadas.

Na Experiência 2, o número total de reidentificações de todos os exemplos soma um total de 225 reidentificações e o número total de reidentificações erradas foi 40 reidentificações. Desta forma, a taxa de acerto é igual a 0.82(2). Ou seja, o algoritmo de reidentificação desenvolvido na Experiência 2 conseguiu reidentificar aproximadamente 82% dos clientes identificados nos exemplos considerados. Na Experiência 3 o número total de reidentificações dos exemplos contabilizou um total de 236 reidentificações e um número total de reidentificações erradas igual a 49. Desta forma, a taxa de acerto do algoritmo desta experiência é aproximadamente igual a 0.792. Por outras palavras, o algoritmo da Experiência 3 conseguiu reidentificar aproximadamente 79% dos clientes. É importante referir que, em ambos os resultados, não é contabilizada a redução de identificadores resultante do pré-processamento aplicado ao conjunto de dados, ou seja, o valor apresentado é apenas referente ao processo de reidentificação.

A taxa de acerto comprova a comparação entre os resultados de performance analisados anteriormente, demonstrando que o algoritmo desenvolvido na Experiência 2, face ao algoritmo apresentado na Experiência 3, obteve melhores resultados na reidentificação dos clientes presentes nos exemplos considerados.

No que diz respeito às limitações, considerando apenas os algoritmos das experiências 2 e 3, as oclusões causadas por objetos como prateleiras ou balcões são um dos principais obstáculos para a ferramenta de reidentificação criada, uma vez que a mesma trabalha com distâncias e intervalos de tempo entre pontos de trajetória, cujo funcionamento é afetado quando existem grandes perdas no que seria o percurso original desempenhado pelos clientes.

Outra das limitações prende-se com o facto de existir algum fator que impossibilite a deteção contínua das pessoas presentes em zonas de Saída/Entrada, como por exemplo a oclusão ou incapacidade de deteção devido à distância a que a pessoa se encontra da câmara. Isto faz com que as deteções das pessoas sejam registadas com um intervalo temporal acima do considerado para reidentificação. Uma vez que as mesmas se encontram numa zona de Saída/Entrada, o intervalo de tempo considerado é reduzido porque apenas podem ser reidentificadas pela primeira parte do algoritmo, impossibilitando desta forma a sua reidentificação.

CAPÍTULO 5

Conclusão

Nesta dissertação, foi desenvolvido um sistema de detecção e *tracking* de clientes num ambiente comercial de retalho, através de vídeos capturados por câmaras de videovigilância inseridas num ambiente multi-câmara. Estes vídeos encontram-se desfocados na zona da cara dos clientes, de forma a manter a anonimização dos mesmos, o que dificulta a reidentificação de pessoas com base em imagens.

Desta forma, este estudo propôs um mecanismo de reidentificação de pessoas com base nas trajetórias percorridas pelas mesmas na loja. As trajetórias foram obtidas com base na informação proveniente do sistema de detecção e *tracking* de pessoas desenvolvido que, após detetar e associar um identificador a cada pessoa, efetua o cálculo e extração da trajetória desempenhada por cada uma delas. No capítulo 4 são apresentados e comparados 3 processos alternativos originais de fazer a reidentificação, tendo o método da Experiência 2 apresentado os melhores resultados.

Este método efetuou um total de 225 reidentificações, com uma taxa de sucesso de 82%, demonstrando a eficácia do mecanismo de reidentificação proposto na associação de um identificador único para cada cliente, independentemente da câmara em que foi detetado. Esta abordagem pode ser vista como uma ajuda para melhorar a precisão dos algoritmos de *tracking*, permitindo-lhes minimizar o número de identificadores gerados, e apresenta uma alternativa para a utilização de processos de reidentificação de pessoas baseados em imagens, num ambiente comercial com várias câmaras. É também uma solução para lidar com algumas oclusões e erros causados pelos algoritmos de detecção e *tracking* de objetos, uma solução a considerar quando os dados que temos têm uma parte da pessoa ocluída devido a questões de privacidade ou quando simplesmente só temos acesso aos dados da trajetória realizada pelos clientes.

Para além do mecanismo de reidentificação proposto, neste trabalho é refeito todo o processo de aquisição de trajetórias que, na versão de dados fornecida, continha incoerências que impediam a abordagem pretendida do problema. Todo o software desenvolvido encontra-se disponível no GitHub em: <https://github.com/DiogoMendess/master-thesis>

Para aumentar o desempenho pretendemos melhorar o processo de reidentificação, testando-o com diferentes conjuntos de dados. Uma vez que o processo de detecção e *tracking* tem um grande impacto na qualidade das trajetórias inicialmente geradas, para trabalho futuro seria importante experimentar a utilização de outros tipos de algoritmos ou técnicas de obtenção de dados de trajetórias para obter as trajetórias desempenhadas pelas pessoas na loja.

No projeto em que a dissertação se encontra inserida, foi paralelamente desenvolvido um mecanismo que visa reduzir o impacto negativo que as oclusões têm na extração de trajetórias [50]. A sua utilização pode fazer a diferença na melhoria da qualidade das trajetórias e consequentemente facilitar o processo de reidentificação em casos onde existem muitas oclusões, devendo ser considerado numa abordagem futura. É possível que num ambiente multicâmara diferente, no qual o campo de visão de cada câmara não tem uma complexidade de oclusão tão elevada, o mecanismo de reidentificação possa obter melhores resultados. Para além disso, a inexistência de anotações no conjunto de dados fornecido introduziu uma dificuldade acrescida na avaliação do método proposto, pelo que seria interessante avaliar o algoritmo desenvolvido com um conjunto de dados anotado para o efeito.

Finalmente, seria interessante associar uma pontuação de confiança às reidentificações efetuadas, de forma que outros mecanismos possam tirar benefícios mediante a sua aplicação a outros processos de reidentificação.

Referências

- [1] F. Rahmanov, M. Mursalov, and A. Rosokhata, "Consumer behavior in digital era: impact of COVID 19," *Marketing and Management of Innovations*, vol. 5, no. 2, pp. 243–251, 2021, doi: 10.21272/mmi.2021.2-20.
- [2] N. M. Larsen, V. Sigurdsson, and J. Breivik, "The Use of Observational Technology to Study In-Store Behavior: Consumer Choice, Video Surveillance, and Retail Analytics," *Behavior Analyst*, vol. 40, no. 2, pp. 343–371, Nov. 2017, doi: 10.1007/s40614-017-0121-x.
- [3] Z. Feng and Y. Zhu, "A Survey on Trajectory Data Mining: Techniques and Applications," *IEEE Access*, vol. 4. Institute of Electrical and Electronics Engineers Inc., pp. 2056–2067, Apr. 13, 2016. doi: 10.1109/ACCESS.2016.2553681.
- [4] R. Wirth and J. Hipp, "CRISP-DM: Towards a Standard Process Model for Data Mining."
- [5] W. Zhao, Y. Zuo, L. Zhao, and J. Jiang, "Application of LSTM Models to Predict In-store Trajectory of Customers," in *IEEE International Conference on Data Mining Workshops, ICDMW*, IEEE Computer Society, 2021, pp. 288–294. doi: 10.1109/ICDMW53433.2021.00043.
- [6] J. S. Larson, E. T. Bradlow, and P. S. Fader, "An exploratory look at supermarket shopping paths," *International Journal of Research in Marketing*, vol. 22, no. 4, pp. 395–414, Dec. 2005, doi: 10.1016/j.ijresmar.2005.09.005.
- [7] M. Kholod, T. Nakahara, H. Azuma, and K. Yada, "The influence of shopping path length on purchase behavior in grocery store," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2010, pp. 273–280. doi: 10.1007/978-3-642-15393-8_31.
- [8] B. Wu and K. Yada, "The Effect of Crowding on Visit Ratio at an Product Area: Based on RFID Data in a Japanese Supermarket."
- [9] Y. Kaneko, S. Miyazaki, and K. Yada, "The Influence of Customer Movement between Sales Areas on Sales Amount: A Dynamic Bayesian Model of the In-store Customer Movement and Sales Relationship," in *Procedia Computer Science*, Elsevier B.V., 2017, pp. 1845–1854. doi: 10.1016/j.procs.2017.08.225.
- [10] L. Zhao, Y. Zuo, and K. Yada, "Sequential classification of customer behavior based on sequence-to-sequence learning with gated-attention neural networks," *Adv Data Anal Classif*, 2022, doi: 10.1007/s11634-022-00517-3.
- [11] Y. Zuo and K. Yada, "Using Bayesian network for purchase behavior prediction from RFID data," in *Conference Proceedings - IEEE International Conference on Systems, Man and Cybernetics*, Institute of Electrical and Electronics Engineers Inc., 2014, pp. 2262–2267. doi: 10.1109/smc.2014.6974262.

- [12] N. Sano, R. Tsutsui, K. Yada, and T. Suzuki, "Clustering of Customer Shopping Paths in Japanese Grocery Stores," in *Procedia Computer Science*, Elsevier B.V., 2016, pp. 1314–1322. doi: 10.1016/j.procs.2016.08.176.
- [13] Z. Li, K. Ishibashi, K. Takai, and K. Yada, "Shop area visit ratio, stay time, and sales outcomes: In-depth analysis based on RFID data," in *2015 2nd Asia-Pacific World Congress on Computer Science and Engineering, APWC on CSE 2015*, Institute of Electrical and Electronics Engineers Inc., May 2016. doi: 10.1109/APWCCSE.2015.7476231.
- [14] M. Liu, L. Zhao, F. Sun, W. Zhao, Y. Zuo, and K. Yada, "Customer Behavior Analysis and Classification Based on Process Mining," in *Conference Proceedings - IEEE International Conference on Systems, Man and Cybernetics*, Institute of Electrical and Electronics Engineers Inc., 2021, pp. 1000–1005. doi: 10.1109/SMC52423.2021.9659063.
- [15] B. Shen, Q. Zheng, X. Li, and L. Xu, "A framework for mining actionable navigation patterns from in-store RFID datasets via indoor mapping," *Sensors (Switzerland)*, vol. 15, no. 3, pp. 5344–5375, Mar. 2015, doi: 10.3390/s150305344.
- [16] W. T. So and K. Yada, "A framework of recommendation system based on in-store behavior," in *ACM International Conference Proceeding Series*, Association for Computing Machinery, Jul. 2017. doi: 10.1145/3092090.3092130.
- [17] Y. Kaneko and K. Yada, "Bayesian Hidden Markov Model for Evaluating the Influence of In-Store Stationary Time of Customers on Their Purchase Behavior," in *Proceedings - 2018 5th Asia-Pacific World Congress on Computer Science and Engineering, APWC on CSE 2018*, Institute of Electrical and Electronics Engineers Inc., Dec. 2018, pp. 142–148. doi: 10.1109/APWConCSE.2018.00031.
- [18] D. Zaim, A. Benomar, and M. Bellafkih, "Analyzing the Trajectories of Customers by Using LCSS Approach," in *4th International Conference on Intelligent Computing in Data Sciences, ICDS 2020*, Institute of Electrical and Electronics Engineers Inc., Oct. 2020. doi: 10.1109/ICDS50568.2020.9268721.
- [19] O. Dogan, J. L. Bayo-Monton, C. Fernandez-Llatas, and B. Oztaysi, "Analyzing of gender behaviors from paths using process mining: A shopping mall application," *Sensors (Switzerland)*, vol. 19, no. 3, Feb. 2019, doi: 10.3390/s19030557.
- [20] O. Dogan and B. Oztaysi, "Genders prediction from indoor customer paths by Levenshtein-based fuzzy kNN," *Expert Syst Appl*, vol. 136, pp. 42–49, Dec. 2019, doi: 10.1016/j.eswa.2019.06.029.
- [21] M. Oziom, S. Bachl, C. Wimmer, and T. Grechenig, "Grofin: Enhancing in-store grocery shopping with a context-aware smartphone app," in *ACM International Conference Proceeding Series*, Association for Computing Machinery, Nov. 2019. doi: 10.1145/3365610.3365630.
- [22] H. Hwangbo, J. Kim, Z. Lee, and S. Kim, "Store layout optimization using indoor positioning system," *Int J Distrib Sens Netw*, vol. 13, no. 2, Feb. 2017, doi: 10.1177/1550147717692585.

- [23] Y. Liu *et al.*, “Inferring gender and age of customers in shopping malls via indoor positioning data,” *Environ Plan B Urban Anal City Sci*, vol. 47, no. 9, pp. 1672–1689, Nov. 2020, doi: 10.1177/2399808319841910.
- [24] D. Cheng, G. Yue, T. Pei, and M. Wu, “Clustering indoor positioning data using e-dbscan,” *ISPRS Int J Geoinf*, vol. 10, no. 10, Oct. 2021, doi: 10.3390/ijgi10100669.
- [25] M. I. Alipio, K. M. T. Peñalosa, and J. R. C. Unida, “In-store customer traffic and path monitoring in small-scale supermarket using UWB-based localization and SSD-based detection,” *J Ambient Intell Humaniz Comput*, 2020, doi: 10.1007/s12652-020-02236-z.
- [26] N. Ferracuti, C. Norscini, E. Frontoni, P. Gabellini, M. Paolanti, and V. Placidi, “A business application of RTLS technology in Intelligent Retail Environment: Defining the shopper’s preferred path and its segmentation,” *Journal of Retailing and Consumer Services*, vol. 47, pp. 184–194, Mar. 2019, doi: 10.1016/j.jretconser.2018.11.005.
- [27] E. Vildjiounaite, S. M. Makela, S. Jarvinen, T. Keranen, and V. Kyllonen, “Predicting consumers’ locations in dynamic environments via 3d sensor-based tracking,” in *Proceedings - 2014 8th International Conference on Next Generation Mobile Applications, Services and Technologies, NGMAST 2014*, Institute of Electrical and Electronics Engineers Inc., Dec. 2014, pp. 100–105. doi: 10.1109/NGMAST.2014.11.
- [28] Y. Kamiya, Y. Gu, and S. Kamijo, “Indoor Positioning in Large Shopping Mall with Context based Map Matching; Indoor Positioning in Large Shopping Mall with Context based Map Matching,” 2019.
- [29] M. C. Popa, T. Gritti, L. J. M. Rothkrantz, C. Shan, and P. Wiggers, “Detecting customers’ buying events on a real-life database,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2011, pp. 17–25. doi: 10.1007/978-3-642-23672-3_3.
- [30] M. C. Popa, L. J. M. Rothkrantz, C. Shan, T. Gritti, and P. Wiggers, “Semantic assessment of shopping behavior using trajectories, shopping related actions, and context information,” *Pattern Recognit Lett*, vol. 34, no. 7, pp. 809–819, 2013, doi: 10.1016/j.patrec.2012.04.015.
- [31] N. M. Larsen, V. Sigurdsson, J. Breivik, and J. L. Orquin, “The heterogeneity of shoppers’ supermarket behaviors based on the use of carrying equipment,” *J Bus Res*, vol. 108, pp. 390–400, Jan. 2020, doi: 10.1016/j.jbusres.2019.12.024.
- [32] M. Popa *et al.*, “Analysis of shopping behavior based on surveillance system,” in *2010 IEEE International Conference on Systems, Man and Cybernetics*, 2010, pp. 2512–2519. doi: 10.1109/ICSMC.2010.5641928.
- [33] R. Arroyo, J. J. Yebes, L. M. Bergasa, I. G. Daza, and J. Almazán, “Expert video-surveillance system for real-time detection of suspicious behaviors in shopping malls,” *Expert Syst Appl*, vol. 42, no. 21, pp. 7991–8005, Jul. 2015, doi: 10.1016/j.eswa.2015.06.016.
- [34] J. Liu, Y. Gu, and S. Kamijo, “Customer Behavior Recognition in Retail Store from Surveillance Camera,” in *Proceedings - 2015 IEEE International Symposium on Multimedia, ISM 2015*,

- Institute of Electrical and Electronics Engineers Inc., Mar. 2016, pp. 154–159. doi: 10.1109/ISM.2015.52.
- [35] S. Ren, K. He, R. Girshick, and J. Sun, “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks,” Jun. 2015, [Online]. Available: <http://arxiv.org/abs/1506.01497>
 - [36] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You Only Look Once: Unified, Real-Time Object Detection,” Jun. 2015, [Online]. Available: <http://arxiv.org/abs/1506.02640>
 - [37] W. Liu *et al.*, “SSD: Single Shot MultiBox Detector,” Dec. 2015, doi: 10.1007/978-3-319-46448-0_2.
 - [38] N. Wojke, A. Bewley, and D. Paulus, “Simple Online and Realtime Tracking with a Deep Association Metric,” Mar. 2017, [Online]. Available: <http://arxiv.org/abs/1703.07402>
 - [39] Y. Zhang *et al.*, “ByteTrack: Multi-Object Tracking by Associating Every Detection Box,” Oct. 2021, [Online]. Available: <http://arxiv.org/abs/2110.06864>
 - [40] Y. Cai, L. Wen, L. Zhang, D. Du, and W. Wang, “Rethinking Object Detection in Retail Stores,” 2021. [Online]. Available: <https://isrc.iscas.ac.cn/gitlab/research/locount-dataset>.
 - [41] K. Nguyen, M. Le, B. Martin, I. Cil, and C. Fookes, “When AI meets store layout design: a review,” *Artif Intell Rev*, vol. 55, no. 7, pp. 5707–5729, Oct. 2022, doi: 10.1007/s10462-022-10142-3.
 - [42] D. Heimsch, Y. H. Lau, C. Mishra, S. Srigrarom, and F. Holzapfel, “Re-Identification for Multi-Target-Tracking Systems Using Multi-Camera, Homography Transformations and Trajectory Matching,” in *2022 Sensor Data Fusion: Trends, Solutions, Applications, SDF 2022*, Institute of Electrical and Electronics Engineers Inc., 2022. doi: 10.1109/SDF55338.2022.9931703.
 - [43] P. Köhl, A. Specker, A. Schumann, and J. Beyerer, “The MTA Dataset for Multi Target Multi Camera Pedestrian Tracking by Weighted Distance Aggregation.”
 - [44] IEEE Staff and IEEE Staff, *2008 IEEE Conference on Computer Vision and Pattern Recognition*.
 - [45] M. S. I. Siam and S. Biswas, “A Deep Learning Based Person Detection and Heatmap Generation Technique with a Multi-Camera System,” in *12th International Conference on Electrical and Computer Engineering, ICECE 2022*, Institute of Electrical and Electronics Engineers Inc., 2022, pp. 260–263. doi: 10.1109/ICECE57408.2022.10089044.
 - [46] R. E. Thomson and W. J. Emery, “Chapter 4 - The Spatial Analyses of Data Fields,” in *Data Analysis Methods in Physical Oceanography (Third Edition)*, R. E. Thomson and W. J. Emery, Eds., Boston: Elsevier, 2014, pp. 313–424. doi: <https://doi.org/10.1016/B978-0-12-387782-6.00004-1>.
 - [47] H. W. Kuhn, “THE HUNGARIAN METHOD FOR THE ASSIGNMENT PROBLEM’.”
 - [48] Y. Wang and H. Yang, “Multi-target Pedestrian Tracking Based on YOLOv5 and DeepSORT,” in *2022 IEEE Asia-Pacific Conference on Image Processing, Electronics and Computers, IPEC 2022*,

Institute of Electrical and Electronics Engineers Inc., 2022, pp. 508–514. doi: 10.1109/IPEC54454.2022.9777554.

- [49] N. S. Pun, S. K. Sonbhadra, S. Agarwal, and G. Rai, “Monitoring COVID-19 social distancing with person detection and tracking via fine-tuned YOLO v3 and Deepsort techniques,” May 2020, [Online]. Available: <http://arxiv.org/abs/2005.01385>
- [50] S. Correia, D. Mendes, P. Jorge, T. Brandão, P. Arriaga, and L. Nunes, “Occlusion-Aware Pedestrian Detection and Tracking,” in *2023 30th International Conference on Systems, Signals and Image Processing (IWSSIP)*, IEEE, Jun. 2023, pp. 1–5. doi: 10.1109/IWSSIP58668.2023.10180296.

Anexos

A. Número de identificadores em cada exemplo da Experiência 1

Experiência 1			
Nº do Exemplo	Nº Pessoas no Exemplo	Nº IDs Cliente	Nº Re-IDs Cliente
1	2	20	4
2	2	17	3
3	2	10	4
4	1	19	6
5	5	33	11
6	2	15	3
7	3	13	4
8	3	18	6
9	4	30	9
10	6	20	10
11	3	11	4
12	6	29	8
13	2	27	5
14	2	13	3
15	2	11	2
16	2	14	2
17	1	7	2
18	3	38	13
19	3	24	4
20	3	26	6
21	5	24	8
22	6	31	12
23	1	10	3
24	3	20	6
25	4	32	6
Total	76	512	144

B. Número de identificadores em cada exemplo da Experiência 2

Experiência 2					
Nº do Exemplo	Nº Pessoas no Exemplo	Nº Reidentificações Efetuadas	Nº Reidentificações Erradas	Nº IDs Cliente	Nº Re-IDs Cliente
1	2	16	1	20	2
2	2	5	1	17	3
3	2	8	1	10	3
4	1	13	3	19	4
5	5	15	2	33	7
6	2	4	0	15	2
7	3	3	1	13	4
8	3	12	2	18	5
9	4	15	3	30	7
10	6	11	1	20	7
11	3	2	2	11	5
12	6	7	0	29	6
13	2	15	2	27	4
14	2	6	0	13	2
15	2	4	0	11	2
16	2	8	0	14	2
17	1	2	1	7	2
18	3	6	8	38	11
19	3	7	0	24	3
20	3	13	1	26	4
21	5	13	3	24	8
22	6	20	5	31	5
23	1	5	1	10	2
24	3	8	1	20	4
25	4	7	1	32	5
Total	76	225	40	512	109

C. Número de identificadores em cada exemplo da Experiência 3

Experiência 3					
Nº do Exemplo	Nº Pessoas no Exemplo	Nº Reidentificações Efetuadas	Nº Reidentificações Erradas	Nº IDs Cliente	Nº Re-IDs Cliente
1	2	15	2	20	3
2	2	5	1	17	3
3	2	8	2	10	3
4	1	14	2	19	3
5	5	16	1	33	6
6	2	3	1	15	3
7	3	3	2	13	4
8	3	14	2	18	5
9	4	17	3	30	5
10	6	12	2	20	6
11	3	4	0	11	3
12	6	7	1	29	7
13	2	16	4	27	3
14	2	5	1	13	3
15	2	3	1	11	3
16	2	8	1	14	3
17	1	2	1	7	2
18	3	8	7	38	8
19	3	7	0	24	3
20	3	13	2	26	5
21	5	14	4	24	7
22	6	21	4	31	4
23	1	6	1	10	2
24	3	7	3	20	6
25	4	8	1	32	5
Total	76	236	49	512	105