



INSTITUTO
UNIVERSITÁRIO
DE LISBOA

Aplicação de algoritmos de *deep learning*: Modelo de reconhecimento de imagem e análise do comportamento dos consumidores em lojas de *checkout* autónomo

Miguel Costa Lima Lobo

Mestrado em Business Analytics

Professora Doutora Patrícia A. Filipe, Professora Auxiliar, ISCTE Business School, Departamento de Métodos Quantitativos para Gestão e Economia

Novembro, 2022

“You can't solve a problem with the ways of thinking that created it. So why not change our ways of thinking, and see the world as our only homeland?” (Albert Einstein)



BUSINESS
SCHOOL

Aplicação de algoritmos de *deep learning*: Modelo de reconhecimento de imagem e análise do comportamento dos consumidores em lojas de *checkout* autónomo

Miguel Costa Lima Lobo

Mestrado em Business Analytics

Professora Doutora Patrícia A. Filipe, Professora Auxiliar, ISCTE Business School, Departamento de Métodos Quantitativos para Gestão e Economia

Novembro, 2022

Agradecimentos

À minha orientadora, Professora Doutora Patrícia Filipe, por todo o apoio, dedicação e confiança ao longo de toda esta etapa. Por todos os ensinamentos e oportunidades que me facultou ao longo do meu percurso académico, por ser um exemplo de rigor que levo para a vida;

Ao Diretor de Mestrado, Professor Doutor Raul Laureano, pela disponibilidade e imprescindível ajuda que sempre demonstrou ao longo deste meu percurso académico;

Ao ISCTE, por ter sido a minha segunda casa nestes dois últimos anos, pelos princípios e valores que transmite aos seus discentes, e pela excelência de ensino;

Ao Márcio Ozal, pela ajuda e disponibilidade indispensáveis que demonstrou em todos os momentos, pela influência que teve na minha escolha do tema, e pelos ensinamentos que nunca me negou e que me fizeram ir conseguindo ultrapassar todos os obstáculos;

Ao André Rocha, por me auxiliar numa fase, em que todas as portas se fechavam, e por ter sempre demonstrado disponibilidade para me ajudar;

Aos meus amigos, que apesar de não estarmos constantemente juntos, nunca deixaram de me apoiar nos meus objetivos, e que ao longo destes anos propiciaram memórias que irão perdurar muitos anos;

À minha irmã, que todos os dias me inspira só por ser o tipo de pessoa que é. Por ser o modelo a seguir mais próximo que tenho, e pelo orgulho que demonstra sempre ter pelos feitos que atinjo. Mas principalmente, por ser a minha companheira de vida;

Aos meus avós, por, apesar de muitas vezes não perceberem totalmente os temas com que trabalho, estiveram sempre lá para mim. Pelo amor que sempre demonstraram e pelo carinho e vontade que têm de passar a maior parte do tempo em família. Eles sim, transmitem o real valor da palavra;

Aos meus pais, por me educarem a ser como sou. Por me instruírem os valores necessários para crescer, mas sem nunca deixarem de estar presentes em todas as etapas da minha vida. Ao meu pai por me ter guiado neste caminho, à minha mãe por todos os dias fazer de tudo para me ver suceder, aos dois, por em conjunto, me mostrarem que não há impossíveis.

Resumo

Recentes avanços em campos como *deep learning* e *computer vision*, trouxeram novas possibilidades à área do retalho. Tais avanços forneceram novas ideias e capacidades às lojas de explorarem novos tipos de informação, devido às múltiplas câmaras implementadas no estabelecimento.

Nesta dissertação, é desenvolvida uma investigação que estuda possíveis comportamentos de clientes que esses estabelecimentos podem explorar a partir da análise de imagens.

Neste seguimento, é necessário perceber como se pode analisar e tirar proveito das mesmas e ter a capacidade de reconhecer intervenientes no processo de compra, através da parametrização do melhor modelo, utilizando técnicas de *deep learning*.

Com os objetivos da investigação em mente, é realizada uma extensa pesquisa através de uma Revisão Sistemática da Literatura na área de *Big Data* aplicada a *computer vision*, *deep learning* e os métodos de análise de sensores. Adicionalmente, é efetuada uma análise aos resultados de um questionário, aplicado a consumidores nacionais, que pretendia criar empatia com os mesmo, de modo a tentar traduzir as suas posições em propostas de valor.

Paralelamente, foram recolhidos e analisados dados de imagem provenientes de um retalhista português e utilizados os algoritmos *Yolo* e *Hidden Markov Model*, considerando a sua rapidez e características, mas também a presença em estudos modernos. Os resultados permitiram identificar pessoas com uma *accuracy* média de 85%.

Foram ainda construídos mapas de frequências e *heat maps*, com o principal contributo de enaltecer padrões de consumo e afluências na loja, para aferir o impacto que este tipo de informação pode ter na experiência do consumidor.

Palavras-Chave: *Deep Learning*; *Big Data*; Deteção Objetos; *Checkout* Autónomo; Comportamento Consumidor

JEL Classification System: C45, C55, M15

Abstract

Recent advances in fields such as deep learning and computer vision have brought new possibilities in the retail area. Such advances have provided new insights and capabilities for shops to exploit new types of information due to the multiple cameras deployed in the establishment.

In this dissertation, an investigation is developed that studies possible customer behaviours that these establishments can exploit from image analysis.

In this follow up, it is necessary to understand how one can analyse and take advantage of them and have the ability to recognise actors in the shopping process by parameterising the best model using deep learning techniques.

With the research objectives in mind, an extensive literature review is carried out in the area of Big Data applied to computer vision and deep learning, but also in sensor analysis methods. Additionally, a survey was implemented to a portuguese consumers sample that intended to create empathy with consumers in order to understand their wishes to try to translate them into value propositions.

In parallel, image data from a portuguese retailer was collected and the Yolo algorithm was used together with Hidden Markov Model to analyse it, considering its speed and characteristics, but also its presence in modern studies. The results allow identifying people with an average accuracy of 85%.

Frequency maps and heat maps were also generated, with the main contribution of highlighting consumption patterns and affluence in the shop, to assess the impact that this type of information can have on consumer experience.

Keywords: Deep Learning; Big Data; Object Detection; Autonomous Checkout; Customer Behaviour

JEL Classification System: C45, C55, M15

Índice Geral

Agradecimentos.....	iii
Resumo.....	v
Abstract.....	vii
Índice Geral.....	ix
Índice de Figuras.....	xi
Índice de Gráficos.....	xiii
Índice de Tabelas.....	xv
Índice de Anexos.....	xvii
Glossário de Acrónimos e Siglas.....	xix
1. Introdução.....	1
1.1. Tema e sua importância.....	1
1.1.1. Evolução.....	1
1.1.2. <i>Big Data Analytics</i>	2
1.1.3. <i>Deep Learning</i> e <i>Computer Vision</i>	3
1.2. Problema e questão de investigação.....	5
1.3. Objetivos e contributos.....	6
1.4. Abordagem metodológica.....	7
1.5. Estrutura e organização da dissertação.....	10
2. Revisão da Literatura.....	11
2.1. Protocolo para a revisão sistemática.....	11
2.2. Detecção de atividades de clientes.....	16
2.3. Contexto do estudo.....	16
2.4. Tecnologias <i>wireless</i>	17
2.5. Tipologia de modelos.....	18
2.5.1. Detecção de objetos.....	18
2.5.2. Localização e Análise de comportamentos.....	19
2.6. Avaliação de resultados.....	21
2.7. Avaliação artigos científicos relevantes.....	24
2.8. Contributos da RSL.....	25
3. Metodologia.....	27
3.1. Empatia e Perceção.....	27
3.2. Definição.....	33
3.3. Compreensão dos dados.....	34

3.4	Preparação dos dados.....	35
3.5	Idealização e Modelação	37
3.5.1	Detecção de objetos	37
3.5.2	Comportamento dos consumidores	37
3.5.3	Bibliotecas	38
3.6	Teste e Avaliação.....	39
3.6.1	<i>Accuracy</i>	39
3.6.2	Mapa de frequências e <i>Heat map</i>	41
3.6.3	Cruzamento dos resultados com o questionário	44
3.6.4	Tempo de execução	45
3.7	Implementação.....	46
4.	Conclusões	47
4.1	Contributos e Trabalhos futuros de investigação	48
4.2	Limitações	49
	Referências Bibliográficas	51
	Anexo A: Relatório do questionário.....	57
	Anexo B: Análise de imagens	67

Índice de Figuras

Figura 1 - Evolução dos métodos de análise de dados como forma de potenciar o negócio ...	2
Figura 2 - 5 V's <i>Big Data Analytics</i>	2
Figura 3 - Dimensões de <i>Big Data</i> no retalho	3
Figura 4 - Arquitetura de um sistema de análise de imagens.....	4
Figura 5 - Metodologia CRISP-DM + <i>Design Thinking</i>	9
Figura 6 - Critérios de elegibilidade de artigos	13
Figura 7 - <i>Heat map</i> de afluência nas secções do estabelecimento.....	33
Figura 8 - <i>Frame</i> câmara nº1.....	36
Figura 9 - Caixas de delimitação no reconhecimento de objetos	38
Figura 10 - Exemplo de um <i>frame</i> alterado para mapa de frequências.....	39
Figura 11 - Classes de classificação de objetos	40
Figura 12 - Exemplo de transição de um <i>frame</i> original para <i>heat map</i>	42
Figura 13 - Mapa de frequências e <i>heat map</i> na "zona aberta"	42
Figura 14 - Mapa de frequências e <i>heat map</i> dos corredores.....	43

Índice de Gráficos

Gráfico 1 - Evolução da influência de técnicas de análise de dados	5
Gráfico 2 - Faixa etária dos inquiridos	27
Gráfico 3 - Rendimento mensal dos inquiridos.....	28
Gráfico 4 - Familiaridade com o conceito de loja sem caixas.....	29
Gráfico 5 - Motivação para visitar lojas deste conceito	29
Gráfico 6 - Percepção do processo de compra	30
Gráfico 7 - Características do processo de compra	31
Gráfico 8 - Experiência nos estabelecimentos comerciais	31
Gráfico 9 - Posição na consideração do <i>layout</i> das lojas	32
Gráfico 10 - Relação de rotas e disposição na otimização do <i>layout</i>	32
Gráfico 11 - Histograma de frequências de classes de <i>accuracy</i> na detecção de pessoas	41

Índice de Tabelas

Tabela 1 - Diferenças entre CRISP-DM e <i>Design Thinking</i>	8
Tabela 2 - Critérios de inclusão e exclusão na RSL.....	12
Tabela 3 - Artigos incluídos na RSL entre 2018 a 2022	14
Tabela 4 - Artigos incluídos na RSL entre 2015 e 2017	15
Tabela 5 - Critérios de qualidade para avaliação de artigos.....	15
Tabela 6 - Tipologia de modelos de deteção e localização	20
Tabela 7 - Avaliação resultados dos artigos alvo de RSL.....	22
Tabela 8 - Avaliação da qualidade dos artigos sujeitos a RSL	24
Tabela 9 - Descrição da localização e secção da loja das diferentes câmaras.....	35
Tabela 10 - <i>Accuracy</i> na deteção de pessoas nas imagens	40
Tabela 11 - Tempo de execução das fases do processo	45

Índice de Anexos

Anexo 1 - Informação acerca da familiaridade do conceito - questionário	57
Anexo 2 - Informação sobre a motivação para visitar lojas deste conceito – questionário	58
Anexo 3 - Informação sobre a perceção do processo de compra - questionário.....	60
Anexo 4 - Informação das características do processo de compra – questionário.....	61
Anexo 5 - Informação sobre experiência em estabelecimentos comerciais - questionário.....	62
Anexo 6 - Informação sobre posição na consideração do <i>layout</i> das lojas - questionário.....	63
Anexo 7 - Informação da afluência nas secções no estabelecimento – questionário.....	63
Anexo 8 - Informação sobre faixa etária - questionário.....	64
Anexo 9 - Informação sobre rendimento mensal - questionário	65
Anexo 10 - <i>Frame</i> câmara nº1	67
Anexo 11 - <i>Frame</i> câmara nº 2	67
Anexo 12 - <i>Frame</i> câmara nº3	67
Anexo 13 - <i>Frame</i> câmara nº4	68
Anexo 14 - <i>Frame</i> câmara nº 5	68
Anexo 15 - <i>Frame</i> câmara nº6	69
Anexo 16 - <i>Frame</i> câmara nº7	69
Anexo 17 - <i>Frame</i> câmara nº8.....	70
Anexo 18 - <i>Heat map</i> câmara nº1	70
Anexo 19 - <i>Heat map</i> câmara nº2	70
Anexo 20 - <i>Heat map</i> câmara nº3	71
Anexo 21 - <i>Heat map</i> câmara nº4	71
Anexo 22 - <i>Heat map</i> câmara nº5	71
Anexo 23 - <i>Heat map</i> câmara nº6	72
Anexo 24 - <i>Heat map</i> câmara nº7	72
Anexo 25 - <i>Heat map</i> câmara nº8	73

Glossário de Acrónimos e Siglas

ACC - *Accuracy*

AI – *Artificial Intelligence*

AP – Algoritmo Próprio

BoW – *Bag of Words*

CNN – *Convolutional Neural Network*

CRM – *Customer Relationship Management*

HMM – *Hidden Markov Model*

HOG – *Histogram of Oriented Gradients*

HOI – *Human-Object Interaction*

IA – Inteligência Artificial

IoT – *Internet of Things*

LSTN – *Location-based Social Networking*

PC – *People Counting*

PLN – Processamento de Linguagem Natural

POS – *Point of Sale*

R-CNN – *Region-based Convolutional Neural Network*

RFID – *Radio Frequency Identification*

RGB – *Red, Green and Blue*

RGPD – Regulamento geral sobre a proteção dos dados

RL – Revisão da Literatura

ROI – *Region of Interest*

RSL – Revisão Sistemática da Literatura

SLR – Systematic Literature Review

SVM – *Support Vector Machine*

UWB – *Ultra-wideband*

WoS – *Web of Science*

WLAN – *Wireless Local Area Network*

YOLO – *You Only Look Once*

1. Introdução

A análise de imagens como forma de explorar novos tipos de informação é, cada vez mais, um tópico com maior magnitude. Neste estudo, esse tipo de capacidades é utilizado para analisar vídeos e imagens de clientes em ambientes de retalho.

Desta forma, esta introdução tem como propósito apresentar um enquadramento ao tema, através da elucidação da sua importância, mediante tópicos que lhe são circundantes, como a inserção em *Big Data*, *deep learning* e *computer vision*. Por outro lado, tem também como finalidade evidenciar o problema e questões de investigação, assim como os objetivos da mesma. De referir ainda que é igualmente apresentada uma metodologia assente numa combinação de outras duas metodologias distintas, assim como a estrutura desta dissertação.

1.1. Tema e sua importância

1.1.1. Evolução

Na evolução da análise de dados, de forma a potenciar o negócio, Wedel and Kannan, (2016), no seu estudo, referem que, logo após uma disponibilização de novas formas e fontes de informação, foram sempre introduzidos e desenvolvidos métodos para as trabalhar e prontamente apresentar resultados.

Adicionalmente, acompanhando essa mesma evolução presente na Figura 1, os mesmos destacam a importância dos métodos desenvolvidos, desde os anos 60, na prática atual da tomada de decisão, como *CRM*, *marketing mix* e personalização. Esses métodos servem também como forma de complemento a técnicas de *Big Data Analytics* e métodos de *deep learning* assentes nas técnicas de análise de vídeo, localização e comportamento dos consumidores. Através da Figura 1 é também possível enquadrar a presente investigação no espaço temporal adequado, pela conciliação de técnicas de análise de vídeo com a localização e análise social dos consumidores.

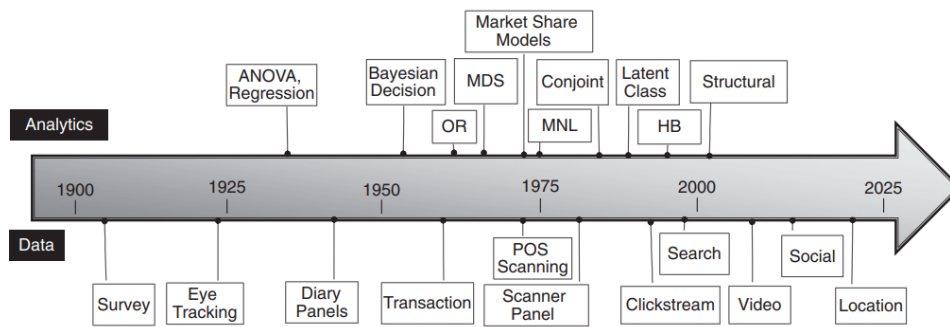


Figura 1 - Evolução dos métodos de análise de dados como forma de potenciar o negócio (Wedel & Kannan, 2016)

Nota: MDS: *Multidimensional scaling*; OR: *Operations Research*; POS: *Point of sale*; MNL: *Multinomial logit model*; HB: *Hierarchical Bayes*.

1.1.2. *Big Data Analytics*

No estudo de Hadi *et al.* (2015), é referido que o *Big Data Analytics* é frequentemente caracterizado pelas suas cinco características, também conhecidas pelos 5 V's, identificados em detalhe na Figura 2, sendo eles: *Volume* (de *terabytes* a *petabytes*); *Velocidade* (análise em tempo real; *streams*); *Variedade* (dados numéricos, texto, imagens, vídeo); *Veracidade* (fiabilidade e validade) e *Valor* (extrair informação valiosa).

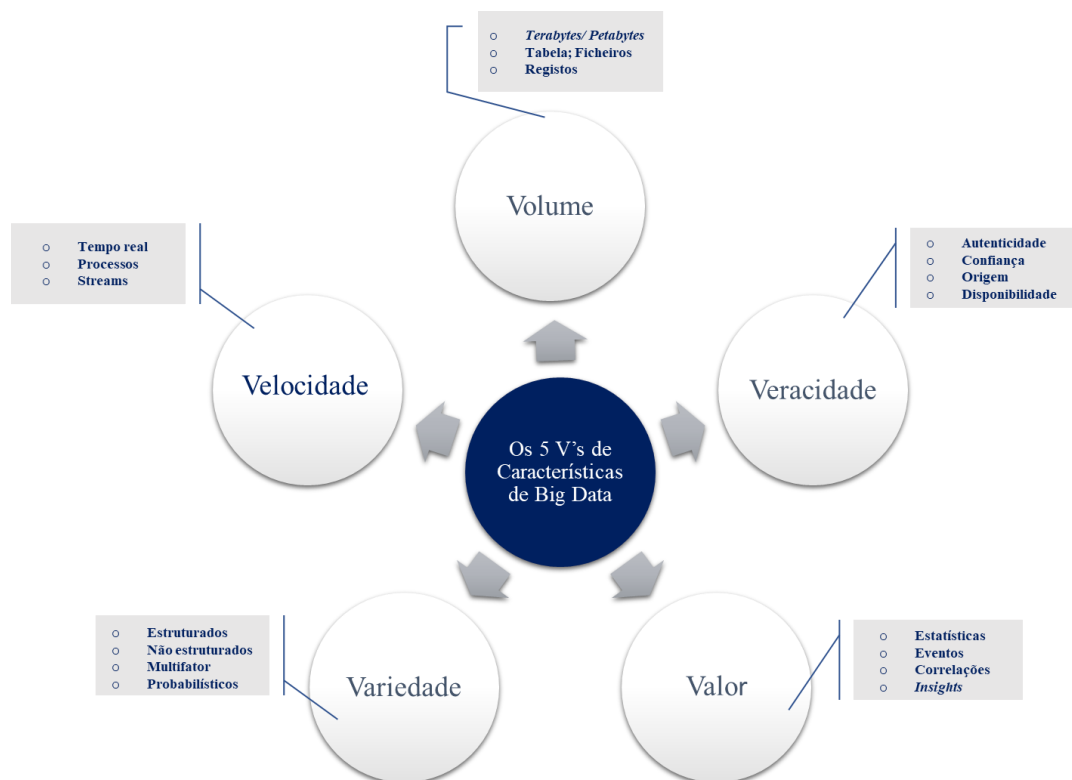


Figura 2 - 5 V's *Big Data Analytics*

Bradlow *et al.* (2017), descrevem cinco dimensões, representadas na Figura 3, onde acreditam que existe potencial para explorar vastos fluxos de informação através de fontes típicas de dados relacionados com *Big Data* na área do retalho.



Figura 3 - Dimensões de *Big Data* no retalho

Nesse estudo, é ainda referido que, a análise de imagens e vídeos para a compreensão de comportamentos de clientes tem a capacidade de ligar todas estas dimensões entre si e, com o aproveitamento da mesma, será possível, não só adotar estratégias mais específicas, mas também medir os seus efeitos de forma mais precisa. Adicionalmente, o estudo de Paolanti *et al.* (2020), descreve as análises dos comportamentos dos consumidores como preocupações perpétuas que têm de ser sempre alvo da análise em estudos deste âmbito.

1.1.3. *Deep Learning* e *Computer Vision*

O *deep learning* tem sido irrefutavelmente bem sucedido nos campos como *computer vision* e reconhecimento de imagens e vídeos (Chai *et al.*, 2021).

As técnicas baseadas em *computer vision* permitem a análise e interpretação de dados relativos a imagens e vídeos (Choudhary & Chaudhury, 2016).

Com as componentes de *Big Data Analytics*, a inteligência artificial e os seus subcampos de *computer vision*, *machine learning* e *deep learning*, atuam como um cérebro a processar e interpretar os vídeos e gravações referentes aos seus “olhos”, neste caso as câmaras (Nguyen *et al.*, 2022). Segundo Choudhary and Chaudhury, (2016), os sistemas de análise de vídeos, em geral, apresentam uma estrutura idêntica à apresentada na Figura 4.

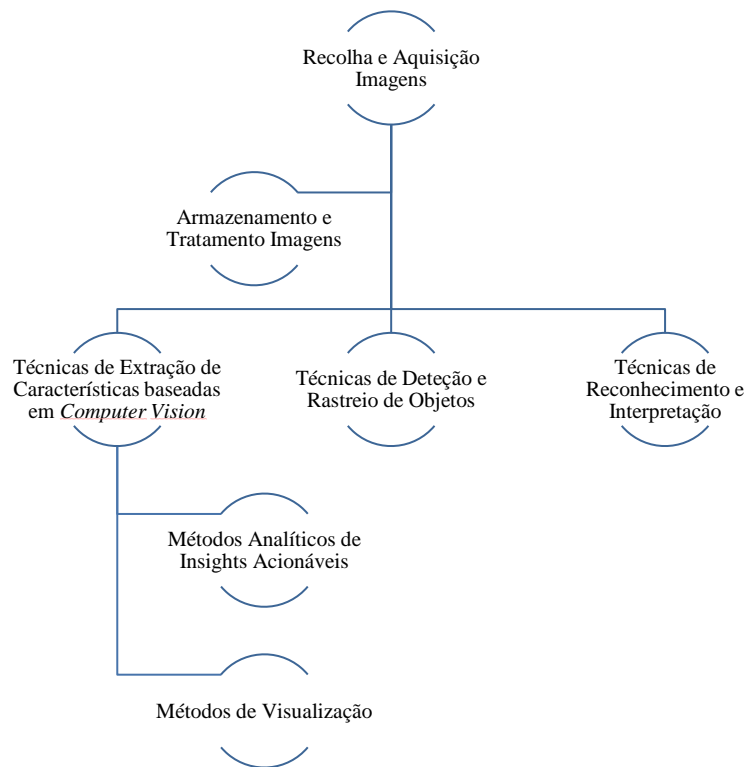


Figura 4 - Arquitetura de um sistema de análise de imagens

É importante referir ainda que Benaich *et al.* (2022), têm vindo a publicar o seu relatório anual sobre o estado da Inteligência Artificial denominado *State of AI Report*. Este relatório analisa os desenvolvimentos mais relevantes em matéria de IA, com o objetivo de desencadear uma conversa informada sobre o estado da mesma e as suas implicações para o futuro, mas também resumir e partilhar os seus conhecimentos com o mundo. Neste mesmo relatório de 2020, foi previsto que as entidades que empregam projetos de IA se expandiriam para além do Processamento de Linguagem Natural (PLN) para alcançar o estado da arte em *computer vision*. No relatório de 2022 a análise de documentos relacionados com esta matéria mostra como esta arquitetura dos modelos se tornou omnipresente, como demonstrado no Gráfico 1.

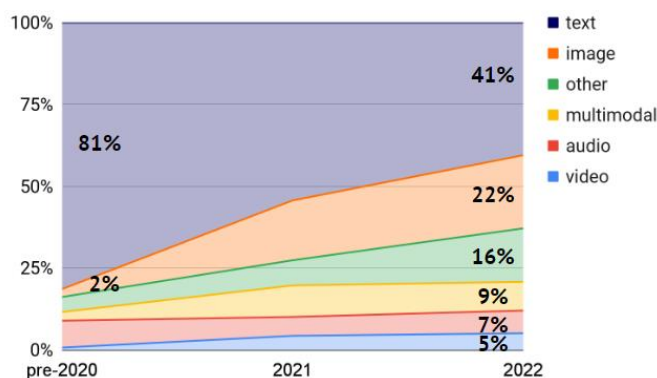


Gráfico 1 - Evolução da influência de técnicas de análise de dados (Bellini *et al.*, 2022)

Relativamente às informações referentes ao Gráfico 1, é de realçar a perda de importância de estudos assentes na análise de informação proveniente de dados de texto, com uma descida de metade dos seus pontos percentuais. Por outro lado, é evidente um maior peso nas investigações que têm por base a análise de imagens, com a sua influência a subir mais de dez vezes, com um aumento de vinte pontos percentuais.

Neste seguimento, a área do retalho está a evoluir a um ritmo acelerado, com as empresas cada vez mais centradas na forma de utilizar as tecnologias de inteligência artificial para reformular a natureza da indústria retalhista (Wei *et al.*, 2020). Todavia, esta causa também se deve à pressão da digitalização no setor onde os retalhistas necessitaram de lidar com estas mudanças através da implementação de novos modelos de negócio e estratégias competitivas com o objetivo de satisfazer as necessidades dos consumidores (Gazzola *et al.*, 2022).

1.2. Problema e questão de investigação

Nos dias que correm, cada vez mais as deslocações ao supermercado não são uma prática desejada, isso devido a simples fatores como, o elevado número de aglomerados, tempo de fila de espera, e o excessivo tempo passado dentro da loja, onde, por vezes, nem metade é considerado tempo útil (Rengarajan *et al.*, 2019).

A partir de 2018, um novo paradigma impulsionou uma mudança na situação corrente, com o lançamento da Amazon Go, a primeira loja no mundo sem caixas de pagamento. Neste seguimento, é de enaltecer que, em Portugal, já várias entidades adotaram esse conceito, com a abertura de lojas como o Continente Labs, Pingo Doce & Go e Galp Smart Store.

Ives *et al.* (2019), enaltecem que, a capacidade dessas mesmas lojas explorarem novos tipos de informação e reduzir constrangimentos para o consumidor é infinda. Isto torna-se possível devido às centenas de câmaras e sensores implantados no interior do estabelecimento que poderão fornecer novos dados e informação para, por exemplo, definir afinidades.

Todavia, apesar de muitos estudos recentes terem em consideração a análise de atividades de clientes, ainda são escassos aqueles que utilizam essa informação para melhorar a experiência do consumidor (Ziaie *et al.*, 2021). Posto isto, neste âmbito, a presente investigação tem como principal finalidade perceber como é que o uso de uma grande variedade de componentes de *IoT*, como câmaras e sensores, poderiam ajudar a retirar informações dos comportamentos dos clientes dentro das lojas, de forma a fornecer uma melhor experiência aos mesmos, simplificando o ato de compra em lojas físicas.

Não obstante, aplicações de tecnologias “*just walk out technology*”, implementadas por retalhistas, que têm como objetivo simplificar o processo de compra em lojas físicas, têm experienciado dificuldades com os modelos/algoritmos que tratam imagens do espaço. Alguns desses constrangimentos são nomeadamente a oclusão, onde uma pessoa é bloqueada de vista por algo na loja (Nguyen *et al.*, 2022); o facto dos indivíduos se fixarem muito próximo uns dos outros (Dogan *et al.*, 2019), ou vários itens agarrados simultaneamente pelos clientes (Merad *et al.*, 2016).

Neste seguimento, esta investigação procura responder à seguinte questão de investigação: Como se podem detetar nitidamente pessoas nas imagens de ambientes de retalho e analisar os seus comportamentos, de modo a medir o impacto que esse tipo de informação pode ter na experiência do consumidor.

1.3. Objetivos e contributos

Conforme a questão de investigação apresentada, esta investigação está assente em três principais objetivos, sendo eles: O1) Avaliar técnicas de recolha, tratamento e processamento de imagens, O2) Reconhecer, com precisão, elementos e objetos presentes em imagens de vídeo, e O3) Avaliar o impacto e os indicadores, que o aproveitamento das características de lojas autónomas podem fornecer.

Adicionalmente, para certos objetivos principais, foram delimitados objetivos concretos:

- ✓ O1: Avaliar técnicas de recolha, tratamento e processamento de imagens;
- ✓ O2.2: Avaliar diferentes técnicas de *deep learning* de deteção de objetos em relação à base de dados disponível;

- ✓ O2.2: Criar um modelo de deteção de objetos parametrizado para identificar pessoas;
- ✓ O3.1: Caracterizar a perceção da população para ambientes de retalho autónomo, mas também tradicionais;
- ✓ O.3.2: Criar um modelo de localização e análise de comportamento dos consumidores.

A resposta às questões de investigação, através da materialização dos objetivos apresentados, numa vertente ligada à literatura, poderá auxiliar a colmatar as dificuldades dos algoritmos que tratam imagens, nomeadamente, em fatores como a correta identificação de objetos e capacidade de reconhecer elementos em ambientes de carácter dinâmico. Adicionalmente, este estudo, ao se apresentar aplicado a uma entidade de retalho portuguesa, sendo que são poucos os estudos relacionados com análise de imagens nestes ambientes, vem contribuir para a divulgação de resultados mais genuínos, no âmbito nacional. Noutra vertente, na perspetiva académica, esta investigação poderá potenciar outros estudos propínquos, seja em contornos de uma análise mais aprofundada da aceitação da população destes ou formatos similares de lojas, seja em aperfeiçoamento ou atualização de novas técnicas que devam ser sujeitas a investigação.

1.4. Abordagem metodológica

De forma a tirar o maior proveito das metodologias, uma combinação de CRISP-DM com *Design Thinking* pode-se apresentar como uma metodologia que conjuga um profundo entendimento do estudo em análise, através da observação das necessidades dos consumidores e empatia que o *Design Thinking* procura, com a fundamentação técnica necessária do CRISP-DM.

O *Design Thinking* é uma metodologia que procura perceber o consumidor, utilizando a empatia humana para redefinir o problema (Pande & Bharathi, 2020). Esta metodologia divide-se em cinco fases sendo elas: Empatia: Perceber e criar empatia com o consumidor ou cliente e desvendar as suas reais necessidades; Definição: Desmontar e sintetizar os dados recolhidos na etapa anterior e definir de forma clara o problema a lidar; Idealização: Criação de ideias e alternativas, abertura de horizontes, exploração de alternativas e soluções; Protótipo: Criação de soluções e experimentação; Teste: Teste de várias soluções (Kwon *et al.*, 2021).

O CRISP-DM, por outro lado, é uma metodologia compreensiva concebida para relacionar e guiar as etapas mais comuns em projetos de *data mining* com seis fases onde descreve o ciclo de vida de um projeto ligado a dados e a base para desenvolver esse mesmo projeto (Martinez-Plumed *et al.*, 2021). Assim, as seguintes fases caracterizam-se por: Compreensão do Negócio: Fase inicial para perceber os objetivos do projeto e objetivos numa perspectiva do negócio; Compreensão dos dados: Recolha inicial dos dados e formas de familiarização com os mesmos, desde perceber a sua qualidade até os primeiros *insights* gerados; Preparação dos Dados: Abrange as atividades necessárias de tratamento dos dados para gerar o melhor *dataset* possível para a fase de modelação, desde seleção dos atributos mais relevantes e limpeza dos dados até à sua transformação e construção de novos; Modelação: Onde são aplicadas diversas técnicas e parâmetros definidos; Avaliação: Avaliação dos modelos e revisão dos passos executados para a elaboração dos mesmos; Implementação: Conhecimento gerado organizado e apresentado de uma forma útil para o utilizador final (Plotnikova *et al.*, 2020).

Ahmed *et al.* (2019), no seu estudo, apresentam um conjunto de parâmetros diferenciadores entre as duas metodologias que se pretendem conciliar, como apresentados na Tabela 1.

Tabela 1 - Diferenças entre CRISP-DM e *Design Thinking*

	<i>Design Thinking</i>	CRISP-DM
Propósito	Criar ideias inovadoras e resolver problemas mal definidos.	Criar um processo fiável e repetível para a entrega de valor.
Problema	Investigado pelas equipas e baseado nas necessidades.	Definido pelo Cliente ou pelo Negócio.
Estratégia	Envolvimento com pessoas (centrado no utilizador).	Priorização baseada em valor para o negócio.
Objetivos	Conceitos definidos para produtos ou serviços.	<i>Software</i> de trabalho com características finalizadas.
Características	Entrevistas a clientes, mapeamento de empatia, modelos e testes rápidos.	Análise de dados e seleção de características, tarefas de previsão, avaliação.
Medição Sucesso	Encontrar potenciais soluções a partir da exploração do problema na ótica do cliente.	Construir a solução adequadamente.

Fonte: Adaptado de Ahmed *et al.* (2019).

A combinação destas metodologias permitirá uma maior flexibilidade nos processos e oportunidades de explorar as vantagens de ambas. Esta combinação poderá ser útil na medida que, em fases do projeto onde se sentiria que o CRISP-DM restringiria de certo modo os objetivos pretendidos, complementar-se-á com as qualidades do *Design Thinking*, e vice-versa.

Um exemplo disso é ter a possibilidade de complementar a fase de avaliação com a de teste do *Design Thinking*. Adicionalmente, no caso da compreensão do negócio, onde é preciso que se tenha um profundo entendimento da situação, através da observação das necessidades das pessoas, a empatia em que o *Design Thinking* se baseia, poderia ser um ponto base para entender melhor o problema onde o projeto se insere. Isto porque, embora muitas vezes as pessoas não possam dizer quais são suas necessidades, os seus comportamentos poderão revelar *insights* valiosos sobre necessidades não satisfeitas (Kwon *et al.*, 2021).

Todavia, existem fases no *Design Thinking* onde se poderia sentir que este é demasiado genérico e necessita de uma estrutura para o sustentar, como nas fases de recolha e preparação dos *datasets*, ou regras na avaliação dos modelos, de modo a ir de encontro a uma solução ótima, oposto a inúmeras potenciais soluções.

O objetivo final poderia passar por cada uma das metodologias servir de solução, onde a outra pecaria em apresentar os resultados mais desejados.

Assim, esta combinação de metodologias poderia ser dividida nas fases seguintes, presentes na Figura 5.

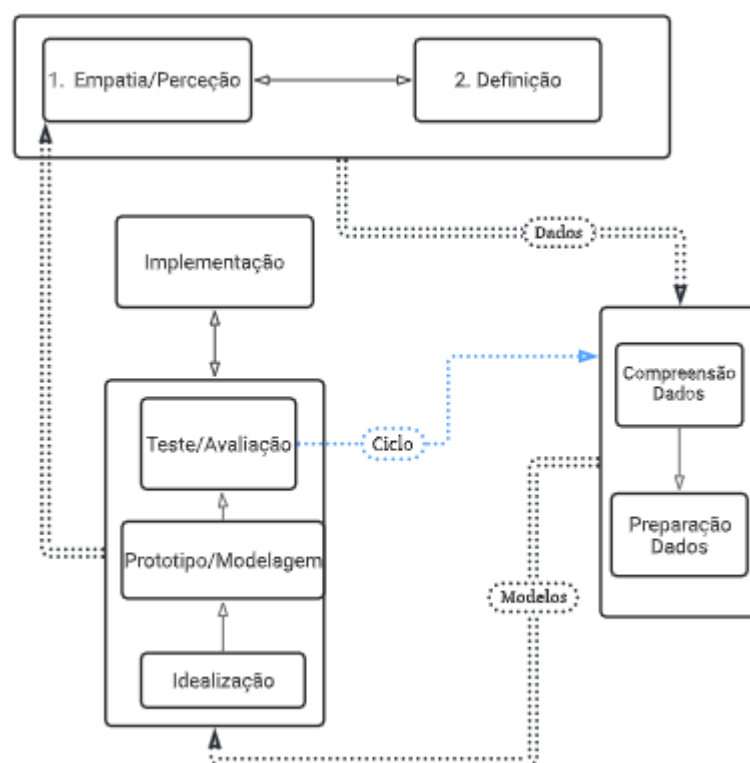


Figura 5 - Metodologia CRISP-DM + *Design Thinking*

1.5. Estrutura e organização da dissertação

Esta investigação foi desenvolvida com o propósito de perceber, como se poderá tirar proveito de imagens relativas aos comportamentos dos consumidores dentro das lojas, de modo a melhorar a sua experiência. A mesma encontra-se organizada em quatro capítulos que contemplam as diversas fases deste estudo.

O primeiro capítulo é referente a tópicos como o tema, questões de investigação e objetivos. Nestes tópicos procedeu-se a uma breve apresentação da evolução do tema e enquadramentos em áreas como de *Big Data Analytics*, *deep learning* e *computer vision*. Adicionalmente foi também descrita e fundamentada a escolha da interligação de duas metodologias (CRISP-DM e *Design Thinking*).

No segundo capítulo está presente a revisão da literatura, em formato de revisão sistemática da literatura (RSL) conduzida de forma a encontrar outros estudos relativos a técnicas de análise de imagens e vídeos em ambientes de retalho. Este capítulo também se encontra dividido em oito subcapítulos, entre eles o protocolo para a RSL, tipologia de modelos mas também capítulos que apresentam aspetos relacionados, tanto com a avaliação de resultados, como com a qualidade dos artigos identificados pela *query* de pesquisa.

A metodologia identificada em 1.4 Abordagem metodológica, foi desenvolvida no terceiro capítulo, onde foi apresentado todo o processo, desde o método de criar empatia com os consumidores, até às fases de modelação e avaliação dos modelos, passando pelas etapas essenciais de compreensão e preparação dos dados. No quarto e último capítulo são apresentadas as conclusões da investigação, assim como, os contributos e limitações identificadas. Este capítulo é, também, complementado com a apresentação de possíveis trabalhos futuros.

2. Revisão da Literatura

2.1. Protocolo para a revisão sistemática

A revisão sistemática da literatura foi conduzida de forma a encontrar outros estudos relativos aos métodos de análise de dados de sensores, de modo a perceber o comportamento dos consumidores em ambientes de retalho. Este processo foi dividido em sete fases como demonstrado na Figura 6. Devido a esta revisão ser relativa a áreas de investigação como de Negócio e Ciências da Computação, a procura automática foi restringida à base de dados científica *Scopus*, devido a ser considerada uma das maiores bases de dados, onde os mesmos são selecionados através de um processo de seleção de conteúdos seguido de um processo contínuo de avaliação. Adicionalmente, segundo Singh *et al.* (2021), a *Scopus* tem um maior número de revistas exclusivas do que a *Web of Science* e quase todos os jornais e publicações indexados na *WoS* também estão presentes na *Scopus*. Para este método de pesquisa foram identificados apenas artigos de jornais. Capítulos de livros, *papers* de conferências e *posters* foram excluídos.

A revisão da literatura apresentada tem como propósito responder a questões de investigação, com base no estudo do estado da arte em modelos de *computer vision* e/ou *deep learning*, orientados para a área do retalho. Desta forma pretende-se responder à questão: Como se podem detetar nitidamente pessoas nas imagens de ambientes de retalho e analisar os seus comportamentos de modo a medir o impacto que esse tipo de informação pode ter na experiência do consumidor. Posto isto, é necessário responder especificamente às seguintes questões: (1) Em que atividades de clientes a análise se deverá incidir e fornecem mais valor de *insights*?; (2) Qual o contexto do estudo?; (3) Que/Quais tecnologias *wireless* são mais utilizadas para o âmbito e qual a sua evolução?; (4) Quais os modelos utilizados para a deteção e localização de objetos?; (5) Quais os modelos utilizados para a análise de comportamentos dos consumidores?; (6) Quais as maiores limitações nas análises de sensores?; (7) Como são analisados os resultados obtidos?

A busca foi conduzida pela construção de blocos de *keywords*. Essas *keywords* foram encontradas através de um processo simultâneo de identificação na literatura de termos sinónimos dos títulos de cada bloco, e validação dessas mesmas *keywords* por especialistas. Um artigo teria então que corresponder pelo menos a uma *keyword* de cada um dos blocos, de modo a ser um possível alvo de revisão. Bloco 1: Reconhecimento de objetos; Bloco 2: *Deep learning*; Bloco 3: *Checkout* Autónomo. A *query* definida foi então:

(("Object detect*" OR "Gesture recogn*" OR "Grasp detection" OR "People track*" OR "Shopper preference*" OR "Customer Preference*" OR "Customer Behav*" OR "People Behav*" OR "Shopper Behav*" OR " Purchase Behav*" OR "Human act*" OR "Motion estimation" OR "Customer Navigation*" OR "Gaze" OR "Gesture Observ*") AND ("Deep learning" OR "Machine learning" OR "Artificial intelligence" OR "Sensing tech*" OR "Computer Vision" OR "Machine vision" OR "IoT" OR "Internet of Things") AND ("Smart store*" OR "Intelligent store*" OR "Just walk out" OR "In-store" OR "Autonomous store" OR "Cashierless" OR "Self checkout" OR "Intelligent Retail" OR "Smart Retail" OR "Physical store*"))

De referir que o asterisco presente (*) toma a posição de *wildcard* para termos desconhecidos e diferentes inflexões da mesma palavra.

O processo de revisão da seleção de artigos foi elaborado em dois grandes processos, o processo preliminar onde foi analisado, numa fase inicial, a aplicação dos critérios de inclusão e exclusão apresentados na Tabela 2, e posteriormente confrontados com os seus títulos, *abstracts* e *keywords* de modo a garantir que iriam de encontro aos estudos pretendidos com base nos requisitos enunciados.

Tabela 2 - Critérios de inclusão e exclusão na RSL

Critérios de Inclusão	Artigos publicados em Journals Artigos que abordam lojas sem caixas Artigos que relacionam <i>IoT</i> e <i>Big Data</i> com lojas sem caixas Artigos fundamentados com componente empírica
Critérios de Exclusão	Capítulos de Livros, <i>Papers</i> de Conferências, Posters e Editoriais Artigos Duplicados Artigos que não estão na língua inglesa Artigos de áreas fora de <i>Computer Science</i> e <i>Business</i> Artigos sem componente empírica <i>Artigos anteriores a 2015</i>

Foram ainda criados critérios próprios adicionais, de forma a restringir de forma personalizada os resultados provenientes da pesquisa e excluir artigos através do seu *abstract* e *keywords*. Cada artigo deverá preencher os requisitos de abordar os seguintes tópicos:

- Espaço físico: Os artigos foram direcionados para análise de comportamentos dos consumidores e a sua localização dentro de um espaço físico e não a sua análise em ambientes virtuais ou *online*;
- Sensores: Os artigos tiveram em conta e abordaram sensores em contexto visual e comportamental. Artigos sem estas componentes não foram tidos em consideração na medida da falta da componente de *computer vision* ou técnica similar;
- Comportamento Autónomo: Os estudos seleccionados apresentaram a capacidade de ser aplicados em ambientes de *checkout* autónomo, mesmo que no presente estudo sejam projetados para ambientes híbridos.

Assim sendo, dos 677 artigos encontrados pela *query* foram eliminados 648, formando então 29 artigos incluídos na RL desta investigação, como demonstra a Figura 6.

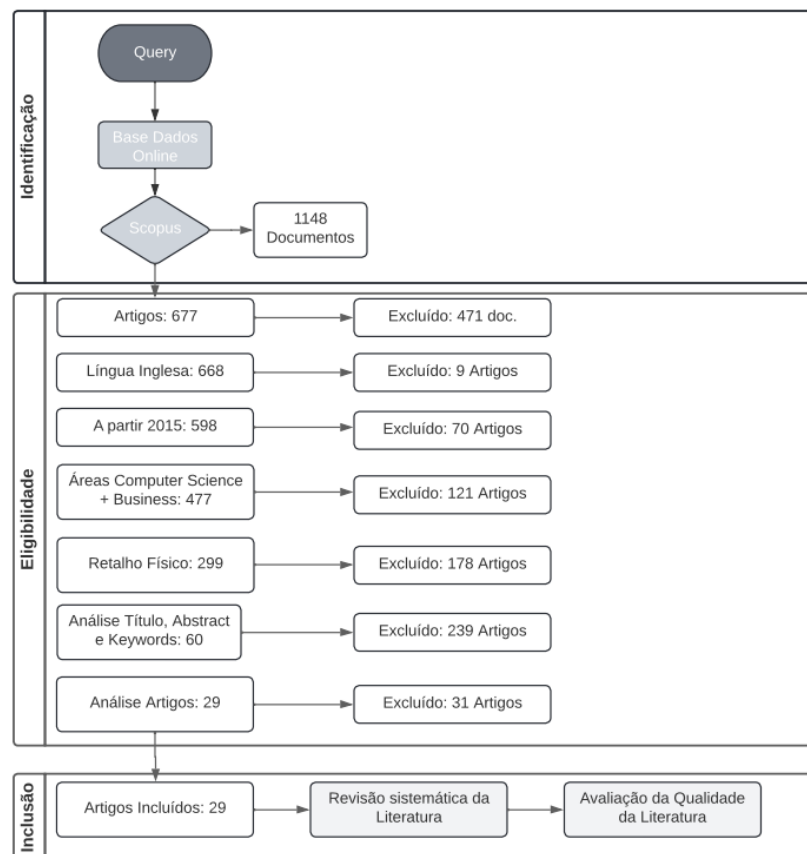


Figura 6 - Critérios de elegibilidade de artigos

A Tabela 3 apresenta os artigos incluídos após o lançamento da Amazon Go (depois de 2018) e a Tabela 4 os artigos publicados antes desse evento.

Tabela 3 - Artigos incluídos na RSL entre 2018 a 2022

ID	Ano	Título	Autores	Journal	Quartil
1	2022	<i>Monitoring Browsing Behavior of Customers in Retail Stores via RFID Imaging</i>	Ali, K., Liu, A.X., Chai, E., Sundaresan, K.	<i>IEEE Transactions on Mobile Computing</i>	Q1
2	2022	<i>Multi Clustering Recommendation System for Fashion Retail</i>	Bellini, P., Palesi, L.A.I., Nesi, P., Pantaleo, G.	<i>Multimedia Tools and Applications</i>	Q2
3	2022	<i>A Dynamic HMM-Based Real-Time Location Tracking System Utilizing UHF Passive RFID</i>	Chowdhury, B.D.B., Masoud, S., Son, Y.-J., Kubota, C., Tronstad, R.	<i>IEEE Journal of Radio Frequency Identification</i>	Q3
4	2022	<i>Cross Validation Voting for Improving CNN Classification in Grocery Products</i>	Domingo, J.D., Aparicio, R.M., Rodrigo, L.M.G.	<i>IEEE Access</i>	Q1
5	2022	<i>Mask R-CNN with New Data Augmentation Features for Smart Detection of Retail Products</i>	Hsia, C.-H., Chang, T.-H.W., Chiang, C.-Y., Chan, H.-T.	<i>Applied Sciences (Switzerland)</i>	Q2
6	2022	<i>When AI meets store layout design: a review</i>	Nguyen, K., Le, M., Martin, B., Cil, I., Fookes, C.	<i>Artificial Intelligence Review</i>	Q1
7	2021	<i>Employing Video-based Motion Data with Emotion Expression for Retail Product Recognition</i>	Alkhdre, A.B., Alshantqiti, A.M.	<i>International Journal of Advanced Computer Science and Applications</i>	Q3
8	2021	<i>Customer segmentation through path reconstruction</i>	Carbajal, S.G.	<i>Sensors</i>	Q2
9	2021	<i>AI-powered service robotics for independent shopping experiences by elderly and disabled people</i>	Ghazal, M., Yaghi, M., Gad, A., El Bary, G., Alhalabi, M., Alkhdher, M., El-Baz, A.S.	<i>Applied Sciences (Switzerland)</i>	Q2
10	2021	<i>Catch me if you scan: Data-driven prescriptive modeling for smart store environments</i>	Hauser, M., Flath, C.M., Thiesse, F.	<i>European Journal of Operational Research</i>	Q1
11	2021	<i>Deep learning based real age and gender estimation from unconstrained face image towards smart store customer relationship management</i>	Islam, M.M., Baek, J.-H.	<i>Applied Sciences (Switzerland)</i>	Q2
12	2021	<i>Implementation of Offline Consumer Behavior Tracking</i>	Kao, Y.-T., Lin, H.-W., Qiu, D.-Y.	<i>International Journal of Pattern Recognition and Artificial Intelligence</i>	Q3
13	2021	<i>Understanding videos at scale: How to extract insights for business research</i>	Schwenzow, J., Hartmann, J., Schikowsky, A., Heitmann, M.	<i>Journal of Business Research</i>	Q1
14	2021	<i>A hierarchy-based system for recognizing customer activity in retail environments</i>	Wen, J.; Guillen, L; Abe, T; Suganuma, T.	<i>Sensors</i>	Q2
15	2021	<i>An intelligent self-service vending system for smart retail</i>	Xia, K., Fan, H., Huang, J., Jian, Q., Wei, D.	<i>Sensors</i>	Q2
16	2020	<i>In-store customer traffic and path monitoring in small-scale supermarket using UWB-based localization and SSD-based</i>	Alipio, M.I., Peñalosa, K.M.T., Unida, J.R.C.	<i>Journal of Ambient Intelligence and Humanized Computing</i>	Q1
17	2020	<i>Deep understanding of shopper behaviours and interactions using RGB-D vision</i>	Paolanti, M., Pietrini, R., Mancini, A., Frontoni, E., Zingaretti, P.	<i>Machine Vision and Applications</i>	Q2
18	2020	<i>A new intelligent supermarket security system</i>	Zhang, Y., Jin, S., Wu, Y., Zhao, T., Yan, Y. Li, Z., Li, Y.	<i>Neural Network World</i>	Q3
19	2019	<i>Analyzing of gender behaviors from paths using process mining: A shopping mall application</i>	Dogan, O., Bayo-Monton, J.-L., Fernandez-Llatas, C., Oztaysi, B.	<i>Sensors</i>	Q2

20	2019	<i>DeepStore: An Interaction-Aware WideDeep Model for Store Site Recommendation with Attentional Spatial Embeddings</i>	Liu, Y., Guo, B., Li, N., Zhang, J., Chen, J., Zhang, D., Liu, Y., Yu, Z., Zhang, S., Yao, L.	<i>IEEE Internet of Things Journal</i>	Q1
21	2019	<i>An effective interest identification technique to enhance sales performance in supermarkets</i>	Priyanka, G., Sathya Bama, S., Aiswaryadevi, V.J., Sruthi, M.S., Soundarya, S.	<i>International Journal of Innovative Technology and Exploring Engineering</i>	Q4
22	2018	<i>Modelling and Forecasting Customer Navigation in Intelligent Retail Environments</i>	Paolanti, M., Liciotti, D., Pietrini, R., Mancini, A., Frontoni, E	<i>Journal of Intelligent and Robotic Systems: Theory and Applications</i>	Q2

Tabela 4 - Artigos incluídos na RSL entre 2015 e 2017

ID	Ano	Título	Autores	Journal	Quartil
23	2017	<i>Mining Customer Preference in Physical Stores from Interaction Behavior</i>	Chen, Y., Zheng, Z., Chen, S., Sun, L., Chen, D.	<i>IEEE Access</i>	Q1
24	2017	<i>Store layout optimization using indoor positioning system</i>	Hwangbo, H; Kim, J; Lee, Z., Kim, S.	<i>International Journal of Distributed Sensor Networks</i>	Q3
25	2017	<i>Customer behavior classification using surveillance camera for marketing</i>	Liu, J., Gu, Y., Kamijo, S.	<i>Multimedia Tools and Applications</i>	Q2
26	2016	<i>Video analytics revisited</i>	Choudhary, A., Chaudhury, S.	<i>IET Computer Vision</i>	Q2
27	2016	<i>Improving retail efficiency through sensing technologies</i>	Quintana, M., Menendez, J.M., Alvarez, F., Lopez, J.P.	<i>Pattern Recognition Letters</i>	Q1
28	2016	<i>Purchase behavior analysis through gaze and gesture observation</i>	Merad, D., Drap, P., Lufimpu-Luviya, Y., Iguernaissi, R., Fertil, B.	<i>Pattern Recognition Letters</i>	Q1
29	2015	<i>Detecting and tracking people in real time with RGB-D camera</i>	Liu, J., Liu, Y., Zhang, G., Zhu, P., Chen, Y.Q.	<i>Pattern Recognition Letters</i>	Q1

De seguida, de modo a avaliar a qualidade dos artigos selecionados após a revisão dos mesmos, com o alvo de identificar a sua contribuição para o estudo proposto, foram realizadas diversas perguntas separadas em 8 grandes temas, como apresentado na Tabela 5.

Tabela 5 - Critérios de qualidade para avaliação de artigos

Atividades dos Clientes	Q1	O estudo está direcionado para a compreensão do consumidor como aspeto chave?
	Q2	Aborda, claramente, o conceito de atividades de clientes?
Contexto do Estudo	Q3	Evidencia de forma clara o contexto do estudo?
	Q4	Aborda a existências ou eventual existência de lojas como a Amazon Go?
Comportamentos Consumidores	Q5	Aborda interações produto-cliente?
	Q6	Evidencia a análise de trajetórias e a sua importância?
Tecnologias Wireless	Q7	Justifica a escolha da(s) tecnologias <i>wireless</i> a utilizar e como são extraídos os dados?
	Q8	Aborda a conciliação de diferentes tecnologias?
Modelos Detecção/Localização	Q9	Descreve técnicas de deteção de objetos?
	Q10	Compara técnicas de deteção de objetos?
	Q11	Aborda vários parâmetros na deteção objetos (ROI; PC; HOI)?

Modelos Análise Comportamento Consumidores	Q12	Aborda várias técnicas de reconhecimento de comportamentos do consumidor?
	Q13	Compara técnicas de reconhecimento de comportamentos do consumidor?
	Q14	Utiliza o modelo para otimização do <i>layout</i> do espaço?
Limitações	Q15	Aborda limitações referentes ao ambiente de análise e à privacidade dos dados?
	Q16	Descreve passos para ultrapassar ou minimizar os impactos das limitações?
Avaliação Resultados	Q17	Realiza testes de avaliação de resultados?
	Q18	Compara com a literatura os resultados obtidos?
	Q19	Descreve os contributos do estudo?

2.2. Detecção de atividades de clientes

Um objeto, no âmbito do retalho, pode referir-se tanto a produtos em si, como clientes, trabalhadores, carros de compras e mesmo obstáculos, todavia a maior parte dos estudos relativos a esta área estão associados e aplicados à deteção de seres humanos num ambiente retalhista (Wen *et al.*, 2021).

Paolanti *et al.* (2018, 2020), referem, nos seus estudos, que ter a capacidade de compreender os comportamentos do consumidor é uma das chaves para o sucesso para retalhistas. Liu *et al.* (2017), reportam a importância de adotar essa direção pois, a melhoria da experiência do consumidor irá levar, por sua vez, a uma maior lucratividade da empresa. Já Merad *et al.* (2016) e Quintana *et al.* (2016), assumem que a perceção dos comportamentos dos consumidores são preocupações perpétuas, que têm de ser sempre o alvo da análise em estudos deste âmbito.

2.3. Contexto do estudo

No início do ano de 2018, recentes avanços na área de inteligência artificial e nos seus campos em *computer vision*, *machine learning* e especialmente *deep learning* levaram a um novo paradigma na área do retalho (Nguyen *et al.*, 2022).

Hauser *et al.* (2021), chegam a referir mesmo que, provavelmente, o exemplo mais proeminente é a cadeia de lojas Amazon Go, nas quais todo o processo de *checkout* é automatizado e não requer a existências de caixas no seu funcionamento. Funcionamento esse que é baseado na combinação complexa de reconhecimento de imagens e sensores, de forma a detetar que produtos os clientes retiram dos expositores e colocam no cesto de compras, mas também como forma de perceber os comportamentos dos mesmos, como as suas trajetórias e seleções de produtos (Ghazal *et al.*, 2021 e Hauser *et al.*, 2021).

O conceito de compras "*Just Walk Out*" significa que as pessoas podem escolher rapidamente os produtos sem necessitarem de filas de espera ou *checkout*. Para conseguir "*Just Walk Out*", são necessárias tecnologias avançadas a serem integradas, tais como *computer vision*, câmaras *RGB* e fusão de sensores. Para atingir o objetivo de implantar em dispositivos, espera-se que as redes de detecção de objetos sejam, não só, suficientemente precisas, mas também rápidas e de pequena dimensão (Xia *et al.*, 2021).

2.4. Tecnologias *wireless*

O conjunto de parâmetros extraídos de cada um dos artigos está relacionado com a extração dos dados e com a plataforma e/ou *hardware* usados no sistema assim como o tipo de sensores utilizados. Esta análise torna-se vantajosa, na medida que reúne os dados presentes em cada estudo e pode revelar as diferentes características e possibilidades das diferentes tecnologias *wireless* de detecção e localização de objetos.

Posto isto, os vários estudos divergem na sua abordagem na aplicação de sensores para a detecção de pessoas. Quintana *et al.* (2016), no seu estudo sobre a melhoria do setor do retalho referem que o aparecimento de novos tipos de sensores *RFID* (*Radio-frequency identification*), uma forma de comunicação *wireless* que incorpora campos eletromagnéticos de modo a identificar e rastrear objetos, está a conduzir a novos sistemas híbridos com boas performances. Contudo Chen *et al.* (2017), contrapõem esta ideologia referindo que tecnologias como *RFID* e *smart glasses* como os propostos por Merad *et al.* (2016), apresentam falta de escalabilidade devido ao elevado custo de implantação e manutenção. Assim sendo, propõem a adoção de tecnologia baseada em WiFi (*WiFi Logs*), mas na perspetiva de grandes superfícies comerciais.

Por outro lado, Paolanti *et al.* (2018), introduzem um sistema baseado em tecnologia *UWB* (*ultra-wideband*), uma tecnologia rádio que consome baixos níveis de energia para comunicações de curto alcance e largura de banda elevada. Adicionalmente, enaltecem que as aplicações *wireless* convencionais, como algumas das enumeradas anteriormente (*RFID* e *WLAN*) já não têm a capacidade, só por si mesmas, de alcançar o valor necessário e de apresentar elevados níveis de precisão em tempo real.

São vários os estudos que abordam a temática do conhecimento do comportamento dos consumidores na área do retalho através da análise de dados visuais provenientes de câmaras *RGB* ou *RGB-D* (*RGB dept camera*). Liu *et al.* (2017) e Wen *et al.* (2021), nos seus estudos utilizam câmaras de vigilância tradicionais, estas que, nos dias que correm, já apresentam incorporadas características *RGB*, isto é capacidade de criar imagens que reproduzem a visão humana, capturando a luz em comprimentos de onda vermelhos (*Red*), verdes (*Green*) e azuis (*Blue*) e, por outro lado, Liu *et al.* (2015); Quintana *et al.* (2016) e Paolanti *et al.* (2018), exploraram as capacidades de câmaras *RGB-D*.

Liu *et al.* (2015), referem que, a maioria dos primeiros estudos da área, utilizavam as câmaras convencionais e que, apesar de muitos desses relatarem que tais métodos podem conduzir a bons resultados de deteção e rastreio nos seus parâmetros de referência, o seu desempenho irá deteriorar-se drasticamente em aplicações do mundo real mais desafiantes. Tal facto deve-se às características obtidas a partir de câmaras convencionais que encontram inúmeras dificuldades em caracterizar os humanos se o fundo for dinâmico e complexo, ou necessitar de ultrapassar desafios como grandes variações nas poses humanas ou oclusões.

Devido à sua natureza descritiva, os dados visuais são a entrada preferida para a deteção humana (Wen *et al.*, 2021). No entanto, ainda são poucos os estudos que combinam a utilização de múltiplos. (Quintana *et al.*, 2016 e Wen *et al.*, 2021). Wen *et al.* (2021), abordam a conciliação entre câmaras *RGB* e GPS, e Alipio *et al.* (2020), a conciliação entre câmaras *RGB* e *UWB*. Quintana *et al.* (2016), também seguem essa linha de pensamento, na medida que, referem que a mistura de características de alto e baixo nível a partir de dispositivos portáteis de aquisições de sinais como *RFID* apresentaram resultados satisfatórios nos aspetos de interação entre ser humano e produto.

2.5. Tipologia de modelos

2.5.1. Deteção de objetos

Os métodos tradicionais tendem a detetar pessoas em dados extraídos de imagens *RGB* através de um histograma de gradientes orientados (*HOG*) (Choudhary & Chaudhury, 2016; Liu *et al.*, 2017). Por outro lado, Merad *et al.* (2016) e Priyanka *et al.* (2019), baseiam-se na ênfase que *SVM's* (*support vector machines*) fornecem uma maior usabilidade em aplicações em larga escala. No entanto, o segundo, de modo a complementar essa técnica utiliza o modelo *BoW* (*Bag of Words*) que pode ser aplicado à classificação de imagens ao tratar as características das mesmas como palavras.

Todavia, mais recentemente, foram desenvolvidas ferramentas de *deep learning* direcionadas para a detecção de objetos, especialmente utilizando redes neurais convolucionais (*CNN*) (Chen *et al.*, 2017; Islam & Baek, 2021; Paolanti *et al.*, 2020 e Quintana *et al.*, 2016), devido ao seu excelente desempenho relativo à precisão de detecção de um vasto conjunto de ambientes, mas também devido à sua velocidade de detecção.

Nguyen *et al.* (2022), referem no seu estudo que os modelos modernos de detecção de objetos que estão atualmente a ser implementados dividem-se em duas categorias: detetores de duas fases (*two-stage detectors*) e detetores de uma fase (*one-stage detector*). Na primeira são extraídas primeiro as regiões independentes da categoria e depois aplicada a classificação nas características (*deep features*) de cada região. São exemplos desses modelos as *Faster R-CNN* e *Mask R-CNN* (Bellini *et al.*, 2022; Hsia *et al.*, 2022 e Xia *et al.*, 2021). Por outro lado, os detetores de uma fase agrupam as duas fases numa só ao diretamente preverem as probabilidades das classes e caixas de delimitação a partir de imagens completas com uma única rede neuronal convolucional. Desta forma, este pode ser executado em tempo real e capta a imagem completa de uma só vez, ao contrário de analisar as propostas de uma região geradas nos métodos anteriores. Um exemplo de um modelo nesta categoria é o *Yolo* (*You only look once*) que, relativamente à velocidade de detecção, apresenta um tempo de detecção ainda mais rápido do que as *CNN* devido à sua arquitetura simplista (Wen *et al.*, 2021; Ghazal *et al.*, 2021 e Schwenzow *et al.*, 2021).

2.5.2. Localização e Análise de comportamentos

No contexto do retalho, os métodos existentes seguem, na maioria dos casos, o caminho de extrair as características espaciais de *frames* consecutivos e reconhecer o comportamento das características espaciais sequenciais através dos modelos, especialmente pelo *HMM* (*Hidden Markov Model*) (Liu *et al.*, 2017; Merad *et al.*, 2016 e Paolanti *et al.*, 2018).

Bellini *et al.* (2022), Liu *et al.* (2019), Paolanti *et al.* (2020) e Wen *et al.* (2021), usam *CNN*'s de modo a extrair características espaciais de cada *frame* de vídeo e reconhecer o comportamento por memória longa de curto prazo (*LSTM*) com as características espaciais. Em comparação com os recursos de análise de pixéis, Merad *et al.* (2016), utilizam *HMM* e máquinas vetoriais de suporte (*SVM*) de modo a analisar trajetórias de coordenadas para reconhecer e perceber os comportamentos.

Ghazal *et al.* (2021) e Liu *et al.* (2015), adotam um esquema de rastreamento simples, mas eficaz. Optam por utilizar um modelo básico de *Kalman Filter* e uma técnica de associação de dados, na qual cada rasto é atribuído a uma resposta no *frame* atual que é mais semelhante e mais próximo da posição prevista pelo modelo. Contudo, a dificuldade consiste em definir uma função de semelhança que tenha a capacidade de refletir da melhor forma as diferenças entre os diferentes seres humanos.

Adicionalmente, um dos estudos alvo de revisão, concebeu modelos baseados em regras autodefinidas através dos dados de *LBSN (Location-based Social Networking)* de modo a reconhecer o comportamento dos consumidores (Bellini *et al.*, 2022; Chen *et al.*, 2017 e Paolanti *et al.*, 2020).

Na Tabela 6 está apresentada a informação resumida relativa ao tipo de tecnologia utilizada em cada estudo e as suas técnicas e modelos, tanto de deteção de objetos como de localização. Adicionalmente, foram introduzidos vários fatores com vista à perceção da qualidade dos métodos de deteção como, se identificavam uma *ROI (Region of Interest)*, se procederam com a contagem de pessoas e, por fim, se analisavam interações entre humanos e outros objetos no espaço. Relativamente aos métodos de localização, foi também analisado se o estudo alvo procedeu só a essa análise comportamental dos clientes, ou se o foco também incidiu numa otimização do *layout* do espaço.

Tabela 6 - Tipologia de modelos de deteção e localização

ID	Tecnologia Wireless	Técnica Deteção	ROI/POI	PC	HOI	Técnica Localização	AnCC	Otimização Layout
1	RFID	Deep CNN	Sim	Sim	Sim	SSD	Sim	N.E
2	RFID	MLP/CNN/R-CNN	Sim	Sim	Sim	LSTM	Sim	N.E
3	RFID	N.E	Sim	Não	N.E	HMM	Sim	Sim
4	N.E	CNN	Sim	N.E	N.E	N.E	N.E	N.E
5	Câmara RGB	R-CNN/Yolo	Sim	N.E	Sim	N.E	N.E	N.E
6	Câmara RGB	R-CNN/Yolo	Sim	N.E	Sim	KFM/PF/KT	Sim	Sim
7	Câmara RGB-D	SVM	Sim	Sim	Sim	FP-Growth	Sim	N.E
8	WLAN RSS	N.E	N.E	N.E	N.E	Lee's Algorithm	Sim	Sim
9	Service Robot	SURF + Yolo	Sim	N.E	Sim	KFM/HMM	Sim	N.E
10	RFID	ANN/SVM	N.E	N.E	Sim	A.P	Sim	Sim
11	Câmara RGB	CNN	N.E	N.E	Sim	N.E	N.E	N.E
12	Câmara RGB	R-CNN	Sim	Sim	Não	N.E	Sim	N.E
13	Câmara RGB	Yolo	N.E	N.E	Sim	N.E	N.E	N.E
14	Câmara RGB/GPS	Yolo	N. E	Sim	Sim	HMM	Sim	N.E
15	N.E	R-CNN	N.E	N.E	Sim	SSD	Sim	N.E
16	UWB	N.E	N.E	N.E	Sim	SSD	Sim	Sim
17	Câmara RGB-D	Deep CNN	Sim	Sim	Sim	LSTM	N.E	N.E
18	Câmara RGB + UWB	CNN	N.E	Sim	N.E	SSD	Sim	N.E
19	Bluetooth	PALIA suite	Sim	N.E	N.E	PALIA suite	Sim	Sim
20	N.E	Deep CNN	Sim	N.E	Sim	LSTM/RNN	Sim	Sim
21	Câmara RGB	BoW +SURF +SVM	Sim	N.E	Sim	SVM	Sim	N.E
22	UWB	N. E	Sim	N. E	Sim	HMM	Sim	Sim
23	WLAN RSS	CNN	Sim	N. E	Sim	N.E	Sim	Sim

24	WLAN RSS	A.P	Sim	Sim	N.E	A.P	Sim	Sim
25	Câmara RGB	HOG/SVM	Sim	Sim	Sim	HMM	Sim	N.E
26	Câmara RGB	SIFT/HOG	Sim	N.E	N.E	KFM/HMM	Sim	Sim
27	Câmara RGB-D / RFID	CNN	Sim	Sim	Sim	SVM	Sim	N.E
28	RFID/SMI eye tracking glasses	SVM	N. E	N. E	Sim	HMM	Sim	N.E
29	Câmara RGB-D	HOG	Sim	Sim	N. E	KFM + SVM	Sim	Sim

Nota: N.E: Não especificado; AP: Algoritmo Próprio; AnCC: Análise Comportamento Consumidor; ANN: *Redes Neurais Artificiais*; BoW: *Bag of Words*; CNN: Rede Neuronal Convolutacional; HMM: *Hidden Markov Model*; HOI: Interação Humano/Objeto; KFM: *Kalman Filtering*; KT: *Kernel-based Tracking*; LSTM: *Long Short-term Memory*; MLP: *Multilayer Perceptron*; PC: Contagem Pessoas PF: *Particle Filtering*; RFID: *Radio Frequency Identification*; SIFT *Scale Invariant Feature Transform*; SSD: *Single Shot Detector*; SURF: *Speeded up Robust Features*; SVM: *Máquina de Vetores Suporte*; YOLO: *You Only Look Once*.

2.6. Avaliação de resultados

Com o propósito de analisar os métodos de avaliação dos resultados obtidos em cada um dos estudos, os seus contributos, mas, principalmente, perceber as limitações encontradas e as suas mitigações, tanto no âmbito do ambiente em análise como a privacidade dos dados, foram sujeitos a análise destas componentes os 29 artigos presentes nesta RSL como está sintetizado na Tabela 7.

Em relação aos testes, como forma de análise da qualidade dos resultados, estudos como os Merad *et al.* (2016) e Wen *et al.* (2021), optaram por proceder a uma reamostragem de novos conjuntos de dados ou alteração de variáveis. Por outro lado, também são utilizadas as matrizes de confusão e transição para analisar classificações de ações ou probabilidade de deslocações de objetos como nos estudos de Liu *et al.* (2017) e Paolanti *et al.* (2018), e métricas de avaliação de deteção objetos como a *accuracy (ACC)* em estudos como os de Kao *et al.* (2021) e Xia *et al.* (2021), e *Precision & Recall (PR Curve)* como nos estudos de Ali *et al.* (2022) e Ghazal *et al.* (2021). Salientar o uso de métricas fora do âmbito numérico como *heat maps*, por Carbajal (2021) e Nguyen *et al.* (2022), devido, tanto à sua componente de teste, como de mitigação de limitações. De referir que todos os artigos analisados apresentam contributos. Uma nova dimensão diz respeito à exploração de técnicas já existentes aplicadas a novos ambientes e o aprofundamento do âmbito tem em consideração os modelos existentes para criação de novas técnicas e formas de exploração dos dados.

Relativamente às limitações, as mesmas foram analisadas em três parâmetros distintos, as limitações referentes ao ambiente em análise, as limitações de privacidade e por fim as formas de mitigação desses constrangimentos. Em relação às limitações referentes à privacidade dos dados, foram vários os artigos que não especificaram qualquer limitação. Dos restantes foram identificados maioritariamente limitações com a análise facial, mitigadas pelo uso de imagens com esses membros desfocados ou mesmo com a remoção de fundos nos vídeos. De referir ainda que, os restantes estudos optaram pelo uso de aplicações *privacy-safe* ou obtiveram o consentimento da amostra.

No que toca às limitações do ambiente em análise, as mais identificada é a volatilidade dos dados, em cerca de um terço dos artigos alvo de análise. Essas características de volatilidade são apresentadas como imagens inconstantes, mudanças de iluminação e fundos (Chen *et al.*, 2017 e Paolanti *et al.*, 2020); mas também volatilidade nos comportamentos impulsivos dos consumidores (Merad *et al.*, 2016 e Paolanti *et al.*, 2018). A outra limitação principal de ambiente foi a oclusão, como apresentada em estudos como os de Quintana *et al.* (2016) e Xia *et al.* (2021), onde a mitigação dessa e similares limitações foi atingida através de sistemas adaptativos, como a conciliação entre câmara e *RFID* ou sistemas com múltiplas câmaras. De referir ainda, como referido anteriormente, estudos como os de Nguyen *et al.* (2022), apresentam soluções, tanto de avaliação de resultados, como de mitigações de limitações através da análise de *heat maps*.

Tabela 7 - Avaliação resultados dos artigos alvo de RSL

ID	Testes	Literatura	Contributos	Limitações Ambiente	Limitações Privacidade	Mitigação Limitações
1	<i>Precision and Recall (PR Curve), Miss Rates, Accuracy (ACC)</i>	N.E	Nova Dimensão de Análise	Ações limitadas; Volatilidade Dados	Consentimento da Amostra	Análise Comportamento Loja Online
2	Reamostragem; <i>Mean Absolute Error (MAE); Regularized Canonical Correlation Analysis (CCA)</i>	N.E	Nova Dimensão de Análise; Aprofundamento do Âmbito	N.E	Consentimento da Amostra	N.E
3	Average Tracking Error; Two-factor ANOVA	Sim	Nova Dimensão de Análise	Volatilidade Dados	N.E	Recalibração Modelo
4	<i>Precision and Recall (PR Curve), Accuracy (ACC) e Matriz Confusão</i>	N.E	Aprofundamento do Âmbito	Imagem	N.E	Recalibração Imagem
5	Matriz Confusão	Sim	Aprofundamento do Âmbito	Imagem	N.E	Recalibração Imagem
6	Heat Map	Sim	Aprofundamento do Âmbito	Volatilidade Dados; Oclusão e Imagem	Proteção Identidade	Limpeza de Dados e Re-identificação Anónima; Heat Map
7	N.E	N.E	Aprofundamento do Âmbito	Similaridades de Ações	N.E	N.E
8	<i>Elbow Inertial Method; Heat Map</i>	N.E	Nova Dimensão de Análise	N.E	Proteção Identidade	Heat Map

9	<i>Precision and Recall (PR Curve)</i>	Sim	Nova Dimensão de Análise	Volatilidade Dados; Oclusão	N.E	Recalibração e Eliminação Imagens
10	<i>Detection Error Tradeoff (DET)</i>	N.E	Nova Dimensão de Análise	Volatilidade Dados	Aplicação <i>Privacy-Safe</i>	N.E
11	<i>Mean Absolute Error (MAE); Cumulative Score (CS); Accuracy (ACC); Precision (PPV) & Recall (TPR)</i>	Sim	Nova Dimensão de Análise; Aprofundamento do Âmbito	Dados Fonte	Consentimento da Amostra	Tratamento de Imagens (<i>Data Augmentation</i>)
12	<i>Accuracy (ACC)</i>	N.E	Aprofundamento do Âmbito	N.E	N.E	N.E
13	Alteração Variáveis independentes; R ² ; P-Value	N.E	Nova Dimensão de Análise; Aprofundamento do Âmbito	Imagem	N.E	N.E
14	Reamostragem; Simulação Ambientes	Sim	Nova Dimensão de Análise; Aprofundamento do Âmbito	Dados Código Fonte; <i>Fish-eye view</i> (Imagem)	Análise de Faces	Faces Desfocadas; Sistema Adaptativo
15	<i>Error & Accuracy</i>	N.E	Aprofundamento do Âmbito	Volatilidade Dados; Oclusão	N.E	Sistema Adaptativo
16	Métricas de Avaliação Segmentação (IoU); Reamostragem	N.E	Aprofundamento do Âmbito	Volatilidade Dados; Imagem; Câmaras	N.E	Sistema Adaptativo
17	Outros métodos Segmentação e Classificação; Reamostragem	Sim	Nova Dimensão de Análise; Aprofundamento do âmbito	Volatilidade Dados	Análise de Faces	Faces Desfocadas; Câmaras <i>RGB-D</i>
18	Reamostragem	N.E	Nova Dimensão de Análise	Volatilidade Dados	N.E	Sistema Adaptativo
19	Reamostragem	Sim	Nova Dimensão de Análise	Oclusão; Similaridades de Ações	N.E	Sistema Adaptativo
20	<i>Error & Accuracy</i>	N.E	Nova Dimensão de Análise; Aprofundamento do Âmbito	Dados Fonte; Volatilidade Dados	Proteção Identidade	Análise Comportamento Loja Online; Anonimização Cliente
21	Matriz Confusão	N.E	Nova Dimensão de Análise; Aprofundamento do Âmbito	N.E	Proteção Identidade	Tratamento de Imagens como Palavras (<i>BoW model</i>)
22	Matriz Confusão e Matriz de Transição; Heat Map	Sim	Nova Dimensão de Análise	Volatilidade Dados	Aplicação <i>Privacy-Safe</i>	<i>Heat Map</i>
23	Métricas de Avaliação de OD & RE, Accuracy	Sim	Nova Dimensão de Análise; Aprofundamento do Âmbito	Volatilidade Dados	Aplicação <i>Privacy-Safe</i>	Análise Comportamento Loja Online
24	Reamostragem	Sim	Aprofundamento do Âmbito	N.E	Aplicação <i>Privacy-Safe</i>	N.E
25	Reamostragem e Matriz Confusão	Sim	Aprofundamento do Âmbito	Similaridades de Ações	Análise de Faces	Faces Desfocadas; N.E (Ambiente)
26	N.E	N.E	Aprofundamento do Âmbito	Oclusão; Volatilidade Dados; Similaridade de Ações	Proteção Identidade	Sistema Multi-câmara; Faces Desfocadas e Vídeos Privacidade Aumentada
27	N.E	N.E	Aprofundamento do Âmbito	Oclusão; Imagem	N.E	Sistema Adaptativo; N.E (Privacidade)
28	Alteração Variáveis independentes	N.E	Nova Dimensão de Análise	Volatilidade Dados	Consentimento da Amostra	Segmentação Padrões Clientes
29	Reamostragem; Métricas de OD e RE, Accuracy	Sim	Nova Dimensão de Análise	Dados Código Fonte; Oclusão; Imagem	N.E	Sistema Multi-câmara + Câmaras <i>RGB-D</i> ; N.E (Privacidade)

Nota: N.E: Não especificado; OD: Deteção Objetos; RE: Eficácia de Recomendação

2.7. Avaliação artigos científicos relevantes

A Tabela 8 tem como objetivo apresentar a avaliação dos 29 artigos presentes na RSL. Pelo facto de ser uma revisão sistemática da literatura, o trabalho de investigadores poderá ser facilitado, na medida em que para cada um dos artigos presentes nesta revisão é avaliada a sua relevância em cada um dos aspetos presentes na Tabela 5. A cada artigo será atribuída uma nota para cada questão de avaliação de qualidade onde: 0 se não responder a aspeto nenhum da questão; 0,5 se responder parcialmente e 1 se responder totalmente.

Tabela 8 - Avaliação da qualidade dos artigos sujeitos a RSL

ID	CA		Características e comportamentos				Tecnologia		Modelos <i>deep learning</i>						Limitações		Avaliação resultados			T
	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15	Q16	Q17	Q18	Q19	
1	1	1	1	0,5	1	1	1	1	1	0,5	1	0	0	0	1	1	1	0	1	14
2	1	0,5	1	0,5	0	1	0,5	0,5	1	1	1	1	0,5	0	0,5	0	1	0	0,5	11,5
3	1	1	1	1	0	1	1	1	0	0	0,5	0,5	0	1	0,5	0	1	1	0,5	12
4	0	0	0,5	0	0	0	0	0	1	1	0	0	0	0	0,5	0,5	1	0	1	5,5
5	0,5	0	0,5	0	1	0	0,5	0	1	1	0,5	0	0	0	0,5	0,5	1	1	0,5	8,5
6	0	1	1	1	1	1	0	0	1	1	0,5	1	1	1	1	1	1	1	0,5	15
7	1	1	0,5	0,5	0,5	0,5	1	1	1	0,5	1	1	1	0	0,5	0	0	0	0,5	11,5
8	1	0,5	0	0	0	1	1	0,5	0	0	0	1	0,5	1	0,5	0,5	1	0	1	9,5
9	0,5	0	1	1	1	0,5	1	1	1	1	0,5	1	0,5	0	0,5	0,5	1	1	1	14
10	0,5	0,5	1	1	1	1	0,5	1	1	1	0	0,5	0,5	1	1	0	1	0	1	13,5
11	0,5	0,5	1	0,5	0,5	0	0,5	0	1	0,5	0	0	0	0	1	1	1	1	1	10
12	1	0,5	1	1	1	0	0,5	0	1	0,5	1	0,5	0	0	0	0	1	0	1	10
13	1	0	0	0	1	0	0,5	0	1	1	0	0	0	0	0,5	0	1	0	1	7
14	1	1	0	0,5	1	1	1	1	1	0	0,5	0,5	0,5	1	1	1	1	1	1	15
15	0,5	0	1	1	1	1	1	1	1	1	0	0	0	0	0,5	0,5	1	0	1	11,5
16	1	0,5	0,5	0,5	0,5	1	1	1	1	1	0	1	1	0	0,5	0,5	1	0	1	13
17	1	1	1	0,5	1	0,5	1	0,5	1	1	1	0,5	0,5	0	1	1	1	1	1	15,5
18	0,5	0,5	1	1	0	0,5	0,5	1	1	0,5	0,5	0	0	0	0,5	0,5	0,5	0	0,5	9
19	1	1	0,5	0,5	0	1	1	1	1	0	0	1	1	1	0,5	0,5	0,5	1	0,5	13
20	1	1	1	0,5	1	1	0	0	1	1	0,5	1	1	1	1	1	1	0	1	15
21	1	0,5	1	0	1	0	0	0	1	1	0,5	0,5	0	0	0,5	0,5	1	0	0	8,5
22	1	1	0	0	1	1	1	0	1	0	0,5	1	0	1	1	1	1	1	1	13,5
23	1	0,5	1	0,5	0,5	0,5	0	0	0	0,5	0,5	1	1	1	1	1	1	1	1	13
24	1	1	1	0	1	0,5	1	1	0,5	0	0,5	0,5	0	1	0,5	0,5	0,5	1	0,5	12
25	1	1	1	0	1	1	0	0,5	1	0,5	1	1	1	0	1	0,5	0,5	1	1	14
26	0,5	0	0,5	0	1	0,5	0,5	0	1	0,5	0,5	1	0,5	0	1	1	0	0	0	8,5
27	1	0,5	1	0,5	1	1	1	1	1	1	1	1	1	0	0,5	0,5	0	1	0,5	14,5
28	1	0,5	1	1	0,5	0	0,5	1	1	0	0	1	0,5	0	1	1	0	0	1	11
29	0,5	1	0	0	0	1	1	0,5	1	0,5	0,5	0,5	0,5	1	0,5	0,5	1	1	0,5	11,5
T	23	17,5	21	13,5	19,5	18,5	18,5	15,5	25,5	17,5	13,5	18	12,5	11	20	16,5	23	14	22	

No que toca à análise das informações das avaliações a que os artigos foram sujeitos nesta revisão sistemática da literatura, é possível concluir que sete dos vinte e nove artigos obtiveram uma pontuação relativamente alta em termos de qualidade, com catorze ou mais pontos em dezanove (3/4 da pontuação), sendo o melhor estudo avaliado o de Paolanti *et al.* (2020), com quinze pontos e meio. De referir ainda que, apesar de terem os seus contributos para a investigação, no âmbito geral, dois estudos não satisfizeram muitos dos parâmetros de avaliação. A investigação de Schwenzow *et al.* (2021), apenas pontuou sete em dezanove, devido à sua utilidade em domínio teórico mas escassez na componente prática. O estudo com pior classificação, o de Domingo *et al.* (2022), com cinco pontos e meio, apresentava quase as características opostas, apresentando apenas os elementos práticos e somente de técnicas de deteção de objetos.

Estudos como os de Carbajal (2021) e Chen *et al.* (2017), são dos poucos que não analisam técnicas de deteção de objetos (Q9 a Q11) e seis artigos não referem quaisquer técnicas de análise do comportamento dos consumidores (Q12 a Q14). Adicionalmente quase todos os artigos presentes a esta RSL descrevem, pelo menos parcialmente, as limitações dos seus estudos (Q15) e direcionam o seu estudo para a compreensão do consumidor como fator chave (Q1).

No lado oposto, é possível concluir que, dos artigos avaliados, apenas cerca de metade, enaltece evidências em analisar especificamente formas de otimizar o *layout* do espaço (Q14) e que muitos estudos se focam apenas em perceber o comportamento dos consumidores, abordando pouco as potencialidades de lojas sem caixas no futuro (Q3).

2.8. Contributos da RSL

Em forma de considerações finais, a revisão sistemática de literatura efetuada forneceu diversos contributos para a elaboração da dissertação em que está inserida. Desta forma, através da mesma foi possível definir diversas fases da metodologia, como entender o tipo de dados analisados e técnicas de preparação dos mesmos. Posteriormente, ainda incidindo na metodologia, perceber que modelos são os mais indicados para a deteção de objetos e análise de comportamentos humanos, mas também como podem ser avaliados os resultados obtidos.

Através desta revisão foi também possível entender a atualidade do tema e dos conceitos e a sua evolução ao longo dos anos, mas também ter a perceção das limitações com que cada estudo se deparou, pois, possivelmente as mesmas poderão ocorrer neste estudo, e por isso também compreender como podem ser mitigadas.

3. Metodologia

3.1 Empatia e Perceção

A proposta diferenciadora do *Design Thinking* é que as ideias de propostas sejam geradas em conjunto com as pessoas que serão impactadas por elas, mas também que os modelos sejam construídos e testados ainda durante o processo (Pande & Bharathi, 2020).

Foi elaborado um questionário publicado entre os dias de 5 outubro 2022 a 1 de novembro 2022, sendo 221 respostas recolhidas e analisadas. O questionário foi criado em *Qualtrics*, um *software* de inquéritos disponível a utilizadores ligados ao ISCTE para apoio ao ensino e investigação. O mesmo foi partilhado via email e através de redes sociais. Os detalhes sobre o questionário estão presentes no Anexo A.

Como referido anteriormente, a amostra é constituída por 221 indivíduos, sendo que 177 responderam ao questionário na sua totalidade. Em relação à faixa etária (Gráfico 2), os grupos que predominam no estudo são os correspondentes às faixas etárias “entre 30 e 59 anos” e “entre 18 e 29 anos”, com 89 indivíduos (50%) e 70 indivíduos (40%) respetivamente, e onde apenas foi registada uma resposta de um indivíduo menor de idade. Em termos de rendimento mensal (Gráfico 3), o intervalo dominante é o de “dois a quatro salários mínimos nacionais” com 62% a se inserirem nesse escalão. Estas características da amostra poderão ser justificadas pelos métodos de partilha do questionário, mas que se adequam perfeitamente ao estilo de consumidor em que se pretende criar empatia.

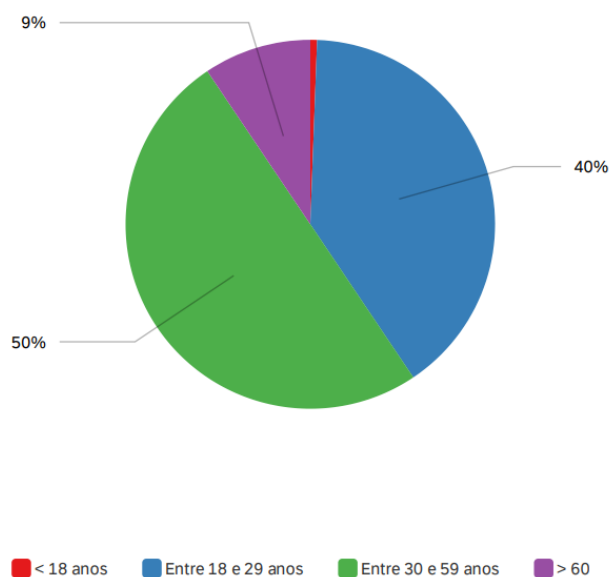


Gráfico 2 - Faixa etária dos inquiridos

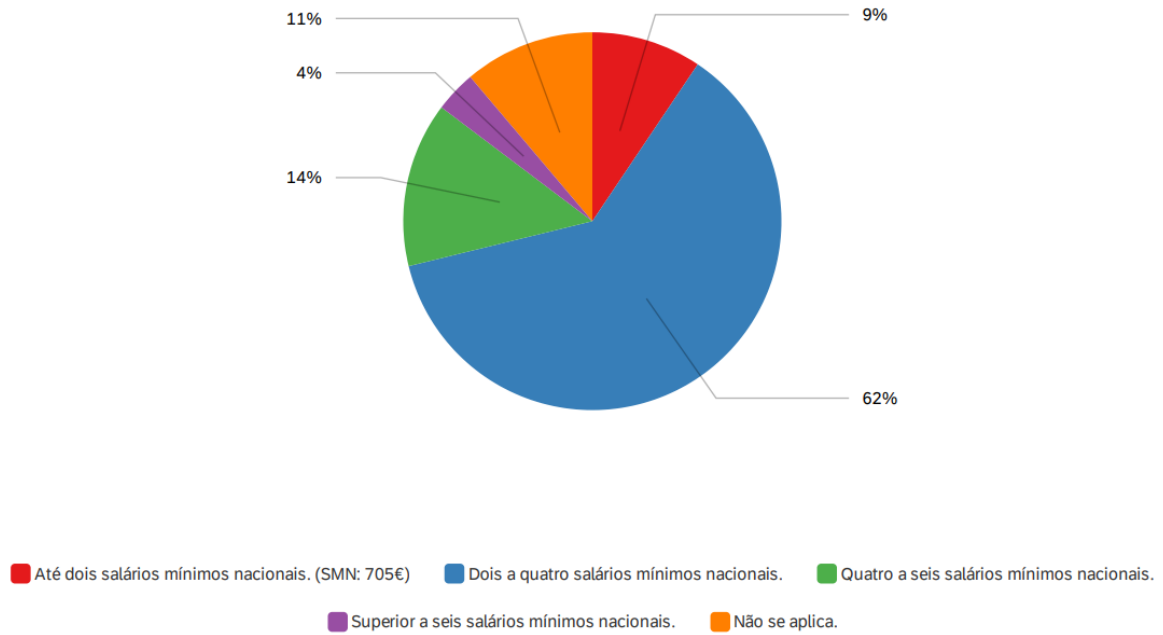


Gráfico 3 - Rendimento mensal dos inquiridos

Devido ao presente estudo ter por objetivo a utilização de técnicas e métodos de *Big Data Analytics* na área de reconhecimento de imagem para melhorar a experiência do consumidor na área do retalho, pretende-se com o questionário criar empatia com os consumidores de modo a perceber as suas reais necessidades relativas ao ambiente em análise.

Posto isto, traduziu-se as necessidades e desejos dos clientes em propostas de valor concretas, através de uma análise descritiva de resultados provenientes do questionário elaborado.

O mesmo foi dividido em dois grandes grupos: a perceção em relação aos conceitos e processos de compra de sistemas de *checkout* autónomo e posteriormente a perceção dos processos para todos os modelos de supermercado (sem caixas, com caixas e híbrido).

As primeiras questões foram formuladas no sentido de perceber o conhecimento dos inquiridos acerca dos formatos de loja sem caixas e posteriormente as suas opiniões em algumas das características das mesmas. De acordo com o Gráfico 4, a resposta dominante à pergunta: “Antes deste questionário, já tinha ouvido falar ou visitado um formato de loja disruptiva?” é que “Apenas ouvi falar” com 62% de inquiridos a seleccioná-la, e em que apenas 12% da amostra nunca ouviu falar deste formato de lojas. De referir ainda que, sensivelmente, um quarto da amostra já ouviu falar e visitou pelo menos um dos estabelecimentos referidos.

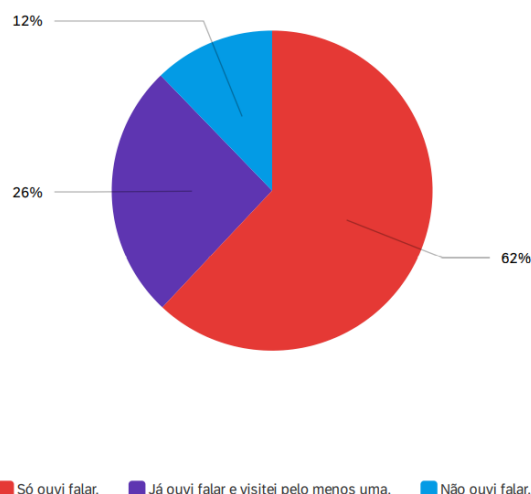


Gráfico 4 - Familiaridade com o conceito de loja sem caixas

De seguida foram inquiridas as principais razões que levariam a uma eventual visita a este tipo de lojas, e que constrangimentos imaginariam ultrapassar. Estudos como de Gazzola *et al.* (2022) e Ives *et al.* (2019), enalteciam como principais características: experimentar novas tecnologias; menor tempo dentro da loja; facilidade de pagamento e, por fim, não perder tempo em filas de espera. Deste modo, como visto no Gráfico 5, foram dadas essas opções de resposta à questão levantada, assim como, uma opção para quem não pretendesse visitar lojas deste conceito. Assim sendo, perto de metade dos inquiridos, 46%, imaginaria uma redução de tempos em fila de espera, seguido de uma expectativa de um menor tempo dentro da loja, com cerca de 20% das respostas. De realçar o carácter inovador do público sondado, com um maior peso do mesmo a favorecer a experimentação de novas tecnologias a uma maior facilidade de pagamento.

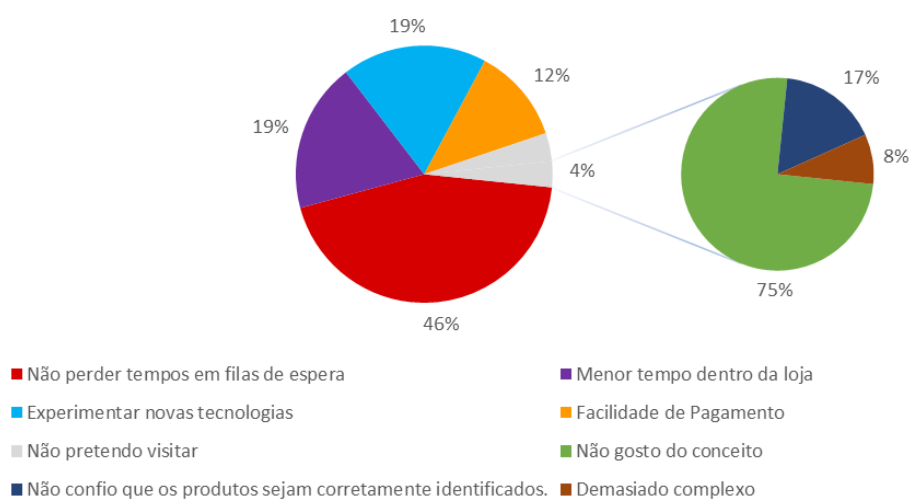


Gráfico 5 - Motivação para visitar lojas deste conceito

Dos 4% que responderam não pretender visitar uma loja, 69% justificou apenas por não apreciar o conceito intrínseco.

O seguinte sub-grupo de questões estava assente na explicação do processo de compra neste novo modelo de supermercados e a percepção dos inquiridos acerca do mesmo. Em relação a todas as fases do processo (Registo; Acesso à loja; Atividades dentro da loja; Saída da loja e Pós-compra), as classes dominantes em todas as fases são sempre “Muito Acessível” ou “Acessível”. Contudo, é importante constatar, no Gráfico 6, a descida de sensação de acessibilidade do processo, das fases de registo e acesso à loja, para o processo de compra em si, onde cerca de um quarto dos inquiridos apenas acha razoável essas fases. De referir ainda, o pós-compra a ser a fase com maior discordância entre classes e, por outro lado, o acesso e saída da loja, com os inquiridos a reconhecerem a sua praticabilidade.

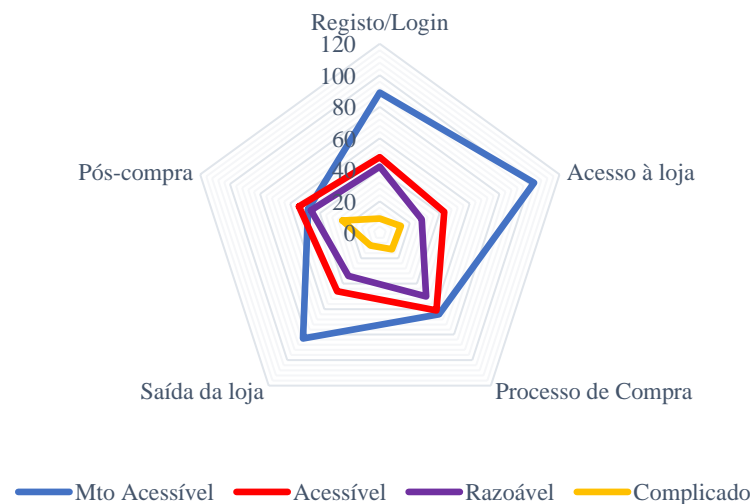


Gráfico 6 - Percepção do processo de compra

Ainda incidindo sobre o processo de compra, e analisando as informações referente ao Gráfico 7, 75% sente que o processo de compra transmite confiança/facilidade ao consumidor, e em mais de metade (60%) dos inquiridos é transmitido que o cliente terá pleno controlo no processo. Todavia, o receio de que os produtos possam não ser corretamente identificados no carrinho de compras é expressivo, com 49% a ter esse fator em conta, para 45%, um menor peso na amostra, a não ter esse receio.

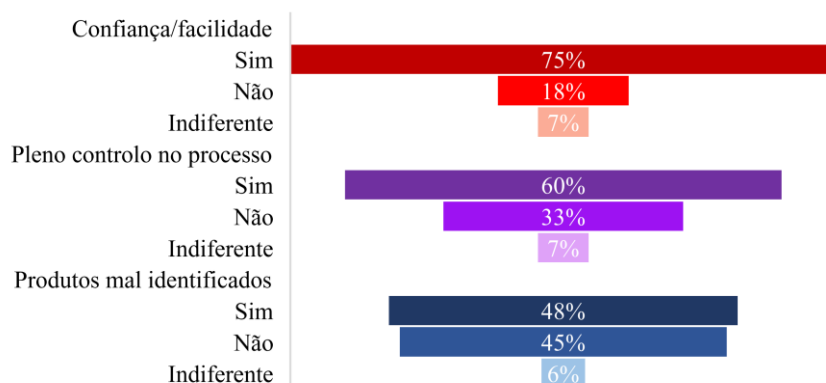


Gráfico 7 - Características do processo de compra

O segundo grande grupo de questões, identificado no Gráfico 8, está relacionado com a experiência dos consumidores em supermercados em formato tradicional. Relativamente às rotas percorridas dentro das lojas, os inquiridos foram questionados se teriam uma rota fluída, onde não se sentiam perdidos dentro da loja; se não se sentiam incomodados pelas rotas que tomavam, e por último se tinham a perceção que percorriam muito espaço da loja à procura do que pretendiam. De realçar que, neste ponto, ainda 75% não considera ter uma rota totalmente intuitiva, e dessa percentagem, 27% chega mesmo a considerar a sua rota totalmente confusa, um peso percentual que chega mesmo a ser superior aos inquiridos que sentem as suas rotas fluídas.

No que toca à atual disposição dos produtos, os resultados são mais polarizados, com pouco mais de um quarto dos inquiridos a ser indiferente à disposição dos produtos nos estabelecimentos.

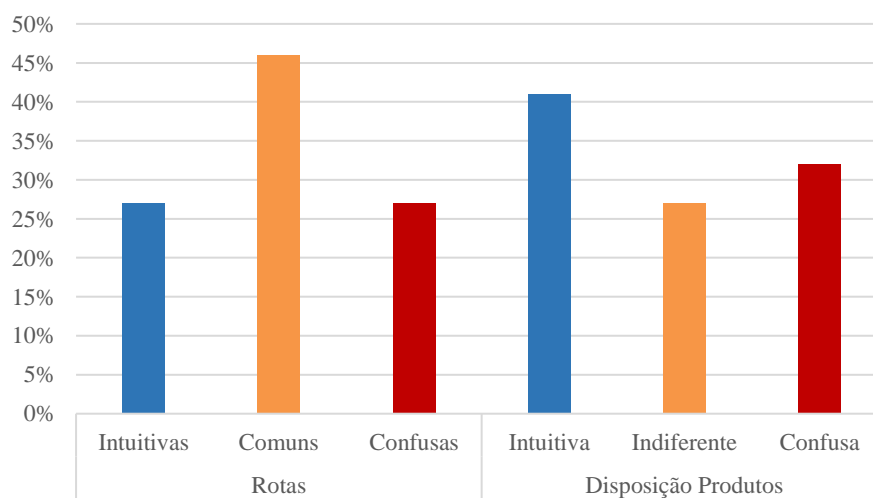


Gráfico 8 - Experiência nos estabelecimentos comerciais

O Gráfico 9 apresenta os resultados à questão relativa à vantagem de as empresas terem em consideração o *layout* das lojas de modo a que o tempo dentro das mesmas seja reduzido, A mesma pode ser considerada umas das questões com resultados mais unilaterais, com 88% dos inquiridos com a percepção que seria vantajoso.

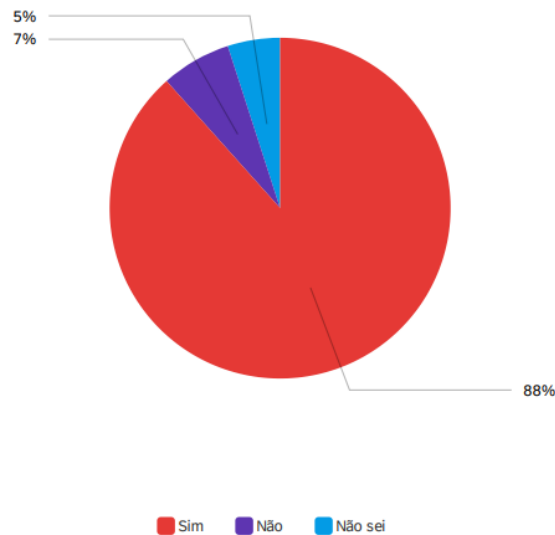


Gráfico 9 - Posição na consideração do *layout* das lojas

De referir ainda, o Gráfico 10 traduz que, apesar de mais de 40% dos inquiridos terem a opinião que os estabelecimentos comerciais têm uma disposição dos produtos intuitiva, e 27% achar que as suas rotas são fluidas, ainda 86% e 91%, dos grupos respetivamente, refletem na vantagem adicional de considerar do *layout* das lojas para um menor tempo dispendido.

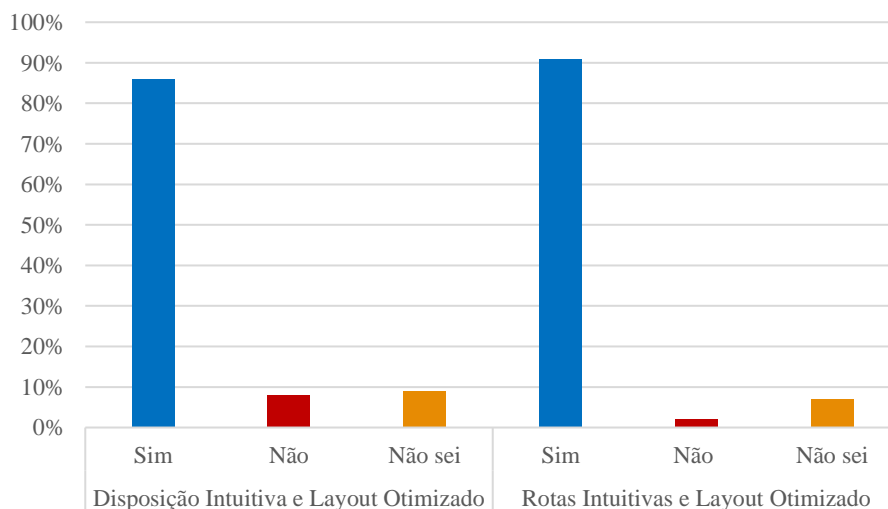


Gráfico 10 - Relação de rotas e disposição na otimização do *layout*

Por fim, foi requerido de forma interativa, que fossem identificadas na figura de um *layout* de um supermercado nacional retirada do estudo de Madureira., 2021, as secções mais regularmente visitadas nas deslocações ao supermercado. Os resultados foram apresentados em forma de *heatmap* como demonstra na Figura 7. Posto isto, é evidenciado a elevada afluência a zonas como frutas, mercearias e congelados, mas também à padaria e aos produtos de higiene e beleza.

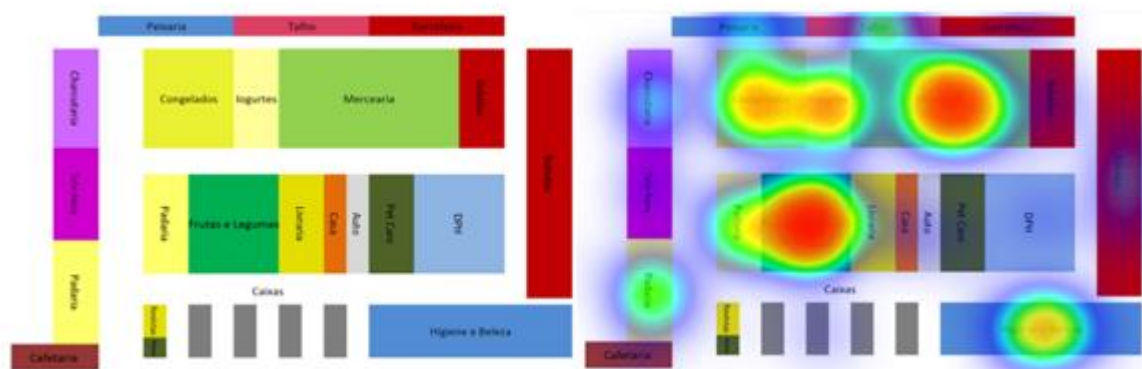


Figura 7 - Heat map de afluência nas secções do estabelecimento

3.2 Definição

Nesta fase, teve-se por base a empatia e estratégias criativas do *Design Thinking* de modo a definir da melhor forma o problema e as soluções iniciais.

Ao longo da fase anterior, foi-se desmontando e sintetizando os dados recolhidos, de modo a reconstruí-los e na presente etapa definir informações pertinentes para a avaliação dos resultados e percepção que a investigação, pode estar numa fase final.

Assim sendo, apesar da fundamentação na literatura das deslocações aos supermercados não serem práticas desejadas, e a confirmação desse mesmo facto pela amostra inquirida, este tipo de conceitos de lojas parece ter as bases para colmatar esses efeitos. Percebeu-se que a amostra inquirida tem confiança nas novas tecnologias e nas possíveis técnicas subjacentes às mesmas, como forma de ultrapassar os seus constrangimentos. Adicionalmente, a sensação das rotas serem confusas ainda é considerável e é quase unânime as vantagens, na ótica do consumidor, do *layout* das lojas ser otimizado para um menor tempo dentro da loja.

Desta forma, para tal efeito, é necessário perceber onde se encontram as maiores aglomerações dentro das mesmas para que depois seja possível fazer sugestões de medidas que possam ser tomadas. A concretização de tal objetivo, pode ter como implicações, melhorar as disposições dos produtos numa ótica de otimização de tempo, aumentar a rotatividade da loja, e melhorar a experiência do consumidor. Posto isto, o objetivo principal número três (O3) foi reformulado, onde foi delimitado mais um objetivo concreto que pretende avaliar padrões de consumo que possam explicar aglomerações nas lojas. Foi criado o objetivo O3.3, onde a plenitude dos mesmos se evidencia como:

- ✓ O1: Avaliar técnicas de recolha, tratamento e processamento de imagens;
- ✓ O2.2: Avaliar diferentes técnicas de deep learning de deteção de objetos em relação à base de dados disponível;
- ✓ O2.2: Criar um modelo de deteção de objetos parametrizado para identificar pessoas;
- ✓ O3.1: Caracterizar a perceção dos consumidores para ambientes de retalho autónomo, mas também tradicionais;
- ✓ O.3.2: Criar um modelo de localização e análise de comportamento dos consumidores;
- ✓ O.3.3: Criar métricas de avaliação de padrões de consumo que possam identificar e explicar aglomerações nas lojas.

3.3 Compreensão dos dados

Na fase de compreensão dos dados, que se poderá caracterizar como uma das fases fulcrais do projeto, a rigidez do CRISP-DM nas fases relativas aos dados preencheria os requisitos necessários para uma boa estrutura de recolha e análise de informação.

Desta forma, nesta fase foram recolhidos os dados provenientes das câmaras de um estabelecimento comercial português, e avaliada a sua qualidade. Esse estabelecimento tem incorporadas dezenas de câmaras, sendo necessário identificar aquelas que sustentam maiores benefícios para a investigação.

Para o efeito de avaliação da abrangência de imagem que cada câmara tinha ao seu dispor, foi efetuada uma breve análise sobre as mesmas e decididas quais, em conjunto, poderiam trazer uma maior capacidade de captar uma maior extensão do espaço, mas também todos os percursos dentro do estabelecimento.

Foram recolhidos conjuntos de quinze minutos de vídeo de oito câmaras distintas, com o objetivo de analisar certas zonas específicas da loja, identificadas por especialistas da loja e do seu ambiente. Foi identificado que em duas das câmaras o vídeo ficou cortado em dois ficheiros distintos, sendo necessário tratá-lo na fase de preparação dos dados.

Posteriormente, os mesmos dados foram divididos em *notebooks* diferenciados de modo a se proceder, numa fase inicial, ao tratamento individualizado de cada vídeo, para numa fase posterior ser possível aglomerar a informação referente aos mesmos. Na Tabela 9, estão presentes *frames* de todos os vídeos, assim como, uma breve descrição dos mesmos e identificação de exemplos de artigos presentes nesses espaços. Os *frames* de todas as câmaras estão apresentados no Anexo B , para melhor compreensão dos mesmos. Adicionalmente, um exemplo de *frame* proveniente da câmara número um está identificado na Figura 8.

Tabela 9 - Descrição da localização e secção da loja das diferentes câmaras

Câmara	Frame	Espaço	Exemplos Secções/Artigos
1	<i>Anexo 10</i>	Corredor nº1	Papelaria; Bebidas; Higiene e Beleza
2	<i>Anexo 11</i>	Corredor nº2	Cereais; Doces; mercearia
3	<i>Anexo 12</i>	Corredor nº3	Mercearia; Iogurtes; Bebidas Frescas; Congelados; Granel
4	<i>Anexo 13</i>	Congelados	Congelados; Higiene e Beleza; Bio
5	<i>Anexo 14</i>	Entradas Corredores	Granel; Iogurtes; Bebidas Frescas; mercearia
6	<i>Anexo 15</i>	Caixas <i>Self-checkout</i>	Frutas e Legumes; Padaria
7	<i>Anexo 16</i>	Padaria	Padaria; Cafeteria; <i>Take Away</i> ; Saída
8	<i>Anexo 17</i>	Espaço Amplo	Padaria; Frutas e Legumes; Caixas <i>Self-checkout</i> ; Entrada; Saída

Nota: As imagens na presente tabela já são *outputs* posteriores à fase de preparação dos dados devido a precauções com a privacidade dos dados.

3.4 Preparação dos dados

Na sequência da compreensão dos dados provenientes do estabelecimento comercial em causa, e sua devida rotulagem, procedeu-se à preparação dos mesmos. A fase de preparação dos dados abrangeu o conjunto de procedimentos necessários de modo a contruir o/os *datasets* finais, provenientes das bases de dados iniciais, com o propósito de ter os melhores dados possíveis para as fases seguintes da metodologia.

Para os modelos utilizados, melhor qualidade de dados de treino e teste disponíveis conduzem geralmente a modelos mais precisos. Para decidir que alterações nos dados são eficazes, foram realizadas experiências, treinando um modelo em diferentes conjuntos de dados correspondentes à utilização de diferentes alterações. As alterações que se mostrarem eficazes para aumentar a precisão do modelo serão utilizados nos modelos finais. As alterações testadas foram as seguintes operações:

- Redimensionamento Imagem e Anotações;
- Brilho;
- Cruzamento de dados de várias câmaras;
- Agregação dos ficheiros de vídeo provenientes da mesma câmara.

Posteriormente, e relativo a tópicos como a privacidade dos dados tanto pessoais como do ambiente em questão, foram efetuados tratamentos adicionais aos dados. Desta forma, de modo ao ambiente de retalho em análise permanecer confidencial, foi removido o fundo do mesmo, como demonstra o exemplo da Figura 8. Por conseguinte, os objetos em análise (consumidor no interior da loja), foram identificados através do contorno das suas formas.

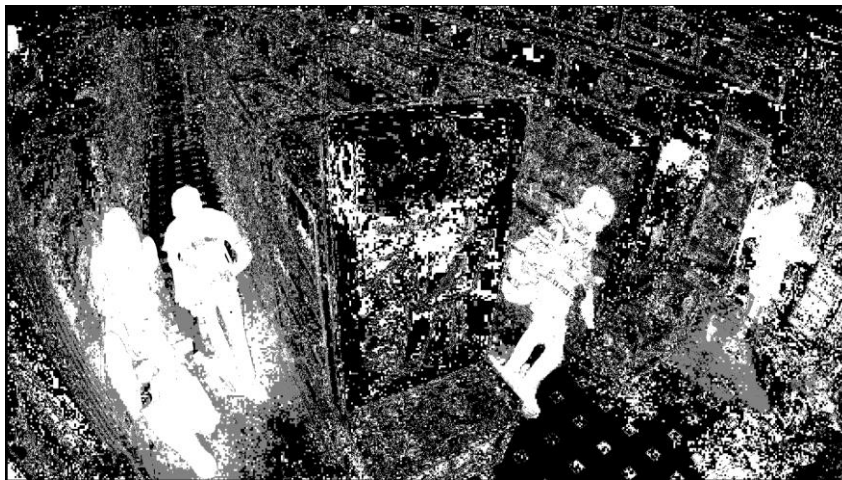


Figura 8 - Frame câmara nº1

A partir de cada vídeo foram gerados milhares de *frames*, em média 9010, passíveis de análise, e ativada uma figura do espaço existente num dos frames do vídeo de modo a representar o layout do espaço, como exemplificado na Figura 8.

3.5 Idealização e Modelação

A quinta e sexta fases da metodologia, consistiram na idealização, na seleção e na aplicação de diversas técnicas de modelação. Nestas fases, promoveu-se a experimentação das técnicas de *deep learning* tanto aplicadas à deteção de objetos como análise de comportamentos dos consumidores extraídos da literatura como as mais relevantes e com melhores resultados para cada um dos objetivos deste estudo. Os modelos foram desenvolvidos na linguagem de programação *Python*.

3.5.1 Deteção de objetos

A deteção de objetos é a tarefa de prever automaticamente a presença de certos objetos numa dada imagem de entrada. A previsão deve produzir a localização e classe de objeto dos então detetados. Esta informação é geralmente incorporada numa caixa de delimitação ou numa máscara de precisão de *pixels* com coordenadas na imagem, juntamente com a classe de objeto correspondente, devido à compreensão dos dados e sua correspondente preparação nas fases anteriores.

As redes neuronais apresentaram um desempenho favorável relativo à precisão de deteção de um vasto conjunto de ambientes, mas também devido à sua velocidade de deteção (Chen *et al.*, 2017; Paolanti *et al.*, 2020 e Quintana *et al.*, 2016). O *Yolo*, por outro lado, pode ser executado em tempo real e capta a imagem completa de uma só vez, ao contrário de analisar as propostas de uma região geradas nos métodos anteriores. Este algoritmo apresenta um tempo de deteção ainda mais rápido que as *CNN* devido à sua arquitetura simplista (Wen *et al.*, 2021).

Desta forma, de acordo com a interpretação da literatura para este estudo, *Yolo* é o algoritmo escolhido devido à sua presença em estudos modernos, tais como os de Ghazal *et al.* (2021), Nguyen *et al.* (2022) e Wen *et al.* (2021). Além disso, é considerado o facto de que estes tipos de modelos (detetores de uma fase) têm a vantagem de prever as probabilidades de classes e caixas de delimitação a partir de imagens completas com uma única rede neuronal convolucional em oposição a detetores de duas fases que analisam partes de uma imagem de cada vez.

3.5.2 Comportamento dos consumidores

Neste estudo, foram utilizados *HMM's* (*Hidden Markov Models*) com o objetivo de elaborar um sistema inteligente que analisa as trajetórias dos clientes num ambiente retalhista.

Ao utilizar a conciliação de caminhos combinados, o algoritmo proposto percorre uma variedade de caminhos (*clusters*) de modo a construir um mapa de localização e trajetórias otimizado com as características pretendidas.

Devido à estrutura do modelo, as probabilidades de escolha individual estão correlacionadas através de um caminho subjacente comum dos estados ocultos. Consequentemente, as probabilidades conjuntas de uma sequência de escolhas são dadas pela soma de todas as rotas possíveis que o indivíduo poderia tomar no tempo entre estados subjacentes (Merad *et al.*, 2016).

3.5.3 Bibliotecas

Em relação às bibliotecas utilizadas para preparação das imagens e posterior modelação, de destacar a importância de duas em específico.

Uma das bibliotecas foi a *OpenCV*, que é caracterizada por ser uma imensa biblioteca *open-source* que se centra maioritariamente em técnicas de *computer vision* e processamento de imagens, com funcionalidades como deteção de objetos ou deteção de faces. Adicionalmente, ao recorrer a esta biblioteca é também possível, ao analisar o vídeo, estimar o movimento no mesmo, subtrair o fundo e localizar objetos.

Foi também utilizada a biblioteca do *Python*, *ImageAI*, uma biblioteca que auxilia estudantes, investigadores e programadores a contruir aplicações com capacidade de *deep learning* e *computer vision*. A partir destas bibliotecas, o modelo foi programado para identificar pessoas a partir de uma probabilidade de confiança de acerto de 75% devido à qualidade de imagem dos vídeos.

Desta forma, cada imagem nos dados foi etiquetada com a localização, o tamanho e a classe de objeto para todas as instâncias de objetos na imagem correspondente, sendo que apenas se pretende detetar objetos como pessoas. As máscaras podem ser facilmente convertidas em etiquetas de caixas de delimitação, como demonstra a Figura 9.

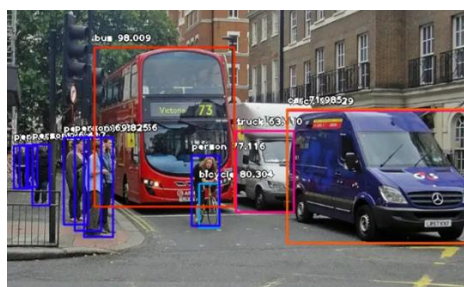


Figura 9 - Caixas de delimitação no reconhecimento de objetos (Olafenwa, M. & Olafenwa, J., 2021)

Após esse reconhecimento, foram geradas *track masks* por cada objeto/pessoa detetado/a de modo a enaltecer padrões de consumo e afluências dentro da loja. Posto isto, foi efetuado o cálculo das frequências e conseqüente mapa, onde as rotas que os clientes mais tomam podem emergir e as secções mais visitadas são realçadas. A Figura 10 demonstra um *frame* antes e após ser efetuado esse cálculo das frequências e construído o seu mapa, onde cada retângulo amarelo representa uma pessoa detetada num dos *frames* do vídeo total.



Figura 10 - Exemplo de um *frame* alterado para mapa de frequências

3.6 Teste e Avaliação

Nesta fase, numa combinação de ambas as metodologias, encontrou-se a vantagem em focar em elementos mais técnicos para encontrar soluções finais resultante de características da fase de avaliação do CRISP-DM, com a experimentação e teste do *Design Thinking*.

Posto isto, foram efetuados variados testes de avaliação, na ótica de avaliação de modelos, como métricas de avaliação de deteção humana como a *accuracy (ACC)* e *heat maps*. Também se procedeu à reamostragem dos dados e conseqüente avaliação de diferentes tempos de processamento. Adicionalmente, e como referido anteriormente, também foi feita a interseção de algumas informações resultantes do questionário efetuado com os resultados dos modelos.

3.6.1 Accuracy

A *accuracy (ACC)* é uma medida que determina a percentagem de imagens corretamente classificadas. Todavia, no contexto de análise de imagens e deteção de objetos, tipicamente a classe de verdadeiros negativos não se aplica. Isto deve-se à mesma implicar uma deteção negativa, ou seja, incluir todas as possíveis deteções erradas que não foram detetadas. Desta forma, a *ACC* é calculada pelo total de verdadeiros positivos a dividir pelo conjunto de verdadeiros positivos, com falsos positivos e falsos negativos.

Neste seguimento, um verdadeiro positivo é um exemplo em que, neste caso a pessoa foi corretamente identificada como uma pessoa, um falso positivo quando é detetada uma pessoa, mas quando a área real de deteção não a representa, ou quando a confiança é inferior a 75% e por fim, um falso negativo, onde num caso real de uma pessoa não é detetado pelo modelo quando devia. Um exemplo deste conjunto de classes está presente na Figura 11.

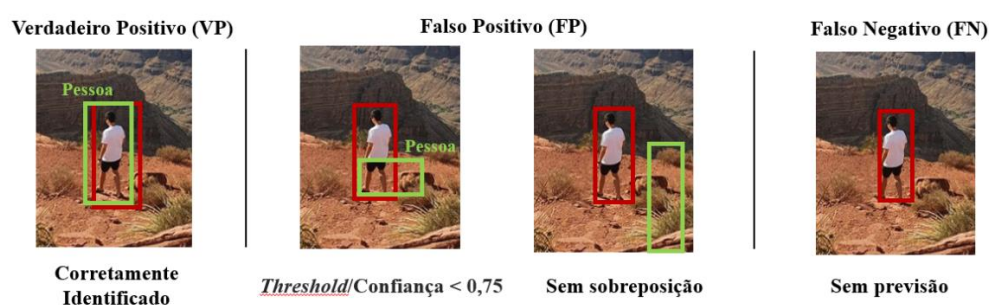


Figura 11 - Classes de classificação de objetos

Optou-se por aplicar esta métrica de avaliação devido à elevada presença da mesma, ao longo dos anos, na literatura, como nos estudos de Ali *et al.* (2022), Kao *et al.* (2021) e Liu *et al.* (2019). Optou-se por não apresentar a matriz de confusão devido à mesma ser mais adequada quando se tem presente mais que uma classe de objeto, mas também pelo facto de aspetos de classes, como os verdadeiros negativos, não se aplicarem na deteção de objetos, pois implicam uma deteção negativa. Para cada vídeo foi avaliado a capacidade de deteção de objetos de modelo através da análise da sua ACC. Posteriormente foi também avaliada a precisão média de deteção de todas as câmaras.

Tabela 10 - Accuracy na deteção de pessoas nas imagens

Câmara	ACC Média	Mediana da ACC	ACC Mínima	ACC Máxima
1	84,66%	84,13%	75,30%	97,20%
2	85,14%	84,02%	75,01%	97,83%
3	84,93%	84,33%	75,06%	98,90%
4	85,12%	85,10%	75,01%	97,73%
5	84,91%	85,56%	75,01%	97,98%
6	83,96%	84,04%	75,20%	96,35%
7	84,48%	83,88%	75,01%	99,07%
8	83,10%	81,92%	75,01%	97,82%
MÉDIA	84,54%	84,12%	75,08%	97,86%

De acordo com as informações referentes à Tabela 10, podemos concluir que o modelo, em média, tem a capacidade de detetar pessoas com uma precisão de 84,54%, e que a sua mediana não se afasta muito sendo de 84,12% o que significa que a ACC média é um bom indicador para a ACC. De referir ainda que a ACC mínima detetada foi de 75%, o limite mínimo de confiança para uma pessoa ser detetada pelo modelo, e a ACC máxima, em média de todas as câmaras, perto de 98%.

O Gráfico 11 tem presente a informação relativa à frequência dos valores de ACC em cinco classes distintas, para o total das câmaras.

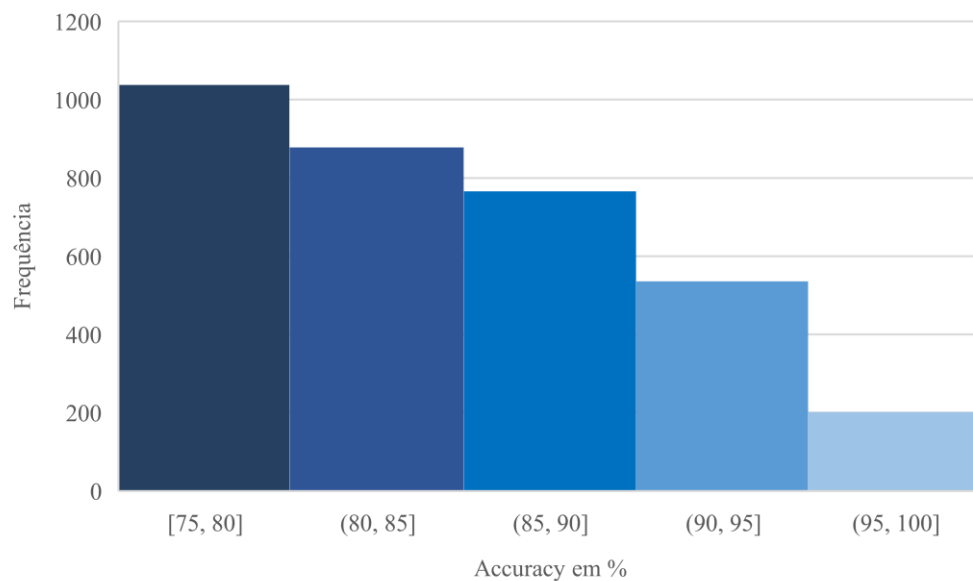


Gráfico 11 - Histograma de frequências de classes de *accuracy* na deteção de pessoas

A maioria das classificações apresenta uma ACC entre 75% e 80% e a frequência dessas mesmas classificações vai diminuindo gradualmente a par do aumento da ACC. Seria de esperar também um menor peso nas classes superiores a 90% de ACC de deteção pois, caso contrário, poderia haver problemas de balanceamento dos dados. Os resultados vão assim de encontro às informações referentes à média e mediana, com a maioria dos dados a se apresentar nas primeiras duas classes.

3.6.2 Mapa de frequências e *Heat map*

Posteriormente, foi construído um *heat map* sobre os resultados do mapa de frequências, de modo a apresentar os dados sobre o *frame* que representaria o *layout* do espaço, como demonstra o exemplo da Figura 12.



Figura 12 - Exemplo de transição de um *frame* original para *heat map*

Os resultados da detecção humana podem ser utilizados para a construção de *heatmaps* e os mesmos podem fornecer um resumo visual da informação através da representação bidimensional dos dados. Com esta forma de apresentação, é possível utilizar cores para desenhar as relações entre os valores dos dados que seriam muito mais difíceis de compreender se apresentados numa folha de valores numéricos (Nguyen *et al.*, 2022).

Os *heat maps* para investigação, devido à privacidade dos dados, são aplicados sobre um *frame*, para cada câmara, com o fundo desfocado como apresentado tanto na Figura 13, como na Figura 14.

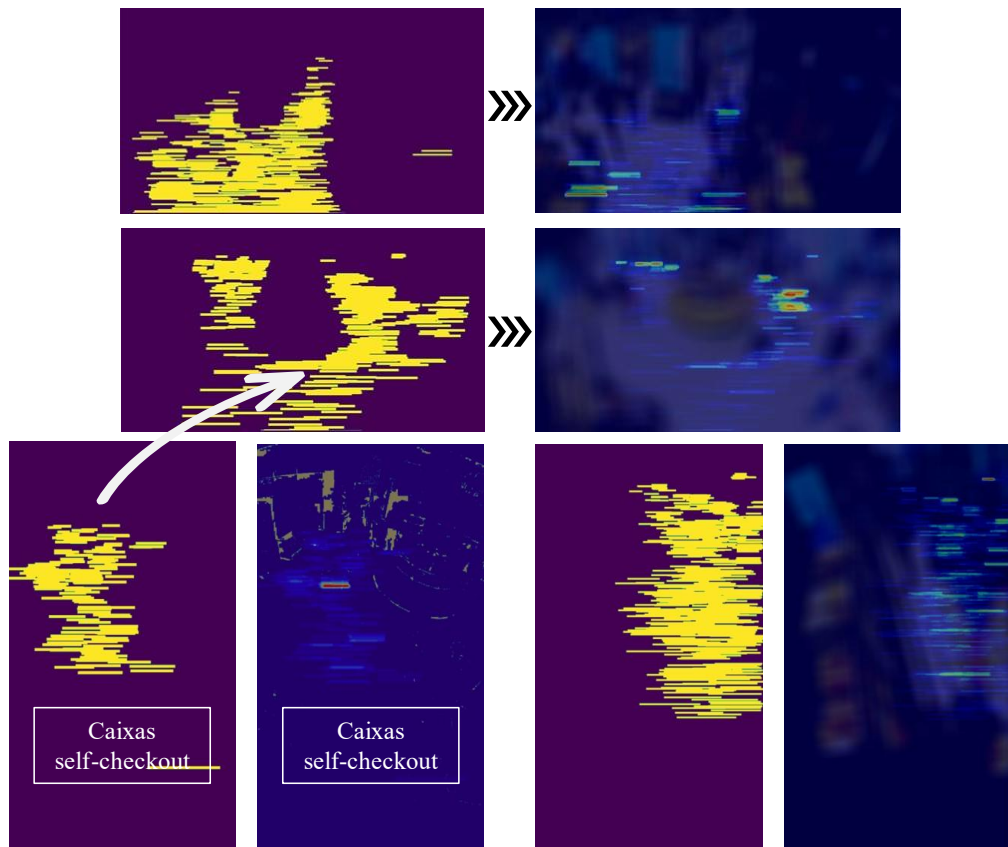


Figura 13 - Mapa de frequências e *heat map* na "zona aberta"

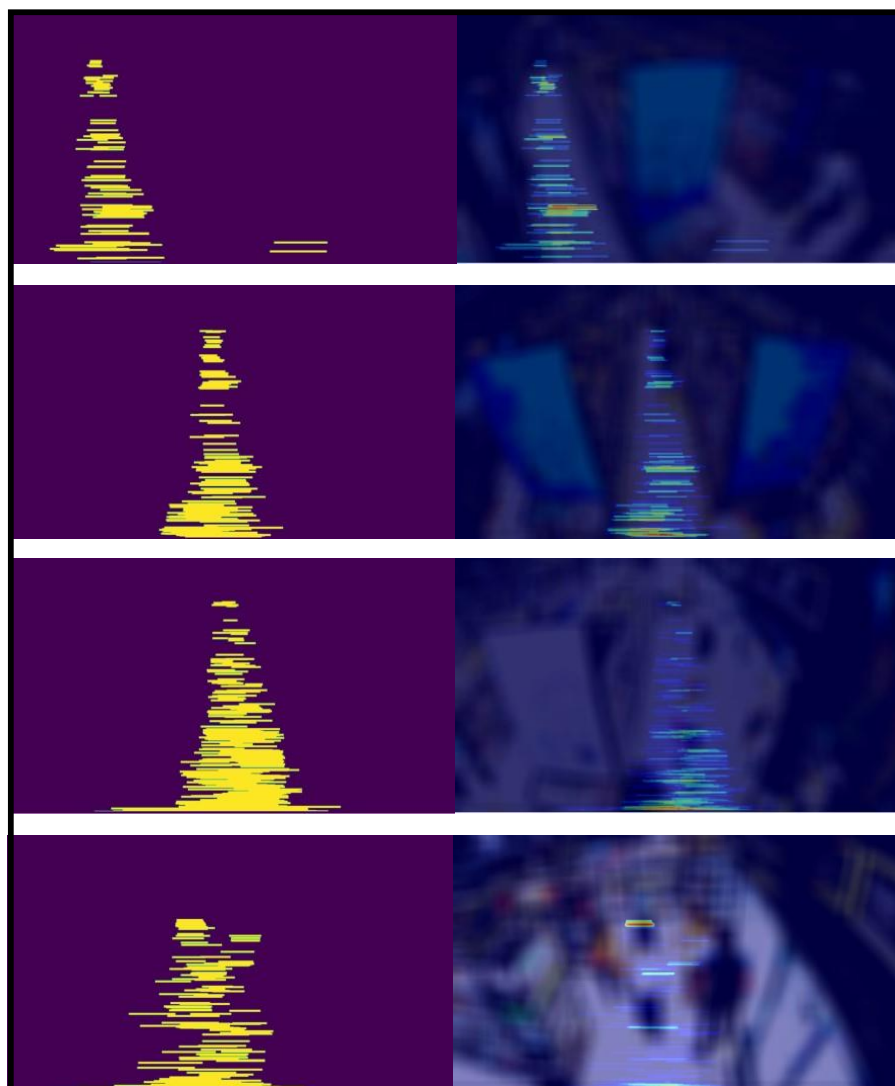


Figura 14 - Mapa de frequências e *heat map* dos corredores

As rotas mais percorridas na loja e a maioria das secções visitadas emergem, e apenas com esta informação, as lojas podem ter acesso a novos tipos de *insights*, tais como onde posicionar os seus produtos, correlações entre as áreas da loja e formas de otimizar o espaço.

Na ótica de tentar melhorar a experiência do consumidor, podem, entre outros, ser encontrados padrões sazonais de consumo, como por exemplo, no primeiro conjunto de mapas da Figura 14, estar presente uma elevada aglomeração na entrada desse corredor, identificada pela cor avermelhada no respetivo *heat map*. De referir que, a secção em causa, tratar-se-ia da papelaria e sendo que o conjunto de dados é referente ao início do ano letivo, pode ser associada uma relação de sazonalidade. Desta forma pode ser interessante analisar estas ligações para perceber os produtos mais requeridos pelos clientes em determinadas épocas e posicioná-los em zonas estratégicas, para promover uma maior rotatividade na loja.

Também é possível ser identificada uma maior afluência de clientes na denominada “zona aberta” do estabelecimento do que nos corredores e na zona de congelados, que liga os mesmos. Podem também ser explorados que tipo de produtos levam a esta discrepância de fluxos de pessoas e como se poderá uniformizar os mesmos, de modo a tentar promover uma menor proporção de aglomerações na loja. Estes são alguns exemplos de potencialidades que as análises de imagens dos comportamentos dos consumidores podem promover de modo a simplificar o processo de compra em lojas físicas e tentar proporcionar uma melhor experiência aos consumidores.

3.6.3 Cruzamento com resultados do questionário

Para além da ótica de avaliação de modelos através de métricas de avaliação da deteção humana, os resultados foram também cruzados com algumas informações pertinentes identificadas na fase de Empatia e Definição da metodologia.

Deste modo, é de refletir em aspetos como 46% dos inquiridos no questionário terem identificado o fator de não perder tempo em filas de espera e ainda se apresentar uma grande afluência às mesmas, não obstante de terem a opção de um processo de compra sem caixas, como demonstra a Figura 13.

Ao comparar os resultados de ambos os *heat maps*, do questionário e do resultante dos modelos, no do primeiro a zona com maior seleção foi a referente a frutas e legumes, o que vai de encontro aos *heat maps* dos modelos onde, na Figura 13, a região central da “zona aberta” representa essa mesma secção. Todavia, devido a essa secção ser muito próxima das caixas de *checkout* autónomo, as conclusões não podem ser absolutas devido à variável de incerteza. Contudo, de realçar também a zona da padaria, identificada pela câmara 7 cuja informação está presente na Tabela 9 onde, ao contrário dos resultados do questionário, parece ser das zonas com maiores afluências, principalmente nos mapas de frequências.

Os cenários apresentados são apenas pequenos exemplos de interligações possíveis, estes que poderão ser aprofundados se for efetuada a ligação de cada tipo de produtos a zonas da loja, e de melhor forma identificados com uma visão clara do estabelecimento.

3.6.4 Tempo de execução

O tempo de execução ainda é um tema pouco discutido em investigações referentes a detecção de objetos, onde a precisão dessa mesma detecção ainda se caracteriza como a referência mais importante. Todavia, em aplicações num ambiente de mundo real, o tempo de execução é crucial para a viabilidade de modelos, pois pode determinar a validade dos resultados possivelmente voláteis a aspetos como atualização das variáveis, sazonalidades, ou mesmo, nas áreas da saúde, questões de vida (Yamashita *et al.*, 2018).

Como forma de avaliação do tempo de execução, o mesmo foi separado em grupos diferentes. A totalidade do tempo do processo é dividido em 5 fases:

- Tempo de transformação do vídeo com características de privacidade dos dados;
- Tempo extração dos *frames* do vídeo;
- Tempo de processamento do modelo de detecção de objetos;
- Tempo de cálculo das frequências de cada objeto;
- Tempo de execução do *heat map*.

O tempo de execução, para um computador com as características abaixo descritas, está evidenciado na Tabela 11.

- ✓ Infraestrutura: *Cloud Oracle*
- ✓ Memória RAM: 60 *gigabytes*
- ✓ Largura de banda de rede: 4.1 *gigabytes*
- ✓ OCPU: 4 (equivalente a 8 CPU Intel físicos)

Tabela 11 - Tempo de execução das fases do processo

Fase	Descrição	Média do Tempo em minutos	Média do Tempo em horas
1	Converter Privacidade	47 min	0,78 horas
2	Extração <i>Frames</i>	33 min	0,55 horas
3	Processamento Modelo	6 min	0,10 horas
4	Cálculo Frequências	2 min	0,03 horas
5	<i>Heat Map</i>	Instantâneo	Instantâneo
Total	<i>Tempo Total</i>	88 min	1 h 28 min

Deste modo, é possível concluir a capacidade de análise de comportamentos e aglomerações em cada secção das lojas em pouco menos de uma hora e meia para muitos *gigabytes* de vídeo e informação. Realçar ainda que a fase que demorou mais tempo foi a de privatização dos dados, ocupando mais de metade do tempo, e que após extraídos os *frames*, processado o modelo e calculadas das frequências, a criação de um *heat map* é instantânea.

3.7 Implementação

Na última etapa, mesmo que o objetivo dos modelos seja aumentar o conhecimento dos dados, e não especificamente ser alvo de utilização como base para decisões ou apoio no processo empresarial, esse mesmo conhecimento adquirido terá de ser organizado, apresentado e distribuído de forma que o utilizador final possa utilizá-lo (Plotnikova *et al.*, 2020). Não obstante, esta dissertação foi apresentada à empresa que forneceu os dados que permitiram estes resultados.

Na mesma, foram desenvolvidos modelos que permitem, numa fase inicial, detetar pessoas em ambientes de retalho, e posteriormente, analisar o seu comportamento através de métricas de padrões de consumo que identificam aglomerações nos estabelecimentos comerciais.

Este tipo modelos e estas variedades de resultados podem ser aproveitados não só por lojas e supermercados com *checkout* autónomo, mas também por negócios de retalho com um formato tradicional. Dependendo do regulamento geral sobre a proteção dos dados (RGPD), mas também das suas condições de privacidade, recolha e tratamento dos dados, as instituições de retalho poderão tornar os seus processos de análise e reformulação do *layout* mais eficientes. Os modelos criados permitem identificar fatores como as rotas que a maioria dos clientes tomam e as secções mais visitadas. E, dito isto, com esta forma de informação pode ser mais fácil compreender os clientes, as suas decisões e também talvez, por exemplo, até a aceitação de certos produtos aos seus olhos. Se este modelo, porventura, for aplicado diversas vezes durante um ano, poderá ser possível que emerjam sazonalidades de escolha de produtos e que se definam afinidades.

4. Conclusões

O processo de compras continua a consistir numa prática que ainda causa transtorno aos consumidores, devido a fatores como elevados tempos dentro da loja e longas filas de espera. De forma a tentar encontrar novas informações referentes, não só aos seus comportamentos, mas também às suas preferências, a análise de imagens relativas a esses comportamentos é uma alternativa viável e eficiente no que toca a essa questão.

Face a esta matéria, pretende-se perceber como é que a análise de imagens de câmaras das lojas poderia ajudar a retirar informações dos comportamentos dos clientes, de forma a fornecer uma melhor experiência aos mesmos, identificar as aglomerações e promover uma maior rotatividade.

Devido à vertente ligada aos consumidores, uma combinação de CRISP-DM com *Design Thinking* apresentou-se como uma metodologia que conjuga um profundo entendimento do estudo em análise, através da observação das necessidades dos consumidores e empatia que o *Design Thinking* procura, com fundamentação técnica necessária do CRISP-DM. Com o conhecimento gerado a partir do questionário, foi possível, não só comprovar o problema relativo ao processo de compra, mas também a perceção do consumidor acerca das potencialidades que podem ser exploradas através destes métodos.

Assim sendo, recorreu-se à avaliação de diversas técnicas de *data mining*, tanto para a recolha, como tratamento e processamento dos dados de imagem e, posteriormente, procedeu-se à criação de modelos, com base em algoritmos de *deep learning* com capacidades, não só de deteção de objetos, mas também de análise de comportamento dos consumidores.

Desta forma, foi possível concluir que o *Yolo* permitiu a análise dos vídeos com um tempo de deteção de objetos muito rápido para o volume de dados em questão e com uma boa capacidade de precisão. Adicionalmente, a sua combinação com o *HMM* permitiu analisar as trajetórias dos clientes num ambiente retalhista e contruir mapas de localização no mesmo e, com o auxílio de bibliotecas como o *OpenCv* e *ImageAI*, foi possível desenvolver estes modelos de forma intuitiva e com linhas de condução simples.

De referir ainda que, as restrições relativas à privacidade dos dados limitaram, de certa forma, a apresentação de alguns dados, mas também de resultados pretendidos. Todavia, em conformidade com os resultados obtidos, também se comprovou que os resultados da deteção humana podem ser avaliados através de diferentes tipos de métricas, de modo a apresentar conclusões que seriam muito mais difíceis de compreender se apresentadas numa folha de valores numéricos. Também com base nesses resultados, pode-se concluir que os objetivos da

investigação foram concretizados. Efetivamente, foram criados tanto modelos de detecção e localização, como métricas de avaliação de padrões de consumo que apresentam vários contributos, ao produzirem resultados passíveis de análise. Foram identificadas aglomerações na loja, assim como identificadas as rotas mais percorridas e a maioria das secções visitadas. Métricas essas, que tiveram valor acrescentado para a caracterização da perceção dos consumidores para os ambientes de retalho.

Por conseguinte, foi possível ao estudo responder à questão de investigação, na medida que foi possível medir o impacto que os resultados provenientes dos modelos podem fornecer, de modo a identificar e mitigar possíveis aglomerações nas lojas, identificar os percursos mais percorridos e proporcionar orientações que possam melhorar a experiência do consumidor.

4.1 Contributos e Trabalhos futuros de investigação

Face ao problema encontrado, a presente investigação, pode contribuir numa vertente empresarial, para os profissionais da área do retalho, como também em vertentes académicas e de investigação.

Antes de mais, foi efetuada uma RSL sobre métodos de análise de dados de sensores, de modo a perceber o comportamento dos consumidores em ambientes de retalho, que, porventura, poderá contribuir para futuros investigadores na área de *Big Data*, *deep learning* e *computer vision* aplicadas ao retalho e não só. Nesta RSL é disponibilizada uma síntese dos principais artigos nestas áreas, e avaliados em várias vertentes, entre elas, o contexto desses estudos, as diversas tecnologias que podem ser utilizadas, tipologias dos modelos e avaliação de resultados. Desta forma, a análise desta RSL pode ser um bom ponto de partida para possíveis novos estudos nesta área.

Adicionalmente, nesta investigação foi explorada a interligação de duas metodologias distintas, o CRISP-DM com o *Design Thinking* de forma a tirar o maior proveito das duas e permitir uma maior flexibilidade nos processos. Também esta combinação de metodologias pode ser útil na vertente académica e de investigação, na medida que, nas fases das mesmas onde se sentiria que uma das metodologias estaria a restringir a concretização dos objetivos, complementar-se-ia com as qualidades da outra.

Também em duas vertentes mais práticas, esta investigação apresenta novos contributos, na vertente ligada à perceção do consumidor para diversos ambientes de retalho por parte do questionário, como na vertente associada aos métodos de análise de dados de imagem e vídeo.

Na primeira vertente, o questionário para além de ter como objetivo criar empatia com os consumidores, também tem por base servir de impulso a quem pretende seguir esta temática, de modo a perceber, de forma mais aprofundada o nível de aceitação da população aos novos formatos de loja sem caixas de pagamento e não só.

Simultaneamente, nesta investigação é também evidenciada a utilidade da análise de imagens provenientes de câmaras de estabelecimentos forma a perceber comportamentos dos consumidores. Porventura, é conjuntamente salientado a utilidade de algoritmos como o *Yolo* devido à sua simplicidade de utilização combinada com a capacidade inerente de reconhecimento de objetos em imagens em ambientes de carácter dinâmico. Do mesmo modo, evidenciou-se a importância, ao contrário do expresso na RSL, da avaliação de resultados através da análise de relações provenientes de *heat maps* que seriam muito mais difíceis de compreender se apresentados numa folha de valores numéricos e que poderão levar a conclusões diferentes.

Por fim, o contributo de aplicabilidade ao contexto nacional, onde a apresentação destes métodos, mas também os resultados, vem contribuir para uma maior genuinidade dos mesmos, na medida que, todos os países têm diferentes consumidores, com características diversificadas, e de serem recolhidos os dados provenientes das câmaras de um estabelecimento comercial português.

Em relação a trabalhos futuros que esta investigação pode tomar, de forma a complementar o que já foi realizado, podem ser testadas hipóteses como a real recomendação de produtos, ou mesmo a sua disposição na loja. Desta forma, poderá ser criado um modelo que, após o cliente escolher um produto na loja, preveja e recomende o próximo produto a ser adicionado ao carrinho de compras, identificando a secção e o percurso até ao mesmo. Assim sendo, em vez de trabalhar apenas com a deteção de pessoas, complementar-se-ia com a deteção de produtos e as suas interações com os clientes. Adicionalmente, no presente estudo, as imagens provenientes das câmaras foram analisadas em separado. Neste âmbito, poderá ser efetuada a incorporação de vídeos de várias câmaras de modo a ser criado um único vídeo, que por si mesmo, tenha a capacidade de analisar a totalidade do espaço.

4.2 Limitações

Em matéria de limitações, as mesmas serão separadas em duas perspetivas diferentes, como identificadas na RSL, limitações relativas ao ambiente em análise e relativas à privacidade.

No que toca a limitações relativas ao ambiente em análise, de referir que, no contexto nacional, ainda está presente muita escassez de lojas sem caixas para serem alvo de estudo. Adicionalmente, e como previamente identificado na literatura, ainda perduram fatores como a oclusão, e o facto de os dados fontes serem provenientes de diversos ficheiros. Em termos de mitigação, a oclusão, foi parcialmente mitigada pelo aumento de *frames* gerados para análise de cada vídeo, de forma aos objetos serem identificados em mais ocasiões e detetar com maior precisão os clientes. No que toca aos dados fontes, a escolha de múltiplas câmaras, e a combinação dos seus vídeos permite que as mesmas tenham a capacidade a captar uma maior extensão do espaço e todos os seus percursos. Posteriormente foi efetuada a agregação dos ficheiros de vídeo provenientes da mesma câmara.

Em relação a limitações referentes à privacidade, é importante referir, a preocupação pela proteção da identidade dos clientes, mas também da empresa e do ambiente em análise. A opção mais evidente para colmatar a privacidade dos clientes seria o desfoque das suas faces, no entanto o espaço comercial continuaria a ser identificado de forma clara. Desta forma, procedeu-se à remoção do fundo dos vídeos e assim sendo os clientes também só seriam expostos os contornos das suas formas, e o espaço não estaria identificado.

Referências Bibliográficas

- Ahmed, B., Dannhauser, T., & Philip, N. (2019). A Lean Design Thinking Methodology (LDTM) for Machine Learning and Modern Data Projects. *2018 10th Computer Science and Electronic Engineering Conference, CEEC 2018 - Proceedings*, 11–14. <https://doi.org/10.1109/CEEC.2018.8674234>
- Ali, K., Liu, A. X., Chai, E., & Sundaresan, K. (2022). Monitoring Browsing Behavior of Customers in Retail Stores via RFID Imaging. *IEEE Transactions on Mobile Computing*, 21(3), 1034–1048. <https://doi.org/10.1109/TMC.2020.3019652>
- Alipio, M. I., Peñalosa, K. M. T., & Unida, J. R. C. (2020). In-store customer traffic and path monitoring in small-scale supermarket using UWB-based localization and SSD-based detection. *Journal of Ambient Intelligence and Humanized Computing*. <https://doi.org/10.1007/s12652-020-02236-z>
- Alkhodre, A. B., & Alshantiti, A. M. (2021). Employing Video-based Motion Data with Emotion Expression for Retail Product Recognition. *International Journal of Advanced Computer Science and Applications*, 12(10), 817–825. <https://doi.org/10.14569/IJACSA.2021.0121091>
- Bellini, P., Palesi, L. A. I., Nesi, P., & Pantaleo, G. (2022). Multi Clustering Recommendation System for Fashion Retail. *Multimedia Tools and Applications*. <https://doi.org/10.1007/s11042-021-11837-5>
- Benaich, N., Hogarth, Ian., Sebbouh, O., Rajkumar, N. (2022). State of AI Report 2022. <https://www.stateof.ai/>
- Bradlow, E. T., Gangwar, M., Kopalle, P., & Voleti, S. (2017). The Role of Big Data and Predictive Analytics in Retailing. *Journal of Retailing*, 93(1), 79–95. <https://doi.org/10.1016/j.jretai.2016.12.004>
- Carbajal, S. G. (2021). Customer segmentation through path reconstruction. *Sensors*, 21(6), 1–17. <https://doi.org/10.3390/s21062007>
- Chai, J., Zeng, H., Li, A., & Ngai, E. W. T. (2021). Deep learning in computer vision: A critical review of emerging techniques and application scenarios. *Machine Learning with Applications*, 6, 100134. <https://doi.org/10.24433/CO.0411648.v1>
- Chen, Y., Zheng, Z., Chen, S., Sun, L., & Chen, D. (2017). Mining Customer Preference in Physical Stores from Interaction Behavior. *IEEE Access*, 5, 17436–17449. <https://doi.org/10.1109/ACCESS.2017.2744263>

- Choudhary, A., & Chaudhury, S. (2016). Video analytics revisited. *IET Computer Vision*, *10*(4), 237–247. <https://doi.org/10.1049/iet-cvi.2015.0321>
- Chowdhury, B. D. B., Masoud, S., Son, Y.-J., Kubota, C., & Tronstad, R. (2022). A Dynamic HMM-Based Real-Time Location Tracking System Utilizing UHF Passive RFID. *IEEE Journal of Radio Frequency Identification*, *6*, 41–53. <https://doi.org/10.1109/JRFID.2021.3102507>
- Dogan, O., Bayo-Monton, J.-L., Fernandez-Llatas, C., & Oztaysi, B. (2019). Analyzing of gender behaviors from paths using process mining: A shopping mall application. *Sensors (Switzerland)*, *19*(3). <https://doi.org/10.3390/s19030557>
- Domingo, J. D., Aparicio, R. M., & Rodrigo, L. M. G. (2022). Cross Validation Voting for Improving CNN Classification in Grocery Products. *IEEE Access*, *10*, 20913–20925. <https://doi.org/10.1109/ACCESS.2022.3152224>
- Gazzola, P., Grechi, D., Martinelli, I., & Pezzetti, R. (2022). The Innovation of the Cashierless Store: A Preliminary Analysis in Italy. *Sustainability (Switzerland)*, *14*(4). <https://doi.org/10.3390/su14042034>
- Ghazal, M., Yaghi, M., Gad, A., el Bary, G., Alhalabi, M., Alkhedher, M., & El-Baz, A. S. (2021). AI-powered service robotics for independent shopping experiences by elderly and disabled people. *Applied Sciences (Switzerland)*, *11*(19). <https://doi.org/10.3390/app11199007>
- Hauser, M., Flath, C. M., & Thiesse, F. (2021). Catch me if you scan: Data-driven prescriptive modeling for smart store environments. *European Journal of Operational Research*, *294*(3), 860–873. <https://doi.org/10.1016/j.ejor.2020.12.047>
- Hsia, C.-H., Chang, T.-H. W., Chiang, C.-Y., & Chan, H.-T. (2022). Mask R-CNN with New Data Augmentation Features for Smart Detection of Retail Products. *Applied Sciences (Switzerland)*, *12*(6). <https://doi.org/10.3390/app12062902>
- Hwangbo, H., Kim, J., Lee, Z., & Kim, S. (2017). Store layout optimization using indoor positioning system. *International Journal of Distributed Sensor Networks*, *13*(2). <https://doi.org/10.1177/1550147717692585>
- Islam, M. M., & Baek, J.-H. (2021). Deep learning based real age and gender estimation from unconstrained face image towards smart store customer relationship management. *Applied Sciences (Switzerland)*, *11*(10). <https://doi.org/10.3390/app11104549>

- Ives, B., Cossick, K., & Adams, D. (2019). Amazon Go: Disrupting retail? *Journal of Information Technology Teaching Cases*, 9(1), 2–12. <https://doi.org/10.1177/2043886918819092>
- Jasim Hadi, H., Hameed Shnain, A., Hadishaheed, S., & Haji Ahmad, A. (2015). Mobility View project BIG DATA AND FIVE V'S CHARACTERISTICS. In *International Journal of Advances in Electronics and Computer Science* (Issue 2). <https://www.researchgate.net/publication/332230305>
- Kao, Y.-T., Lin, H.-W., & Qiu, D.-Y. (2021). Implementation of Offline Consumer Behavior Tracking. *International Journal of Pattern Recognition and Artificial Intelligence*, 35(9). <https://doi.org/10.1142/S0218001421500282>
- Kwon, J., Choi, Y., & Hwang, Y. (2021). Enterprise design thinking: An investigation on user-centered design processes in large corporations. *Designs*, 5(3). <https://doi.org/10.3390/designs5030043>
- Liu, J., Gu, Y., & Kamijo, S. (2017). Customer behavior classification using surveillance camera for marketing. *Multimedia Tools and Applications*, 76(5), 6595–6622. <https://doi.org/10.1007/s11042-016-3342-1>
- Liu, J., Liu, Y., Zhang, G., Zhu, P., & Chen, Y. Q. (2015). Detecting and tracking people in real time with RGB-D camera. *Pattern Recognition Letters*, 53, 16–23. <https://doi.org/10.1016/j.patrec.2014.09.013>
- Liu, Y., Guo, B., Li, N., Zhang, J., Chen, J., Zhang, D., Liu, Y., Yu, Z., Zhang, S., & Yao, L. (2019). DeepStore: An Interaction-Aware WideDeep Model for Store Site Recommendation with Attentional Spatial Embeddings. *IEEE Internet of Things Journal*, 6(4), 7319–7333. <https://doi.org/10.1109/JIOT.2019.2916143>
- Madureira, C. (2021). Planeamento e processo de abertura de loja SONAE MC-Continente Bom dia [MA thesis]. Instituto Politécnico de Setúbal, Escola Superior de Ciências Empresariais.
- Martinez-Plumed, F., Contreras-Ochando, L., Ferri, C., Hernandez-Orallo, J., Kull, M., Lachiche, N., Ramirez-Quintana, M. J., & Flach, P. (2021). CRISP-DM Twenty Years Later: From Data Mining Processes to Data Science Trajectories. *IEEE Transactions on Knowledge and Data Engineering*, 33(8), 3048–3061. <https://doi.org/10.1109/TKDE.2019.2962680>
- Merad, D., Drap, P., Lufimpu-Luviya, Y., Iguernaissi, R., & Fertil, B. (2016). Purchase behavior analysis through gaze and gesture observation. *Pattern Recognition Letters*, 81, 21–29. <https://doi.org/10.1016/j.patrec.2016.04.008>

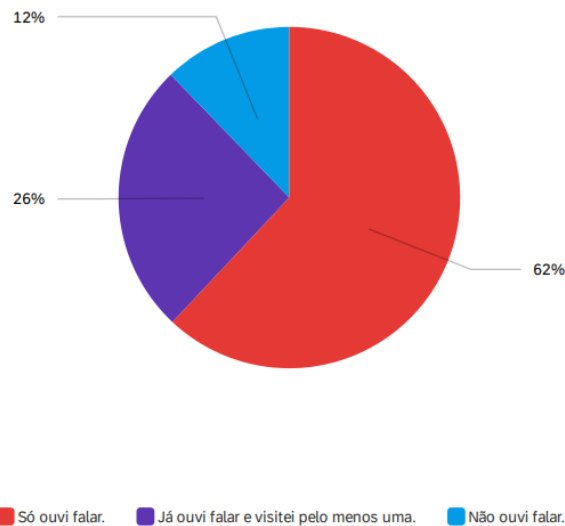
- Nguyen, K., Le, M., Martin, B., Cil, I., & Fookes, C. (2022). When AI meets store layout design: a review. *Artificial Intelligence Review*. <https://doi.org/10.1007/s10462-022-10142-3>
- Olafenwa, M., & Olafenwa, J. (2021). Official English Documentation for ImageAI! — ImageAI 2.1.6 documentation. DeepQuest AI. <https://imageai.readthedocs.io/>
- Pande, M., & Bharathi, S. V. (2020). Theoretical foundations of design thinking – A constructivism learning approach to design thinking. *Thinking Skills and Creativity*, 36. <https://doi.org/10.1016/j.tsc.2020.100637>
- Paolanti, M., Liciotti, D., Pietrini, R., Mancini, A., & Frontoni, E. (2018). Modelling and Forecasting Customer Navigation in Intelligent Retail Environments. *Journal of Intelligent and Robotic Systems: Theory and Applications*, 91(2), 165–180. <https://doi.org/10.1007/s10846-017-0674-7>
- Paolanti, M., Pietrini, R., Mancini, A., Frontoni, E., & Zingaretti, P. (2020). Deep understanding of shopper behaviours and interactions using RGB-D vision. *Machine Vision and Applications*, 31(7–8). <https://doi.org/10.1007/s00138-020-01118-w>
- Plotnikova, V., Dumas, M., & Milani, F. (2020). Adaptations of data mining methodologies: A systematic literature review. *PeerJ Computer Science*, 6, 1–43. <https://doi.org/10.7717/PEERJ-CS.267>
- Priyanka, G., Sathya Bama, S., Aiswaryadevi, V. J., Sruthi, M. S., & Soundarya, S. (2019). An effective interest identification technique to enhance sales performance in supermarkets. *International Journal of Innovative Technology and Exploring Engineering*, 8(12), 2574–2580. <https://doi.org/10.35940/ijitee.K2004.1081219>
- Quintana, M., Menendez, J. M., Alvarez, F., & Lopez, J. P. (2016). Improving retail efficiency through sensing technologies: A survey. *Pattern Recognition Letters*, 81, 3–10. <https://doi.org/10.1016/j.patrec.2016.05.027>
- Rengarajan, V., Vijay Anand, V., Vijayabanu, C., & Thiyagarajan, S. (2019). Customer perception on waiting time in super markets – an exploratory study using one way anova. *International Journal of Scientific and Technology Research*, 8(10), 2649–2659.
- Schwenzow, J., Hartmann, J., Schikowsky, A., & Heitmann, M. (2021). Understanding videos at scale: How to extract insights for business research. *Journal of Business Research*, 123, 367–379. <https://doi.org/10.1016/j.jbusres.2020.09.059>

- Singh, V. K., Singh, P., Karmakar, M., Leta, J., & Mayr, P. (2021). The journal coverage of Web of Science, Scopus and Dimensions: A comparative analysis. *Scientometrics*, *126*(6), 5113–5142. <https://doi.org/10.1007/s11192-021-03948-5>
- Wedel, M., & Kannan, P. K. (2016). Marketing analytics for data-rich environments. *Journal of Marketing*, *80*(6), 97–121. <https://doi.org/10.1509/jm.15.0413>
- Wei, Y., Tran, S., Xu, S., Kang, B., & Springer, M. (2020). Deep Learning for Retail Product Recognition: Challenges and Techniques. In *Computational Intelligence and Neuroscience* (Vol. 2020). Hindawi Limited. <https://doi.org/10.1155/2020/8875910>
- Wen, J., Guillen, L., Abe, T., & Suganuma, T. (2021). A hierarchy-based system for recognizing customer activity in retail environments. *Sensors*, *21*(14). <https://doi.org/10.3390/s21144712>
- Xia, K., Fan, H., Huang, J., Wang, H., Ren, J., Jian, Q., & Wei, D. (2021). An intelligent self-service vending system for smart retail. *Sensors*, *21*(10). <https://doi.org/10.3390/s21103560>
- Yamashita, R., Nishio, M., Do, R. K. G., & Togashi, K. (2018). Convolutional neural networks: an overview and application in radiology. In *Insights into Imaging* (Vol. 9, Issue 4, pp. 611–629). Springer Verlag. <https://doi.org/10.1007/s13244-018-0639-9>
- Zhang, Y., Jin, S., Wu, Y., Zhao, T., Yan, Y., Li, Z., & Li, Y. (2020). A new intelligent supermarket security system. *Neural Network World*, *30*(2), 113–131. <https://doi.org/10.14311/NNW.2020.30.009>
- Ziaie, A., ShamiZanjani, M., & Manian, A. (2021). Systematic review of digital value propositions in the retail sector: New approach for digital experience study. *Electronic Commerce Research and Applications*, *47*. <https://doi.org/10.1016/j.elerap.2021.101053>

Anexo A: Relatório do questionário

Q1. Antes deste questionário, já tinha ouvido falar ou visitado um formato de loja disruptiva como a Amazon Go; Continente Labs; Pingo Doce & Go ou Galp Smart Store?

#	Campos	%	Nº Respostas
1	Só ouvi falar	61,99%	137
2	Já ouvi falar e visitei pelo menos uma	25,79%	57
3	Não ouvi falar	12,22%	27
			221



Anexo 1 - Informação acerca da familiaridade do conceito - questionário

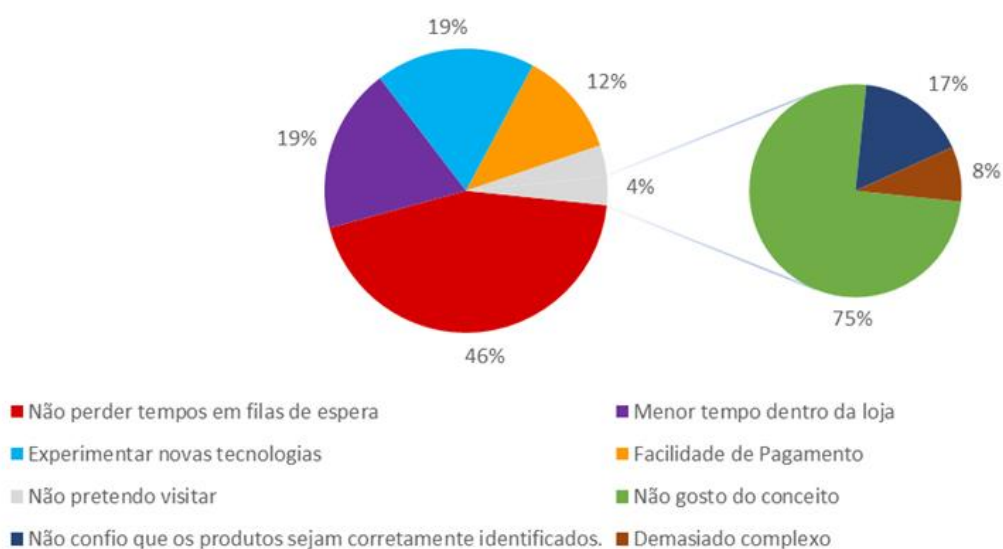
Grupo 2. No início de 2018, a Amazon lançou a primeira loja da sua cadeia Amazon Go, o primeiro espaço no mundo sem a necessidade de caixas de pagamento. É uma loja física, autónoma, suportada por algoritmos de *computer vision* e sensores, que detetam as interações dos clientes com os expositores e respetivos produtos.

Q2.1. Na eventualidade de visitar este tipo de lojas, quais seriam as principais razões para essa visita, que características imaginaria encontrar e constrangimentos ultrapassar? Escolha no máximo 2 opções.

#	Campos	%	Nº Respostas
1	Não perder tempos em filas de espera	45,63%	162
2	Menor tempo dentro da loja	19,44%	69
3	Experimentar novas tecnologias	18,87%	67
4	Facilidade de pagamento	12,39%	44
5	Não pretendo visitar	3,66%	13
			355

Q2.2. Caso não pretenda visitar qual seria a razão para tal eventualidade?

#	Campos	%	Nº Respostas
1	Não gosto do conceito	75,00%	9
2	Não confio que os produtos sejam corretamente identificados	16,67%	2
3	Demasiado complexo	8,33%	1
			12



Anexo 2 - Informação sobre a motivação para visitar lojas deste conceito – questionário

Grupo 3. Atualmente o processo de compra neste conceito de loja segue os seguintes passos:

(1) Registrar ou fazer login na aplicação e inserir os seus dados;

(2) Aceder ao QR Code na aplicação para entrar na loja;

(3) Começar o processo de compra onde:

(3.1) Todos os artigos são automaticamente adicionados ao carrinho de compras após serem retirados das prateleiras ou, quando lidos os códigos de barras;

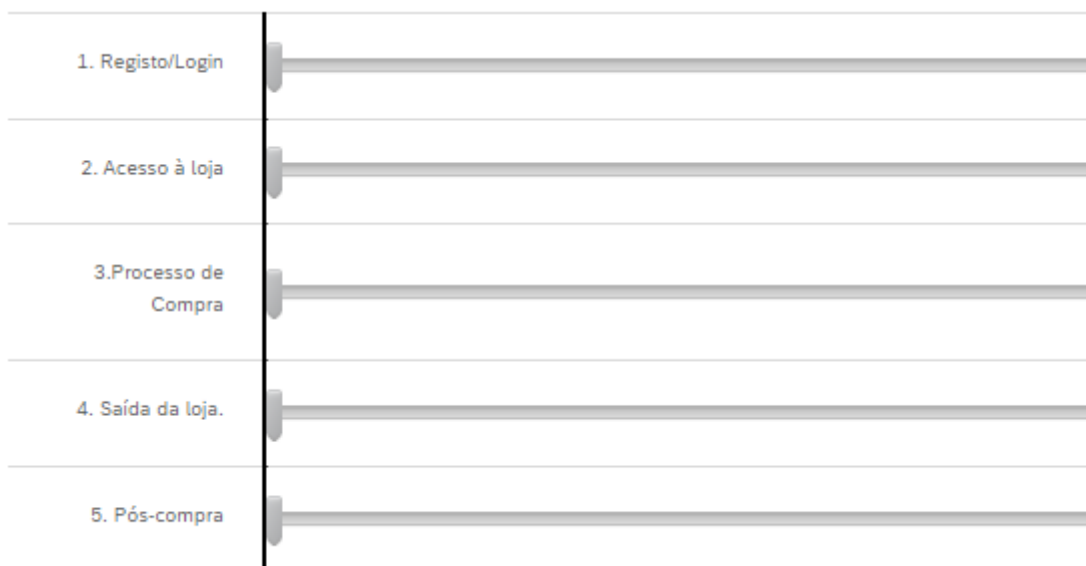
(3.2) Os artigos poderão ser removidos do carrinho de compras, bastando voltar a recolocá-los ou reler o código de barras;

(3.3) Não se deverá retirar produtos das prateleiras para entregar a outros compradores, de modo a evitar a identificação de artigos indesejados;

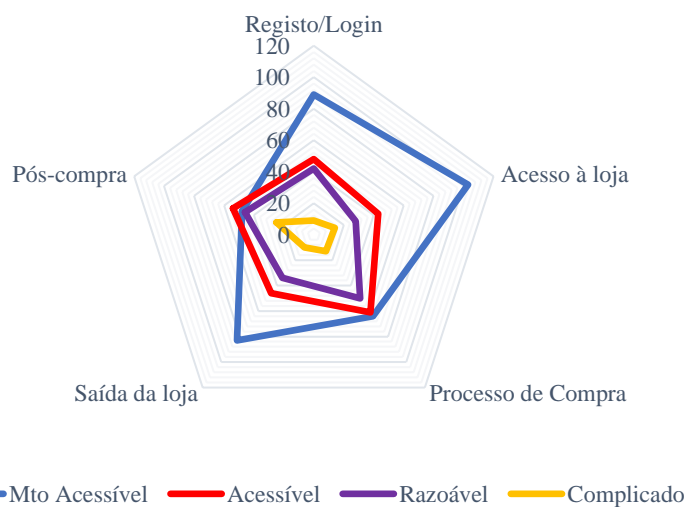
(4) Após o término da compra apresentar de novo o QR Code à saída;

(5) Se, porventura, forem identificados produtos indevidamente cobrados, é possível proceder a uma reclamação via app.

Q3.1. Qual a sua perceção em relação ao processo acima descrito? Deverá avaliar cada uma das etapas, movendo o retângulo. (1-Muito Acessível; 2-Acessível; 3-Razoável; 4-Complicado)



#	Campos	Mto acessível		Acessível		Razoável		Complicado		Total
		%	Nº Respostas	%	Nº Respostas	%	Nº Respostas	%	Nº Respostas	
1	Registo/ Login	47.34%	89	25.53%	48	22.34%	42	4.79%	9	188
2	Acesso à loja	54.79%	103	22.87%	43	14.89%	28	7.45%	14	188
3	Processo de Compra	34.04%	64	32.45%	61	26.60%	50	6.91%	13	188
4	Saída da loja	47.98%	83	26.59%	46	19.65%	34	5.78%	10	173
5	Pós-compra	27.75%	48	31.21%	54	26.59%	46	14.45%	25	173



Anexo 3 - Informação sobre a perceção do processo de compra - questionário

Q3.2. Em relação ao processo de compra dentro deste novo modelo de supermercados.

Q3.2.1. Sente que o processo de compra transmite confiança/facilidade ao consumidor?

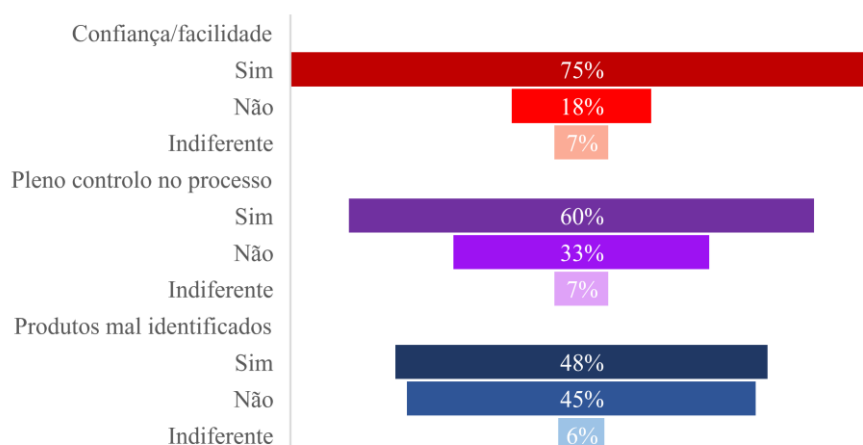
#	Campos	%	Nº Respostas
1	Sim	75,40%	141
2	Não	17,65%	33
3	Indiferente	6,95%	13
			187

Q3.2.1. Este modelo transmite-lhe que o cliente terá pleno controlo no processo?

#	Campos	%	Nº Respostas
1	Sim	60,43%	113
2	Não	32,62%	61
3	Indiferente	6,95%	13
			187

Q3.2.1. Receia que, neste processo de compra, os produtos possam não ser corretamente identificados no seu carrinho?

#	Campos	%	Nº Respostas
1	Sim	49,20%	92
2	Não	44,92%	84
3	Indiferente	5,88%	11
			187



Anexo 4 - Informação das características do processo de compra – questionário

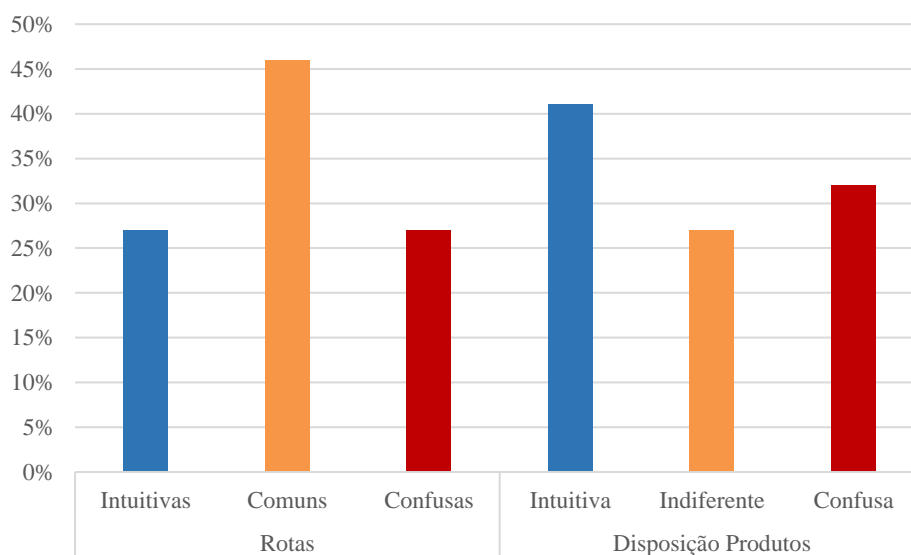
Grupo 4. Em relação aos processos de compra que foi experienciando ao longo da sua vida em qualquer modelo de supermercado (sem caixas, com caixas, híbrido):

Q4.1. Como classifica a sua experiência dentro dos estabelecimentos?

# Campos		%	Nº Respostas
1	Intuitiva: Tenho uma rota fluída, onde não me sinto perdido dentro da loja.	26,92%	49
2	Comum: Não me sinto incomodado pelas rotas que tenho dentro dos estabelecimentos.	45,60%	83
3	Confusa: Sinto que percorro muitos espaços da loja à procura do que pretendo e não tenho a sensação de ter um processo de compras fluído.	27,47%	50
			182

Q4.2. Qual a sua opinião em relação à atual disposição das secções dos supermercados?

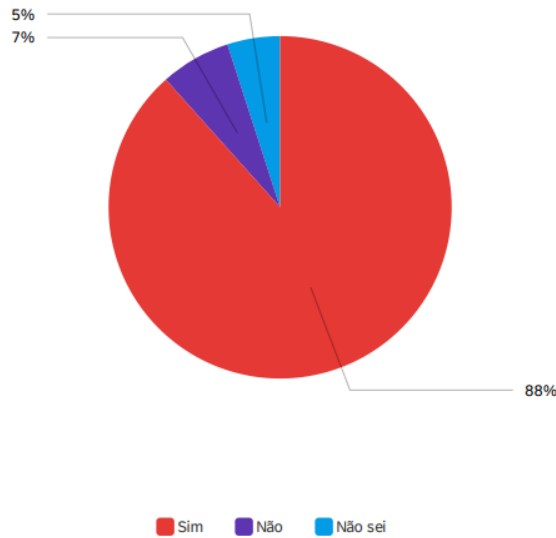
# Campos		%	Nº Respostas
1	Intuitiva	40,66%	74
2	Indiferente	26,92%	49
3	Confusa	32,42%	59
			182



Anexo 5 - Informação sobre experiência em estabelecimentos comerciais - questionário

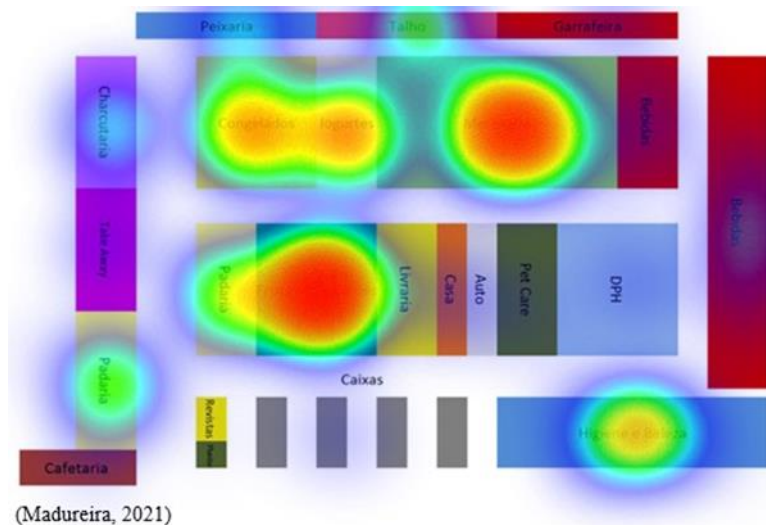
Q4.2. Na sua opinião seria vantajoso que as empresas tivessem em consideração o layout das lojas de modo que o tempo despendido dentro das mesmas seja reduzido?

#	Campos	%	Nº Respostas
1	Sim	88,46%	161
2	Não	6,59%	12
3	Não sei	4,95%	9
			182



Anexo 6 - Informação sobre posição na consideração do layout das lojas - questionário

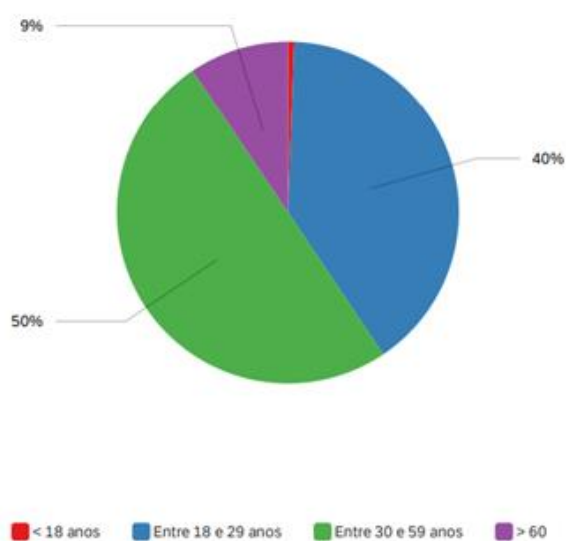
Q5. De uma forma simplificada, que secções visita mais regularmente nas suas deslocações ao supermercado? Indique no máximo 4, clicando sobre as mesmas.



Anexo 7 - Informação da afluência nas secções no estabelecimento – questionário

Q6. Em que faixa etária se insere?

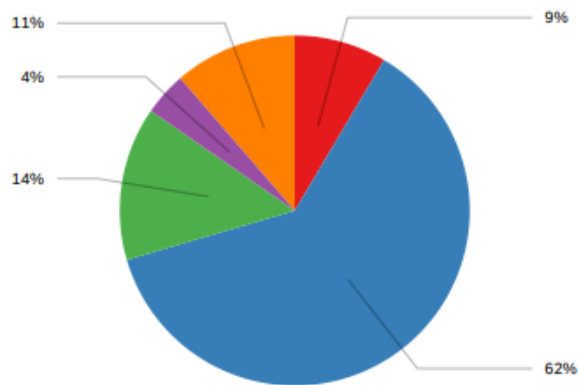
# Campos		%	Nº Respostas
1	Menos de 18 anos	0,56%	1
2	Entre 18 anos e 29 anos	40,11%	71
3	Entre 30 e 59 anos	50,28%	89
4	Mais que 60 anos	9,04%	16
			177



Anexo 8 - Informação sobre faixa etária - questionário

Q7. Qual o seu rendimento mensal?

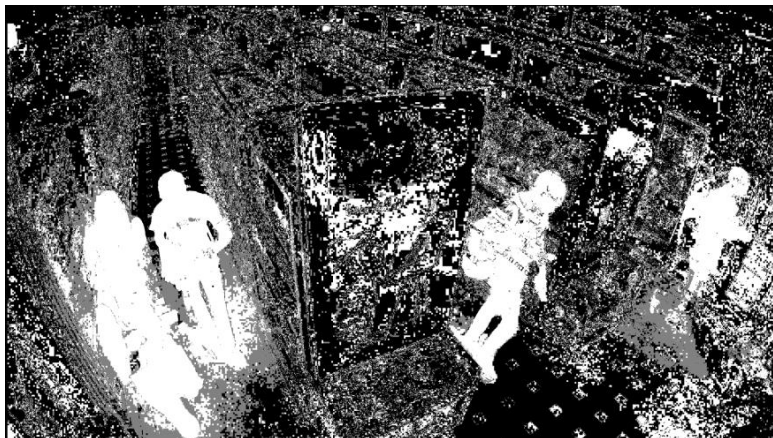
# Campos		%	Nº Respostas
1	Até 2 salários mínimos nacionais. (SMN: 705€)	9,04%	16
2	Dois a quatro SMN	61,58%	109
3	Quatro a seis SMN	14,12%	25
4	Superior a seis SMN	3,95%	7
5	Não se aplica	11,30%	20
			177



■ Até dois salários mínimos nacionais. (SMN: 705€) ■ Dois a quatro salários mínimos nacionais. ■ Quatro a seis salários mínimos nacionais.
■ Superior a seis salários mínimos nacionais. ■ Não se aplica.

Anexo 9 - Informação sobre rendimento mensal - questionário

Anexo B: Análise de imagens



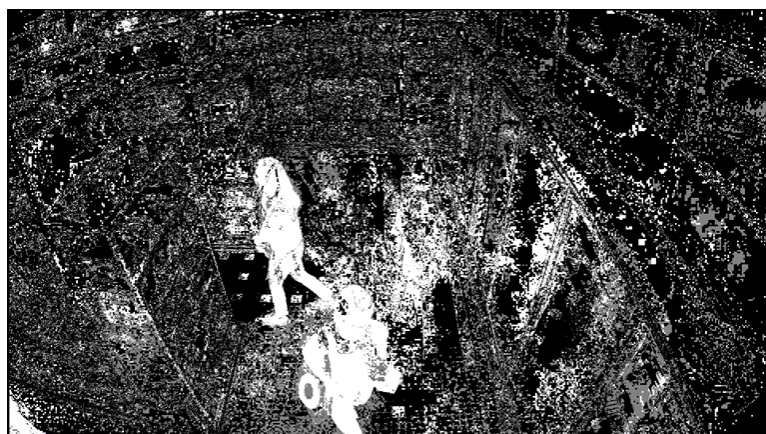
Anexo 10 - *Frame câmara n°1*



Anexo 11 - *Frame câmara n° 2*



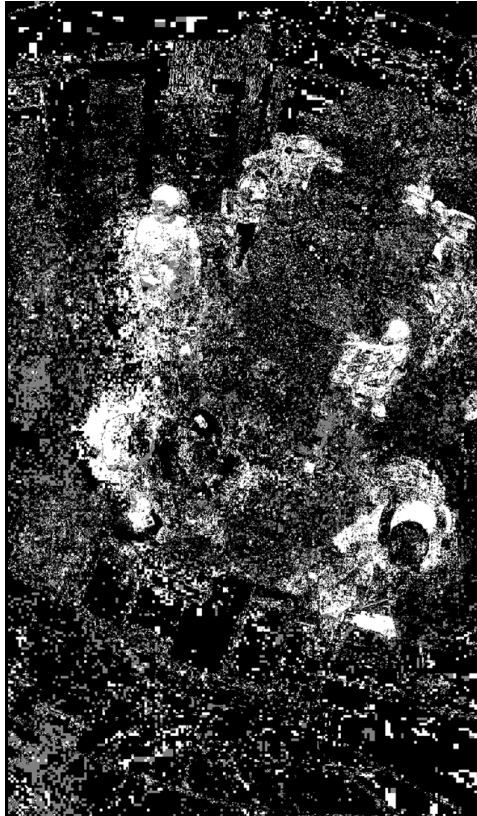
Anexo 12 - *Frame câmara n°3*



Anexo 13 - *Frame* câmara n°4



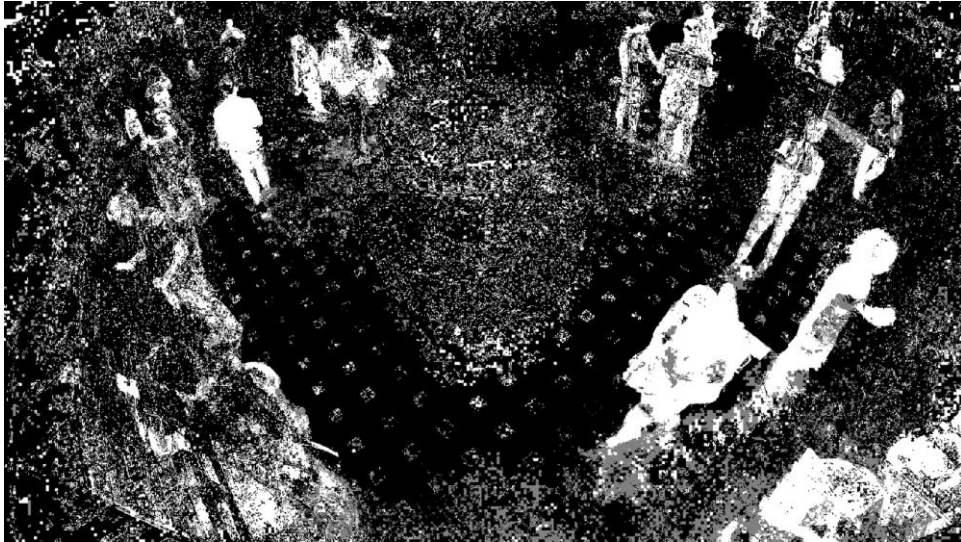
Anexo 14 - *Frame* câmara n° 5



Anexo 15 - *Frame* câmara n°6



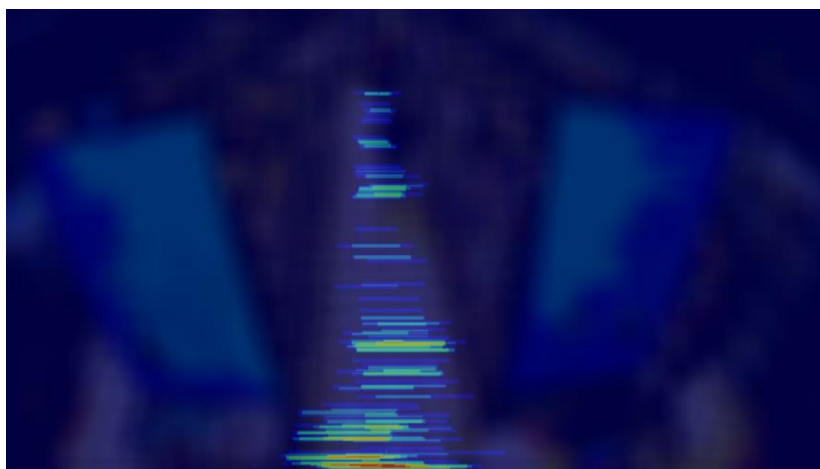
Anexo 16 - *Frame* câmara n°7



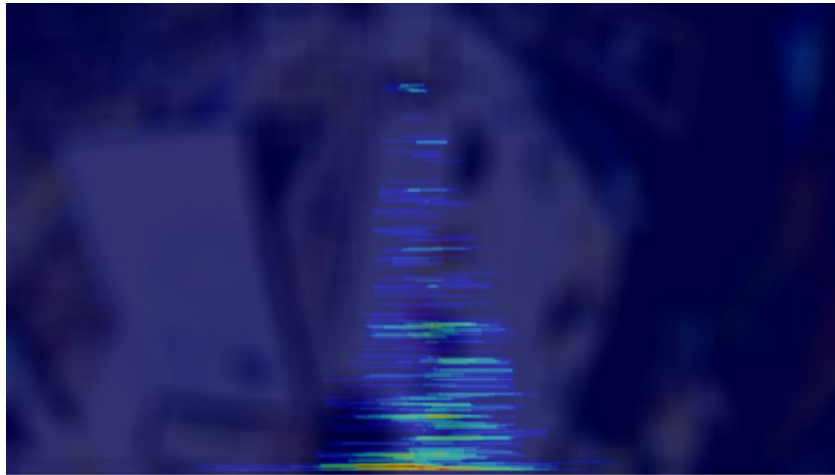
Anexo 17 - Frame câmara n°8



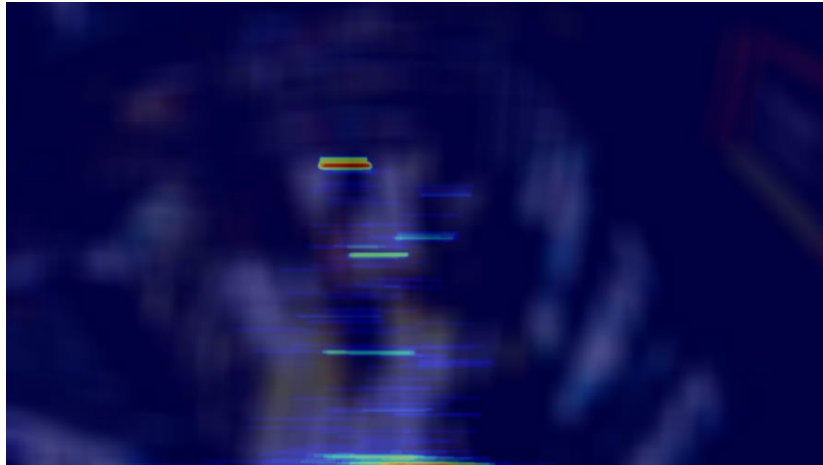
Anexo 18 - Heat map câmara n°1



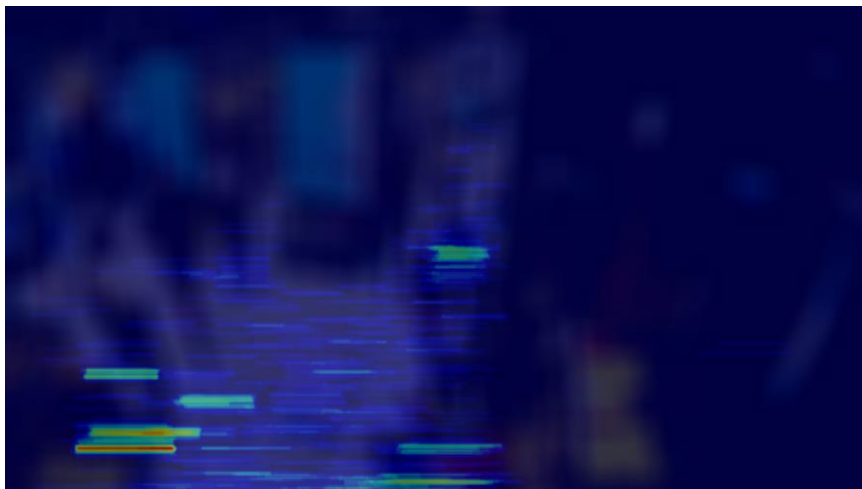
Anexo 19 - Heat map câmara n°2



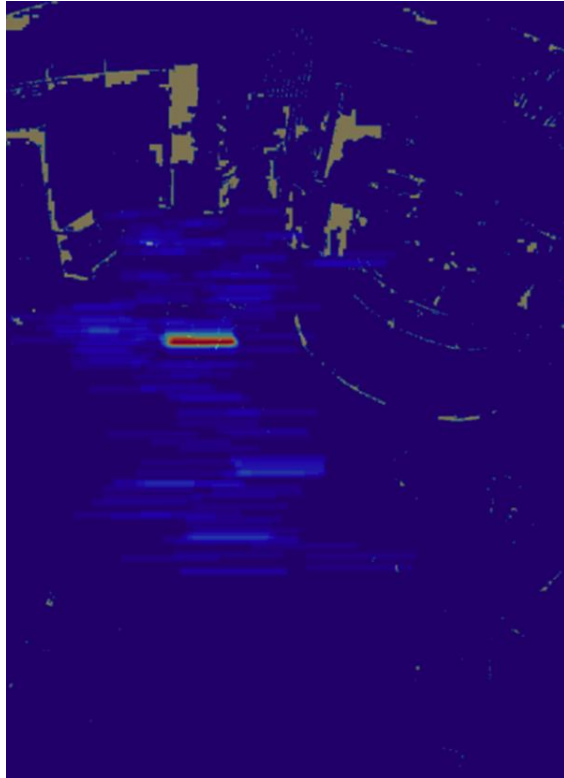
Anexo 20 - *Heat map* câmara n°3



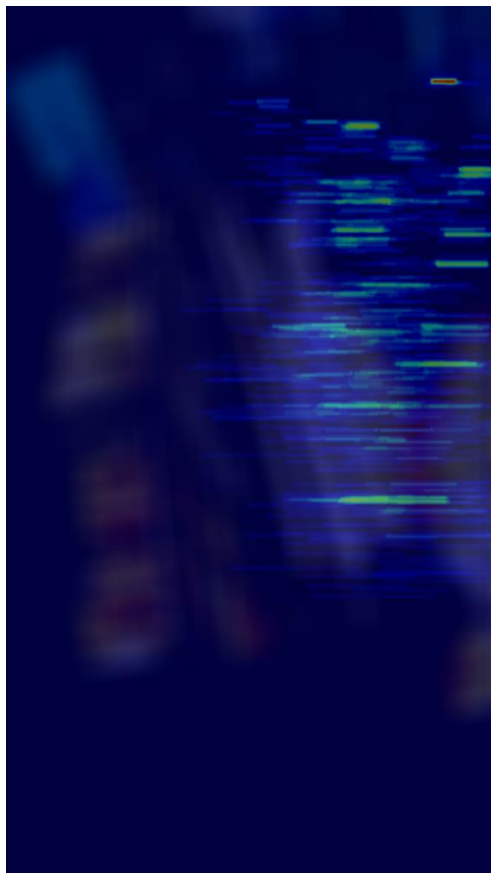
Anexo 21 - *Heat map* câmara n°4



Anexo 22 - *Heat map* câmara n°5



Anexo 23 - *Heat map* câmara nº6



Anexo 24 - *Heat map* câmara nº7



Anexo 25 - *Heat map* câmara nº8