

UNIVERSIDADE DE PASSO FUNDO
Programa de Pós-Graduação em
Computação Aplicada

Dissertação de Mestrado

**SISTEMA DE RECOMENDAÇÃO
PARA O COMÉRCIO BASEADO EM
VISÃO COMPUTACIONAL**

MALOMAR ALEX SEMINOTTI



UNIVERSIDADE DE PASSO FUNDO
INSTITUTO DE CIÊNCIAS EXATAS E GEOCIÊNCIAS
PROGRAMA DE PÓS-GRADUAÇÃO EM COMPUTAÇÃO APLICADA

**SISTEMA DE RECOMENDAÇÃO PARA O
COMÉRCIO BASEADO EM VISÃO
COMPUTACIONAL**

Malomar Alex Seminotti

Dissertação apresentada como requisito parcial à obtenção do grau de Mestre em Computação Aplicada na Universidade de Passo Fundo.

Orientador: Prof. Dr. Rafael Rieder

Passo Fundo
2021

CIP – Catalogação na Publicação

S471s Seminotti, Malomar Alex
Sistema de recomendação para o comércio baseado em
visão computacional / Malomar Alex Seminotti. – 2021.
56 f. : il. ; 30 cm.

Orientador: Prof. Dr. Rafael Rieder.
Dissertação (Mestre em Computação Aplicada) –
Universidade de Passo Fundo, 2021.

1. Inteligência artificial. 2. Percepção facial. 3. Visão por
computador. 4. Software - Desenvolvimento. I. Rieder, Rafael,
orientador. II. Título.

CDU: 004.4

Catalogação: Bibliotecário Luís Diego Dias de S. da Silva – CRB 10/2241

ATA DE DEFESA DO TRABALHO DE CONCLUSÃO DE CURSO DO ACADÊMICO

MALOMAR ALEX SEMINOTTI

Aos vinte dias do mês de agosto do ano de dois mil e vinte e um, às quinze horas, realizou-se, de forma on-line, por meio de videoconferência, a sessão pública de defesa do Trabalho de Conclusão de Curso “Sistema de Recomendação para o Comércio Baseado em Visão Computacional”, de autoria de Malomar Alex Seminotti, acadêmico do Curso de Mestrado em Computação Aplicada do Programa de Pós-Graduação em Computação Aplicada – PPGCA. Segundo as informações prestadas pelo Conselho de Pós-Graduação e constantes nos arquivos da Secretaria do PPGCA, o aluno preencheu os requisitos necessários para submeter seu trabalho à avaliação. A banca examinadora foi composta pelos doutores Rafael Rieder (Orientador), Guilherme Afonso Madalozzo (UPF) e Tales Nereu Bogoni (UNEMAT). Concluídos os trabalhos de apresentação e arguição, a banca examinadora considerou o candidato **APROVADO**. Foi concedido o prazo de até quarenta e cinco (45) dias, conforme Regimento do PPGCA, para o acadêmico apresentar ao Conselho de Pós-Graduação o trabalho em sua redação definitiva, a fim de que sejam feitos os encaminhamentos necessários à emissão do Diploma de Mestre em Computação Aplicada. Para constar, foi lavrada a presente ata, que vai assinada pelos membros da banca examinadora e pela Coordenação do PPGCA.



Prof. Dr. Rafael Rieder – UPF
Presidente da Banca Examinadora
(Orientador)



p.p.
Prof. Dr. Guilherme Afonso Madalozzo – UPF
(Avaliador Interno)



p.p.
Prof. Dr. Tales Nereu Bogoni – UNEMAT
(Avaliador Externo)



Prof. Dr. Carlos Amaral Mölbig – UPF
Coordenador do PPGCA

Dedico este trabalho ao professor Dr. Cristiano Roberto Cervi (*in memoriam*), pelo seu interesse em acompanhar o desenvolvimento do estudo e o desejo de vê-lo concluído.

AGRADECIMENTOS

A todos os professores que contribuíram com seu conhecimento para que mais este passo fosse concluído na minha formação.

Ao meu amigo de longa data, Dr. Rafael Rieder, pelo seu apoio e orientações no decorrer deste trabalho.

À minha família, por toda compreensão e apoio, e que entenderam a necessidade de algumas ausências durante este período.

SISTEMA DE RECOMENDAÇÃO PARA O COMÉRCIO BASEADO EM VISÃO COMPUTACIONAL

RESUMO

Sistemas de recomendação são utilizados em diversos domínios, como sugestão de conteúdo ou de compra de produtos, auxiliando em processos de tomada de decisão. Esses sistemas geralmente analisam conteúdos armazenados em bancos de dados, e conjuntos de dados não estruturados como imagens e vídeos são pouco explorados. Com isso em mente, o presente trabalho apresenta um sistema de recomendação com a capacidade de indicar produtos, em tempo real, com base no comportamento das pessoas, considerando a análise de imagens de vídeo utilizando técnicas de visão computacional e inteligência artificial. As inferências para sugestões de vendas consideram o monitoramento de imagens de uma região pré-definida da loja, com produtos que os usuários observam durante uma visita ao estabelecimento. O estudo foi conduzido com a participação de 15 voluntários, em ambiente de laboratório, em virtude das restrições impostas pela pandemia da Covid-19. A aplicação identificou as pessoas corretamente 100% das vezes, sem gerar falsos positivos, ao mesmo tempo que preservou a identidade dos participantes, já que não houve nomeação das mesmas em nenhum momento. O estudo demonstrou que a aplicação é capaz de registrar o comportamento e o hábito de compra de clientes ao lojista, com 73,4% de eficiência na identificação dos produtos observados. Esse fato pode ajudar o lojista na oferta de produtos similares com preços mais acessíveis, oferecer vantagens ao cliente para agregar valor de venda, melhorar a disposição de produtos nos pontos de venda, além de ajudar no treinamento de seus colaboradores.

Palavras-Chave: inteligência artificial, reconhecimento facial, sistemas de recomendação, visão computacional.

COMPUTER VISION-BASED RECOMMENDER SYSTEM FOR COMMERCE

ABSTRACT

Recommender systems are used in several domains, such as content suggestion or product purchase, helping in decision-making processes. These systems often analyze content stored in databases, and unstructured datasets such as images and videos are underexplored. With this in mind, this work presents a recommender system with the ability to indicate products, in real-time, based on people's behavior, considering the analysis of video images using computer vision and artificial intelligence techniques. The inferences for sales suggestions take monitoring images of a pre-defined area, with products that users observe during a visit to the store. We conducted a pilot study with 15 volunteers, in a laboratory environment, due to the restrictions imposed by the Covid-19 pandemic. The application identified all people correctly 100%, without generating false positives while preserving the participants' identities, not named them any time. The study showed that the application could register customers' behavior and buying habits to the store owner, offering an efficiency of 73.4% in identifying the products observed. This fact can help retailers giving similar products at more affordable prices, offering advantages to the customer to add sales value, improve the availability of products at the points of sale, in addition to helping to train their employees.

Keywords: artificial intelligence, computer vision, face recognition, recommender systems.

LISTA DE FIGURAS

Figura 1	–	Pontos identificados na face	27
Figura 2	–	Diagrama esquemático da aplicação	34
Figura 3	–	Exemplo da obtenção da posição estimada da cabeça	35
Figura 4	–	Exibição do registro da face, produtos de interesse associados e respectivas datas de visualização	37
Figura 5	–	Desenho da posição de instalação da câmera para monitoramento da prateleira.	39
Figura 6	–	Prateleira modelo utilizada no ambiente de experimento	40
Figura 7	–	Exemplo de captura da face mais próxima à câmera	46
Figura 8	–	Aplicação mostrando as prateleiras configuradas e as informações sobre a face	47

LISTA DE TABELAS

Tabela 1	–	Regiões da face considerada por algoritmo de mapeamento facial	26
Tabela 2	–	Arquivo com os dados da captura e a face codificada	38
Tabela 3	–	Arquivo de configuração	41
Tabela 4	–	Dados das pessoas participantes no experimento	43
Tabela 5	–	Resultado da visualização dos produtos	44

SUMÁRIO

1	INTRODUÇÃO	19
2	FUNDAMENTAÇÃO TEÓRICA	23
2.1	SISTEMAS DE RECOMENDAÇÃO	23
2.1.1	Filtragem Baseada em Conteúdo	23
2.2	VISÃO COMPUTACIONAL	24
2.2.1	Reconhecimento Facial	25
2.3	TRABALHOS RELACIONADOS	27
3	MATERIAIS E MÉTODOS	33
3.1	SOFTWARE E HARDWARE	33
3.2	VISÃO DA ARQUITETURA DO SISTEMA	33
3.3	FUNCIONAMENTO DA APLICAÇÃO	35
3.4	AMBIENTE DE TESTE	39
4	RESULTADOS E DISCUSSÕES	43
4.1	RESULTADOS DO EXPERIMENTO	43
4.2	DISCUSSÕES	45
5	CONCLUSÃO	51
	Referências	53

1. INTRODUÇÃO

Empresas, em geral, buscam aumentar seu volume de vendas e de clientes para, conseqüentemente, aumentar seus lucros. Uma das formas de obter isso é melhorar as recomendações para seus produtos, oferecendo itens que realmente sejam de interesse dos consumidores. Este objetivo pode ser alcançado com a utilização de um sistema de recomendação [1].

De acordo com Ricci, Rokach e Shapira [2], um sistema de recomendação pode combinar várias técnicas computacionais para selecionar itens personalizados com base nos interesses dos usuários e conforme o contexto no qual estão inseridos. Existem vários tipos de sistemas de recomendação, sendo o mais comum o baseado em conteúdo. Este sistema utiliza palavras-chave para descrever itens e construir um perfil de usuário com base nos itens que representam as preferências e os interesses dos mesmos. Nesta perspectiva, os algoritmos desenvolvidos para tais sistemas visam recomendar itens similares àqueles que determinado usuário mostrou interesse no passado ou está examinando no presente. Assim, por meio de comparação entre produtos previamente classificados, os itens com maior correlação são recomendados.

Um dos objetivos de um sistema de recomendação é ajudar os usuários a encontrar informações, produtos ou serviços, analisando seus requisitos. Fornecem sugestões de itens a serem recomendados para um usuário, ajudando-o na análise dos dados e na tomada de decisão [3]. Esse tipo de solução tem contribuído no aumento de vendas on-line, principalmente para sites de comércio eletrônico, e tem se tornado um importante campo de pesquisa desde meados da década de 1990 [4]. Eles podem combinar várias técnicas de software para selecionar itens personalizados com base nos interesses dos usuários e conforme o contexto no qual estão inseridos.

Um dos requisitos principais de um sistema de recomendação é trazer itens relevantes para os usuários. O nível de confiança dos usuários no sistema está diretamente relacionado à qualidade das recomendações. Se os usuários não receberem produtos e serviços favoráveis, o mecanismo de recomendação poderá ser considerado inadequado em relação à satisfação do cliente, o que torna evidente para os usuários procurarem sistemas alternativos [5].

Pesquisadores da área de sistema de recomendação procuram estudar e melhorar abordagens capazes de fornecer dados e informações que possam aumentar a assertividade nas sugestões realizadas. Nesse sentido, sistemas de visão computacional, associados aos algoritmos de recomendações, podem ajudar em processos de tomada de decisão e a melhorar a experiência de compra das pessoas [6].

O comércio eletrônico é um domínio famoso de aplicações de recomendação [5]. Os sistemas existentes são criados pensando em atender o comércio on-line e não o físico [7]. De acordo com Mora et al. [8], o comércio off-line ainda carece de meios para enriquecer a experiência de compra na loja para torná-la mais conveniente, personalizada e envolvente. Em alguns casos, as informações são coletadas e analisadas para permitir, *a posteriori*, a recomendação com base no histórico de compras das pessoas. Desta forma, perde-se a possibilidade de oferecer produtos de interesse do cliente no momento em que o mesmo demonstrou este desejo.

A análise do comportamento das pessoas, durante sua visita ao estabelecimento, também pode ajudar a atrair e reter clientes com mais eficiência e eficácia, e os varejistas podem obter vantagem competitiva sobre seus concorrentes [3]. O uso de um sistema de recomendação baseado em visão computacional pode mostrar a um lojista os produtos que mais estão chamando atenção de seus clientes no momento da compra, mesmo sem a sua supervisão ou fora do alcance de um atendente. Estes dados podem se tornar relevantes ao lojista, uma vez que, conhecendo o comportamento dos clientes, ele pode reorganizar os produtos nas prateleiras, deixando em destaque os itens mais vendidos ou com maior margem.

A visão computacional resulta da integração das áreas de processamento digital de imagens e inteligência artificial, tendo como propósito a obtenção de algoritmos capazes de interpretar o conteúdo visual de imagens. Suas aplicações estão presentes em diversos segmentos tecnológicos que envolvem análise de imagens, reconhecimento de padrões e controle inteligente [9].

Embora os sistemas de recomendação existentes sejam bem sucedidos na produção de sugestões, eles ainda enfrentam desafios de desempenho, como precisão, escalabilidade e inicialização a frio (recomendação sem dados históricos) [10]. Para tanto, esses sistemas têm usado algoritmos de *deep learning* para melhorar a qualidade de suas recomendações [5].

O sucesso das práticas de *deep learning* tem afetado significativamente as direções de pesquisa e desenvolvimento não apenas na área de Computação. O uso de técnicas de *deep learning* também demonstram desempenho eficiente e proficiente em recomendações, melhorando os recursos de computação, a coleta e o processamento eficiente das informações [5] - especialmente para identificar padrões a partir da análise de conjuntos de dados não estruturados [11].

Dados estruturados facilitam a manipulação das informações, diferentemente de dados não estruturados, como imagens e vídeos, onde é necessária a utilização de técnicas de visão computacional para extrair as informações. Estas técnicas requerem maior capacidade de processamento para serem executadas. Pesquisadores têm trabalhado constantemente em aprimorar o processo de extração de características de imagens [12]. A

correspondência das imagens de consulta são, geralmente, considerados um modelo, que corresponde aos recursos das imagens no banco de dados.

Desta forma, um sistema que usa elementos de recomendação e de visão pode informar o interesse do cliente em um determinado item no momento em que ele ocorre e, desta forma, ajudar na abordagem do lojista, que poderá ter informações em tempo real sobre o interesse do cliente. O lojista pode, também, sugerir produtos similares, com um custo mais atrativo. Isso tende a gerar um potencial maior na transação comercial, aumentando o valor da venda, além de fidelizar o cliente pela prestatividade e objetividade no atendimento [13].

Nesse contexto, este trabalho apresenta a implementação de um sistema de recomendação denominado **Recommender Vision**, baseado em conteúdo e técnicas de visão computacional para sugerir produtos de interesse de clientes em tempo real, considerando o seu comportamento. As inferências para sugestões de vendas tomam por base o monitoramento de imagens de determinada região da loja, com produtos que os clientes estão observando durante uma compra, e também cruzando com histórico existente de interesses visuais anteriores. A abordagem considera a preservação de identidade dos clientes, a partir da correspondência facial, sem identificação nominal, considerando a Lei Geral de Proteção de Dados Pessoais (LGPD) [14].

O presente trabalho está dividido em cinco capítulos. O Capítulo 2 apresenta breve fundamentação teórica sobre visão computacional e sistemas de recomendação, bem como trabalhos relacionados ao objeto de estudo. O Capítulo 3 destaca os materiais e métodos aplicados para a concepção do Recommender Vision. O Capítulo 4 mostra os resultados e discussão de testes preliminares realizados com o software proposto. Por fim, o Capítulo 5 apresenta as conclusões e os trabalhos futuros para a continuidade da solução.

2. FUNDAMENTAÇÃO TEÓRICA

Este capítulo apresenta a fundamentação sobre conceitos básicos de sistemas de recomendação e visão computacional. Em seguida, são apresentados os trabalhos relacionados a este estudo, mostrando abordagens que procuram combinar visão e recomendação.

2.1 SISTEMAS DE RECOMENDAÇÃO

Sistemas de recomendação (SR) são aplicações que tem por objetivo auxiliar no processo de filtragem e fornecimento de sugestões personalizadas de itens, de forma automática (total ou parcial), de acordo com o potencial interesse de um usuário. O uso de SR tem se mostrado uma abordagem eficiente para reduzir o esforço do usuário na busca de informações de seu interesse [15].

Existem várias aplicações de recuperação de conteúdo que buscam auxiliar os usuários na identificação de itens de interesse. Porém, é comum que estas aplicações tragam muito conteúdo irrelevante [7]. Buscando minimizar este problema, SR têm focado no uso de recursos de inteligência artificial e de análise de grandes volumes de dados para garantir uma busca adequada ao perfil do indivíduo. Características diversas do próprio usuário e determinados requisitos relacionados aos itens que se deseja encontrar são considerados para atingir metas de recomendação precisas e personalizadas [16].

O interesse em SR ainda permanece alto porque constitui uma área de pesquisa rica em problemas e por causa da abundância de aplicações práticas que ajudam os usuários a lidar com a sobrecarga de informações e fornecer recomendações personalizadas, conteúdo e serviços para eles. No entanto, apesar de todos esses avanços, a atual geração de SR ainda requer melhorias para tornar os métodos de recomendação mais eficazes e aplicáveis a uma gama ainda maior de aplicações da vida real [7].

Diferentes técnicas são aplicadas nos SR para que os conteúdos mais adequados para seus usuários sejam encontrados [17]. A abordagem desse trabalho tem por foco recomendações por meio de filtragem baseada em conteúdo.

2.1.1 Filtragem Baseada em Conteúdo

A filtragem baseada em conteúdo (FBC) considera a previsão nas propriedades do item recomendado, concentrando-se na similaridade entre elas [18].

A abordagem de recomendação baseada em conteúdo tem suas raízes na pesquisa de recuperação de informações e na filtragem de informações. Sistemas de conteúdo se concentram em recomendar itens que contenham informações textuais, como documentos, sites (URLs) e mensagens de notícias. A melhoria em relação às abordagens tradicionais de recuperação de informações vem do uso de perfis de usuário que contêm informações sobre os gostos, preferências e necessidades destes [7].

Para Ferretto [19], esta técnica procura prever a avaliação de um item por um usuário que ainda não fez a avaliação deste. Este método procura recomendar itens pertencentes a grupos de itens pelos quais o usuário demonstrou interesse recentemente [20].

Os SR baseados em conteúdo têm algumas limitações, tais como [7]:

- *Análise limitada de conteúdo*: as técnicas baseadas em conteúdo são restringidas pelos recursos explicitamente associados aos objetos que esses sistemas recomendam;
- *Especialização excessiva*: quando o sistema só pode recomendar itens com pontuação alta em relação ao perfil de um usuário, o usuário está limitado a ser recomendado para itens semelhantes aos que já foram classificados;
- *Novo usuário*: o usuário deve avaliar um número suficiente de itens antes que um SR baseado em conteúdo possa realmente entender as preferências do usuário e apresentar recomendações confiáveis.

Para Vivian [21], quando o conteúdo das informações é pouco estruturado, como vídeo, som e imagens, esta técnica necessita de maior poder computacional. Uma alternativa é o uso de metadados, como legendas e *tags*, ao invés dos dados propriamente ditos, que utilizam menor poder de processamento.

2.2 VISÃO COMPUTACIONAL

Visão computacional (VC) é uma área que desenvolve teoria e tecnologia para construir sistemas inteligentes capazes de interpretar e entender o mundo visual. Usa imagens digitais de câmeras e vídeos, junto a modelos de *deep learning*, com o objetivo de usar computadores para emular a visão humana, incluindo aprender, ser capaz de fazer inferências e tomar ações com base em entradas visuais [22].

As primeiras experiências em VC aconteceram nos anos 1950. Foram usadas as primeiras redes neurais para detectar os limites de um objeto e para classificar objetos simples em categorias como círculos e quadrados. Nos anos 1970, o primeiro uso comercial de VC interpretou textos manuscritos e digitados usando reconhecimento óptico de caracteres [23].

Para um computador, uma imagem é apenas uma matriz de valores. Normalmente, é uma matriz tridimensional (RGB) de valores de pixel. Cada pixel tem um valor específico

de vermelho, verde e azul que representa a cor de um determinado pixel. É inútil ler a imagem se o computador não entender o que significa ou se não conseguir descrever o que é a imagem e o que ela contém. Neste ponto, entra o aprendizado de máquina. A máquina é ensinada a entender a imagem e dizer o que ela contém [24].

Para treinar um modelo de aprendizagem de máquina para identificar um objeto, como uma maçã, por exemplo, dá-se a ele vários exemplos de imagens que contém uma maçã. Quando uma nova imagem de uma maçã é apresentada a este computador com esse modelo, ele pode usar o que aprendeu sobre maçãs anteriormente e identificar que essa nova imagem também contém maçã. Isso acontece da mesma forma com o rosto das pessoas para o reconhecimento facial [25].

De acordo com Szeliski [26], a diversidade de algoritmos de VC é extensa, e utilizada em uma variedade de aplicações do mundo real, como soluções para as áreas de segurança, direito e medicina. Nesse contexto, a abordagem deste trabalho tem por foco o uso de técnicas para reconhecimento facial.

2.2.1 Reconhecimento Facial

Reconhecimento Facial (*Face* ou *Facial Recognition*) é uma técnica de identificação biométrica (assim como a impressão digital) em que um software mapeia matematicamente os traços faciais de uma pessoa e, por meio de algoritmos, é capaz de compará-los a uma imagem digital do rosto dessa mesma pessoa, reconhecendo (ou negando) sua identidade [27].

A fase de mapeamento do rosto leva em conta os chamados pontos nodais — nome dado às características que fazem com que as pessoas tenham feições tão distintas umas das outras. Distância entre os olhos, largura do nariz, profundidade das órbitas oculares, forma das maçãs do rosto e comprimento da linha da mandíbula são alguns exemplos de pontos nodais usados pelo Reconhecimento Facial. Há cerca de 80 na face humana.

A medição de um rosto, ou seja, a relação entre esses pontos, cria uma geometria espacial única, que é armazenada em forma de dados (chamada de *template* ou *faceprint*). Quando uma nova imagem digital (que pode ser foto, vídeo ou captura ao vivo) é apresentada, o software faz a comparação.

Conforme Granatyr e da Silva [28], para que uma face possa ser reconhecida, são necessárias algumas etapas, como, por exemplo: coleta de faces, detecção de faces, detecção de pontos faciais, treinamento de uma rede neural convolucional (CNN) e reconhecimento facial.

A etapa de **coleta de faces** consiste na obtenção de faces para treinamento. Deve-se utilizar rostos com expressões faciais variadas. Quanto mais imagens coletadas, mais assertivo será o reconhecimento facial posteriormente.

Em seguida, realiza-se a **detecção de faces** na imagem. As faces detectadas são destacadas com um quadrado identificando o *bounding box* (região de interesse).

Existem diferentes métodos para detecção de faces. Como exemplo, pode-se citar o método de *Histogram of Oriented Gradients* (HOG) [29] [30] [31], geralmente utilizado pelo seu custo benefício em relação à acurácia e ao custo computacional necessário para a execução do algoritmo, menor do que o requisitado por uma CNN.

Para a **detecção de pontos faciais**, foi utilizado um modelo com 68 pontos de referência facial (rede chamada "shape_predictor_68_face_landmarks.dat", treinada com HOG e *Support Vector Machine*, SVM). Para que seja considerada uma face pelo algoritmo, os 68 pontos tem que ser identificados dentro da região de interesse. A Tabela 1 apresenta as regiões de mapeamento, e a Figura 1 exibe a localização dos pontos na face.

Tabela 1. Regiões da face considerada por algoritmo de mapeamento facial [32].

REGIÃO	PONTOS
Linha do queixo	01 a 17
Sobrancelha direita	18 a 22
Sobrancelha esquerda	23 a 27
Ponte nasal	28 a 31
Nariz	32 a 36
Olho esquerdo	37 a 42
Olho direito	43 a 48
Lábio externo	49 a 60
Lábio interno	61 a 68

Para efetuar o **treinamento da rede**, cada imagem só pode ter uma face, para que o algoritmo mantenha a sua assertividade. O reconhecimento facial, nesta etapa, usa uma CNN pré-treinada para identificar os pontos da face, retornando um vetor numérico de 128 posições, com as características do rosto de cada imagem. O treinamento gera arquivos com características das imagens para o reconhecimento, que são utilizadas na etapa seguinte.

Por fim, na etapa de **reconhecimento facial**, efetua-se a comparação da imagem a ser reconhecida com os arquivo da rede pré-treinada, utilizando o algoritmo kNN (vizinho mais próximo). Este, por sua vez, faz uso da distância Euclidiana para retornar o menor intervalo em relação à imagem alvo, sendo este relacionado a imagem "escolhida".

No definição do algoritmo, define-se o limiar (distância máxima) que será utilizado para que a imagem seja reconhecida. Se o resultado for 0, a imagem reconhecida é igual à imagem utilizada no treinamento.

Quanto menor a imagem ou quando mais "distante" estiver a face em uma foto, mais difícil será o algoritmo conseguir reconhecer a pessoa.

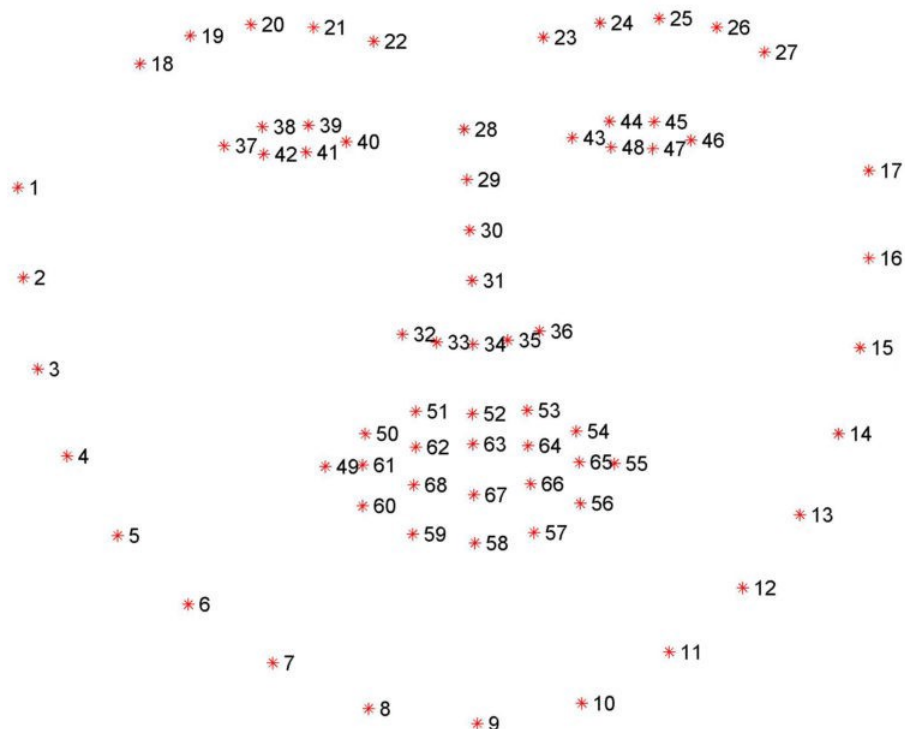


Figura 1. Pontos identificados na face. Fonte: José [32].

2.3 TRABALHOS RELACIONADOS

Arapakis et al. [33] propõem uma nova interface de pesquisa de vídeo que prevê a relevância do tópico de um vídeo por meio da análise de aspectos afetivos do comportamento do usuário. Além disso, apresentam um método para incorporar tais características afetivas no perfil do usuário, para facilitar a geração de recomendações significativas, de vídeos não vistos. No experimento, mostram que o recurso de interação multimodal é uma forma promissora de melhorar o desempenho da recomendação.

Além disso, os autores apresentam uma forma de explorar essas informações para classificar a relevância temática dos vídeos vistos, com o auxílio de uma SVM, e, eventualmente, enriquecer os perfis dos usuários. O valor da interface criada está na combinação de diferentes módulos (sistema de reconhecimento de expressão facial, SR, etc.), na integração de dados sensoriais e, finalmente, na aplicação de fusão de informações. Os autores citam trabalho semelhante, publicado por Yeasin et. al, que aplicou o reconhecimento de expressão facial para identificar seis expressões faciais universais a partir de sequências de vídeo e mediu os níveis de interesse com base em um espaço afetivo de três minutos. De acordo com os autores, teóricos da emoção discreta sugerem a existência de seis ou mais emoções básicas (felicidade, tristeza, raiva, medo, nojo e surpresa) universalmente exibidas e reconhecidas.

No trabalho supracitado, foi utilizado o eMotion, um sistema de reconhecimento de expressão facial que utiliza as emoções básicas. A pesquisa indica que as emoções são

comunicadas principalmente por meio de expressões faciais ao invés de gestos corporais e fornecem pistas faciais (sorrisos, risos, carranca, etc.), considerados um aspecto essencial de nossa interação social. Os sistemas automáticos são uma abordagem alternativa para a análise da expressão facial e têm apresentado desempenho comparável (sob condições controladas) ao reconhecimento humano treinado, que atinge a precisão de 87% [34]. O eMotion aplica um classificador genérico que foi treinado em um conjunto de dados diversificado, combinando dados do banco de dados Cohn-Kanade [35].

Para o estudo, os autores requisitaram a colaboração de 24 pessoas com perfis diferentes para participarem da avaliação. Eles utilizaram dois computadores para o teste: um deles contendo teclado, mouse e uma câmera; e o outro contendo o SR, o modelo SVM, o sistema de reconhecimento de expressões faciais (eMotion) e o software gravador de vídeo.

Para os testes, foi formulado um conjunto de tarefas de pesquisa que diferiam em seu domínio e escopo. Para cada tarefa de pesquisa, os participantes tinham a possibilidade de selecionar, entre uma lista predefinida de opções, o subtema de sua escolha. Para a conclusão das tarefas de pesquisa, foi utilizada uma interface de pesquisa de vídeo personalizada, que funcionava em cima do mecanismo de pesquisa do YouTube, projetada para se parecer com seu layout básico e manter um número mínimo de elementos gráficos. Os participantes puderam facilmente selecionar e visualizar qualquer um dos cliques exibidos. O conteúdo de um clipe foi mostrado em um painel separado. Um feedback é considerado positivo se o usuário encontrar pelo menos um quadro interessante durante o tempo em que está assistindo a um vídeo.

O valor do sistema proposto reside na combinação de diferentes módulos e modalidades, bem como na integração perfeita de componentes afetivos no perfil do usuário. Além disso, foi apresentada uma forma de processar essas informações, a fim de determinar a relevância dos vídeos assistidos e gerar recomendações significativas. Segundo os autores, o sistema é realisticamente aplicável; foi implementado usando uma câmera web de baixo custo e um navegador web padrão, que foi modificado para se comunicar com um sistema de reconhecimento de expressão facial.

Já a abordagem de **Wu et al. [36]** consiste em um SR integrado de reconhecimento facial para lidar com o problema de criação de perfil em uma loja de varejo. De acordo com os autores, os SRs que sugerem produtos, conteúdos ou recursos de aprendizagem aos usuários, têm se tornado um assunto importante e recebido muitos estudos. Os SR existentes utilizam diferentes fontes de conhecimento para traçar o perfil dos usuários, mas em alguns campos de aplicação, como varejo, obter tal fonte pode ser diferente.

Alguns varejistas começaram a utilizar tecnologias de reconhecimento facial para extrair dados demográficos, como sexo e idade, dos clientes em suas lojas físicas. Ainda segundo os autores, existem poucos estudos considerando o efeito da combinação de tecnologias de reconhecimento facial com SR. Neste trabalho, os autores propuseram um SR

que integra o reconhecimento facial para obter informações demográficas em uma loja de varejo.

A filtragem colaborativa (FC), utilizada no estudo dos autores, estima a similaridade entre dois clientes com base em seus registros de consumo, caminhos de navegação na web ou fluxos de cliques de aplicativos. No entanto, confiar em uma grande quantidade de registros do usuário leva ao conhecido problema de inicialização a frio da FC. No sistema, após os processos de reconhecimento facial e identificação do cliente, um cliente frequente, que possua registros suficientes no banco de dados, receberá uma lista de recomendações gerada pela FC. Para um cliente que é estranho ou desconhecido para o sistema, foi adotada a filtragem demográfica para amenizar o problema de inicialização a frio da FC. O recomendador de filtragem demográfica neste sistema fornece recomendações com base em informações demográficas de clientes inferidas pelo processo de reconhecimento facial. Essas informações demográficas são usadas para dividir os clientes em vários grupos. Primeiramente, o sistema atribui um cliente estranho ou desconhecido a um desses grupos e, em seguida, recomenda produtos populares deste grupo a ele. Todas as respostas de recomendações de clientes são coletadas para criar perfis adicionais de clientes e melhorar os resultados das recomendações.

Os autores buscaram avaliar o recomendador de filtragem demográfica do sistema na área de varejo. Existe muito código aberto de reconhecimento facial, e a maioria deles oferece a função de classificação de gênero e idade. Foi selecionado um aplicativo gratuito chamado *CrowdSight* [37] para aplicar o reconhecimento facial. Para construir o banco de dados de perfil do cliente inicial e o banco de dados de seu comportamento, foi utilizado o banco de dados de um varejista de produtos eletrônicos. As transações efetuadas pelos clientes envolviam telefones celulares. O experimento foi conduzido com um grupo de assuntos. Cada assunto foi classificado pelo aplicativo *CrowdSight* primeiro, e o recomendador produziu uma lista de recomendações com base no resultado da análise. A precisão da classificação de gênero é 92,31%, enquanto a precisão da classificação de idade é 61,54%. Este é um resultado razoável, pois a dificuldade de classificação por idade é muito maior.

Para comparação, também foi produzida outra lista de recomendações que foi gerada pela escolha de telefones celulares populares de diferentes marcas em 2013. Cada participante foi convidado a votar em uma lista de recomendações que eles provavelmente comprariam. Houve 69,23% dos sujeitos votando na lista produzida pelo sistema proposto, o que significa que o SR faz recomendações de maneira eficaz.

Conforme os autores, os varejistas podem criar perfis de clientes individuais, aproveitando informações relevantes capturadas da Internet, de sensores e de outros dispositivos. Neste estudo, fez-se a integração de um SR com reconhecimento facial. O resultado da avaliação mostrou que a tecnologia de reconhecimento facial melhorou os resultados das recomendações e, portanto, ajudou os varejistas a aprimorar seu marketing. O sistema

criado demonstrou ter um bom desempenho e pode aliviar o problema de inicialização a frio de forma satisfatória.

Por outro lado, **Alfian et al. [3]** desenvolveram um totem digital (DSOS - Digital Signage Online Store) e o instalaram em locais diferentes, para que os clientes pudessem ter a experiência de navegar e comprar um produto através de um menu em uma tela sensível ao toque, com imagens dos produtos em 3D para interação. O sistema de processamento de dados em tempo real coletou o histórico de navegação dos clientes e os dados de transações, conforme ocorriam. Além disso, os autores utilizaram a regra de associação para extrair informações úteis do comportamento do usuários, de modo que ela pudesse ser usada pelos gerentes para aprimorar, de maneira eficiente, a qualidade do serviço.

Os autores utilizaram a regra de associação para extrair informações úteis do comportamento do cliente. Esses resultados podem ser usados para promoção, bem como recomendações de produtos relevantes para clientes DSOS. Além disso, no atual ambiente de varejo, em constante mudança, a análise do comportamento do cliente em tempo real na DSOS ajuda a atrair e reter os clientes com mais eficiência e eficácia, e os varejistas podem obter uma vantagem competitiva sobre seus concorrentes.

O estudo da análise de comportamento de compra de clientes foi utilizado para promover produtos ecológicos. O resultado obtido ofereceu uma estratégia importante que pode ajudar os varejistas a otimizar a promoção deste tipo de produto, bem como criar um ambiente interessante nas lojas. Dessa forma, pode-se melhorar a experiência de compra e a satisfação do cliente, bem como a rentabilidade por cliente.

Com o aumento do número de SD e de clientes, o sistema proposto foi capaz de processar convenientemente uma grande quantidade de dados de entrada. O conjunto de dados mostrou que, à medida que o número de visitas e a duração das compras aumentaram, a chance dos produtos serem comprados também aumentou. Combinando dados de compra e navegação de clientes, as regras de associação do padrão de transação frequente foram alcançadas. Assim, os produtos apresentaram uma alta possibilidade de serem comprados se forem usados como recomendações.

O fornecimento de imagens 3D interativas aumentou a confiança e o prazer do cliente ao interagir com os produtos. O cliente pode ter uma visão melhor dos produtos, o que impactou positivamente seu comportamento de compra, aumentando, inclusive, a frequência das visitas aos estabelecimentos. Os resultados obtidos puderam ser usados tanto para a criação de promoções, quanto para a recomendação de produtos relevantes para os clientes DSOS.

Como conclusão, os autores relataram que, para os varejistas aumentarem a satisfação do cliente, a DSOS pode ser usada para enriquecer a experiência de compra dos clientes. O estudo propôs a análise do comportamento do cliente com base no processamento de dados em tempo real. A tecnologia de *Big Data*, como NoSQL MongoDB e Apache Kafka, foi utilizada para este propósito. O sistema proposto monitorava o histórico

de navegação do cliente, bem como a aquisição de dados da DSOS e o utilizava como entrada para a tomada de decisões.

Além disso, a regra de associação proposta foi usada para extrair informações úteis do comportamento do cliente e dados de compra e pode ser utilizada para recomendação de produto na DSOS, ajudando os clientes a obter o produto pretendido e, para os lojistas, aumentando a receita de vendas da empresa no futuro.

Diferentemente dos artigos listados acima, o presente trabalho busca fornecer uma solução capaz de sugerir recomendações em tempo real para o lojista, com base no campo visual do cliente durante uma compra. Desta forma, o lojista pode ofertar produtos ao cliente assim que o sistema detectar o interesse deste, realizando um atendimento mais personalizado e, potencialmente, agregando valor à venda. Para tanto, a visão geral da arquitetura, os materiais e métodos utilizados, e o funcionamento da aplicação proposta são descritos no próximo Capítulo.

3. MATERIAIS E MÉTODOS

Os materiais e métodos apresentados nesta seção foram aplicados para criar um sistema de recomendação, como prova de conceito, chamado **Recommender Vision**, que utiliza reconhecimento facial e de padrões para fornecer sugestões de venda de produtos para os lojistas. O sistema tem por objetivo identificar produtos de interesse dos clientes em compra em uma loja física, através do monitoramento de imagens de vídeo, e repassar esta informação a um lojista para ofertar o referido produto ao cliente, buscando aumentar o valor da venda.

Para a obtenção das imagens, foi utilizada uma câmera para monitorar duas prateleiras, uma acima e outra abaixo do dispositivo. A câmera ficou direcionada para a frente, visando identificar a face do cliente, para obter a posição onde ele está olhando e definir o produto de interesse do mesmo.

A recomendação baseada em conteúdo foi o método selecionado para este trabalho, uma vez que as imagens capturadas são analisadas para detectar os itens de preferência do cliente. O histórico de interesses visuais anteriores também é utilizado, caso já exista registro do cliente na base de conhecimento.

Para validar preliminarmente os recursos de software e hardware, e também em virtude da pandemia Covid-19, optou-se em testar a aplicação considerando um conjunto restrito de pessoas e objetos, em ambiente controlado.

3.1 SOFTWARE E HARDWARE

Para o desenvolvimento da aplicação, optou-se pelo uso da linguagem Python e recursos de bibliotecas auxiliares para as funcionalidades referentes à visão computacional, sendo *OpenCV* (<http://www.opencv.org>) para algoritmos de visão de computacional; *Face Recognition* (https://github.com/ageitgey/face_recognition) para detecção da face da pessoa na imagem; e *DLib* (<http://www.dlib.net>) para obter a direção para onde a pessoa está olhando.

Os requisitos mínimos para executar a aplicação são: um computador com processador Intel i3 ou superior, 2 GB de memória RAM, 200 MB de espaço em disco e uma câmera de vídeo, com resolução mínima de 640x480.

3.2 VISÃO DA ARQUITETURA DO SISTEMA

O presente trabalho consistiu no desenvolvimento de uma aplicação para identificar o comportamento dos clientes em compra em uma loja física, com a utilização de visão

computacional, buscando recomendar itens de interesse do comprador. Seu diagrama esquemático encontra-se na Figura 2.

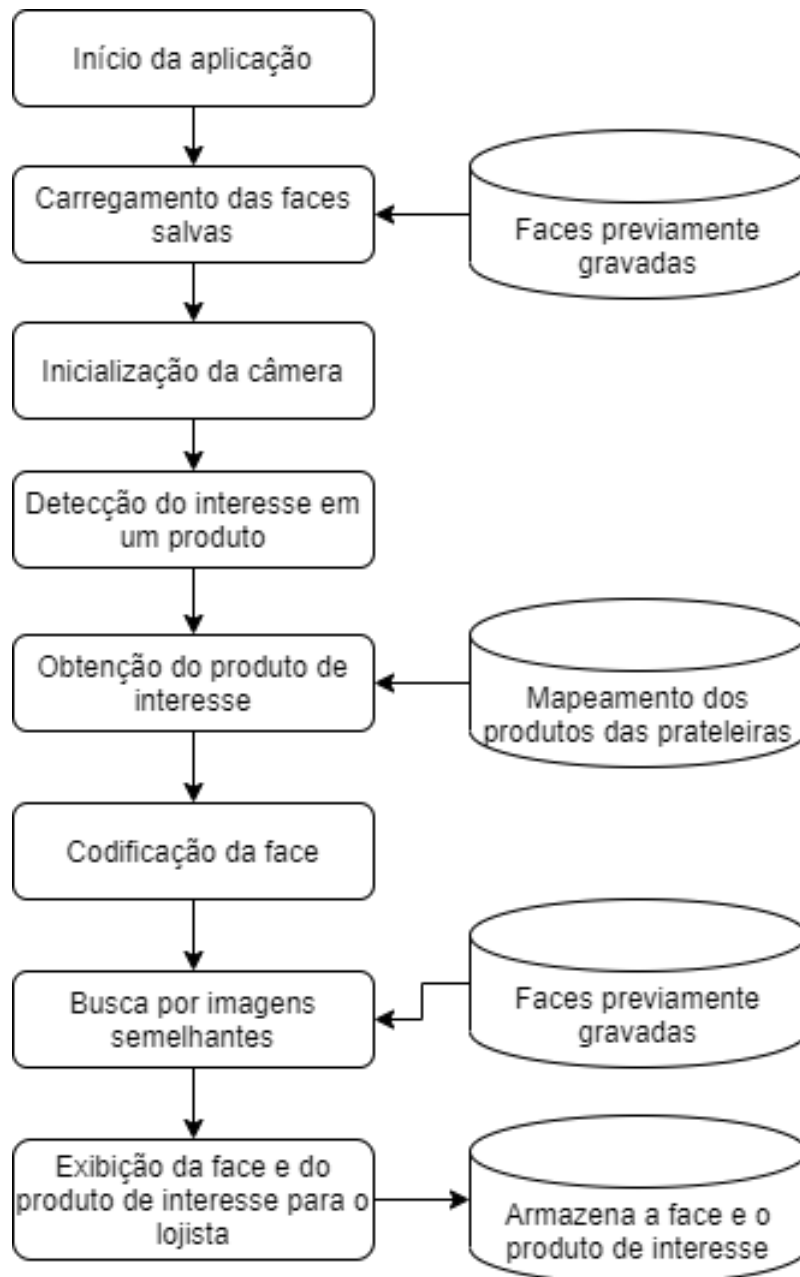


Figura 2. Diagrama esquemático da aplicação. Fonte: o autor

Ao iniciar a aplicação, as faces previamente reconhecidas são carregadas para a memória, juntamente com o produto de interesse do cliente, associado a cada arquivo.

Em seguida, a câmera é inicializada e a imagem da mesma é exibida para o lojista. Para a obtenção da imagem, é utilizado o endereço IP da câmera. A aplicação fica em *loop*, verificando se existe uma face na imagem. Ao detectar o interesse do cliente em um determinado produto, a aplicação identifica o item que foi observado e efetua a codificação da face do cliente (preservando sua identidade visual), e salvando estas informações em disco. Em seguida, é efetuada a busca e a comparação por registros codificados simila-

res, para identificar se a face recentemente codificada já não foi identificada anteriormente. Caso encontre registros, os últimos produtos de interesse deste cliente são retornados e exibidos na tela para o lojista (Figura 4).

3.3 FUNCIONAMENTO DA APLICAÇÃO

Ao iniciar, a aplicação carrega todos os arquivos codificados e os armazena em memória, para que o reconhecimento facial não necessite aguardar a leitura destes arquivos para efetuar o processamento das faces. Da mesma forma, todas as configurações são carregadas e armazenadas na memória.

A câmera é, então, inicializada e sua imagem fica sendo exibida em *loop* para o lojista, no computador centralizador. A cada *frame* lido, a aplicação verifica se existe uma face na imagem, obtendo os seus pontos faciais, conforme Figura 1. Para a definição dos pontos faciais, é utilizado o modelo *shape_predictor_68_face_landmarks.dat*, da biblioteca Dlib. Se existir, obtém o produto que está sendo observado.

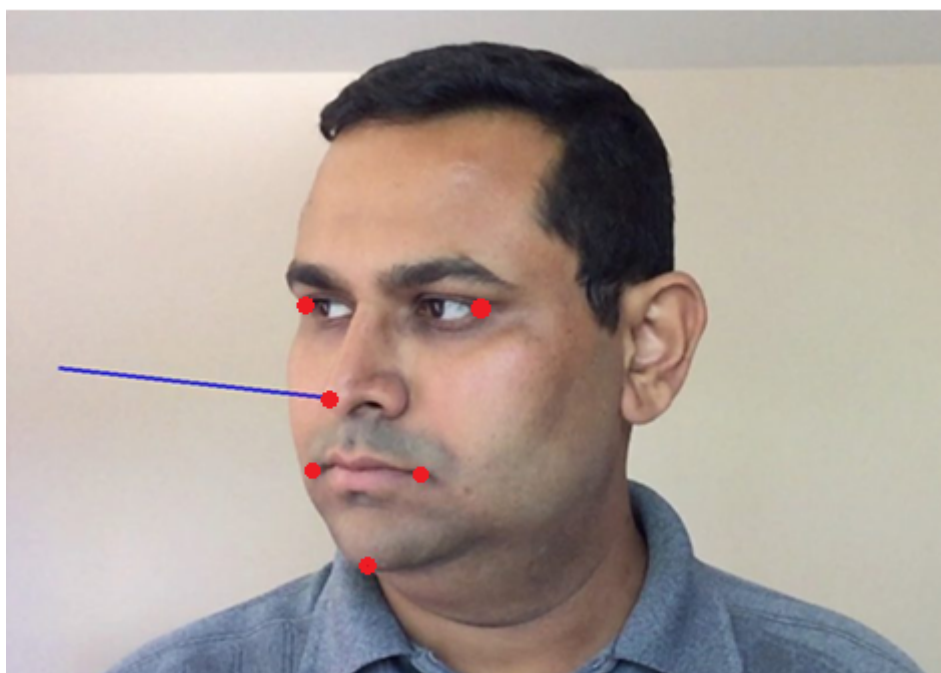


Figura 3. Exemplo da obtenção da posição estimada da cabeça. Fonte: <https://www.learnopencv.com/head-pose-estimation-using-opencv-and-dlib>

Para auxiliar na identificação dos produtos, o software possui um arquivo de configuração com o mapeamento dos produtos de cada prateleira (Tabela 3). Neste arquivo, são incluídos, manualmente e na mesma ordem existente na prateleira, o nome do produto e o espaço em centímetros que o mesmo ocupa. Esta medida deve ser feita a partir da posição onde a câmera se encontra, contendo duas seções, esquerda e direita, para cada

prateleira. Qualquer mudança na ordem dos produtos ou na posição da câmera requer a reconfiguração deste mapeamento.

Para identificar o produto que o cliente está olhando, inicialmente são obtidos os ângulos horizontal e vertical, usando como base o ponto facial que representa o nariz da pessoa.

Para isto, é utilizado um *array* com alguns parâmetros de acordo com a câmera que obteve a imagem, um modelo 3D padrão pré-definido de uma face, e os pontos faciais obtidos do *frame* que está sendo processado. Após, são obtidos dois vetores com a rotação e a translação da face em relação à imagem. Para a obtenção do ângulos de Euler, que representam a orientação da face, é feito o cálculo da distância Euclidiana do vetor de rotação. A distância obtida é utilizada como parâmetro para calcular os ângulos *roll-pitch-yaw* (horizontal, vertical e inclinação) [38].

Com os ângulos obtidos, identifica-se, primeiro, a prateleira que está sendo observada, comparando-se o ângulo vertical com o valor que está informado no arquivo de configurações. A seguir, com base na distância entre a pessoa e a câmera (cateto adjacente) e no ângulo horizontal, é calculada a distância lateral (cateto oposto) entre a câmera e o produto, através das relações trigonométricas do triângulo retângulo (Equação 1).

$$\text{seno}(\text{angulo}) = \frac{\text{cateto oposto}}{\text{cateto adjacente}} \quad (1)$$

A obtenção da distância da pessoa até a câmera é feita utilizando as dimensões da imagem, a distância focal da câmera e a altura da face. Para isso, é necessário efetuar a calibragem da câmera. A forma utilizada foi através do uso de uma folha A4 em branco a uma distância pré-determinada, utilizando um marcador de objetos [39].

Toda vez que a posição da câmera mudar ou um modelo diferente de câmera for utilizado, a distância focal deve ser recalculada e ajustada no arquivo de parâmetros da aplicação. Tais informações são relevantes para o cálculo da distância entre a face e o dispositivo.

A aplicação captura um *frame* a cada segundo, a partir do momento em que identifica o olhar do cliente para uma das prateleiras. O produto observado no primeiro segundo será comparado com o observado nos próximos n segundos (definido em arquivo de configuração). Se nos próximos segundos o produto visualizado for o mesmo, isto confirmará o possível interesse do cliente no item. Caso o cliente passe a observar um produto diferente antes do último segundo, o processo se repete a partir do novo produto visualizado, reiniciando a contagem de segundos. De acordo com Gidlöf *et al.* [40], considerando os últimos cinco segundos anteriores a uma escolha, existe uma tendência maior no comportamento visual de uma pessoa cerca de dois segundos antes dela selecionar o produto que finalmente será escolhido.

Quando o interesse do cliente em um item é efetivado (i.e., observou o mesmo item pelo número de segundos pré-definido), sua face é codificada e salva em disco, juntamente com a data, hora e o produto alvo. Em seguida, é efetuada uma comparação com faces similares codificadas anteriormente - ou seja, verifica se determinada pessoa já esteve (ou não) no estabelecimento previamente e demonstrou interesse em algum item. Caso encontre registros, os últimos m produtos de interesse deste cliente são retornados e exibidos em uma tela de controle para o lojista, com as respectivas datas dos registros, conforme demonstrado na Figura 4. Desta forma, o mesmo pode tomar a decisão de oferecer o produto ou seus similares ao cliente. A definição de quantos registros serão mostrados também está contida no arquivo de configuração.

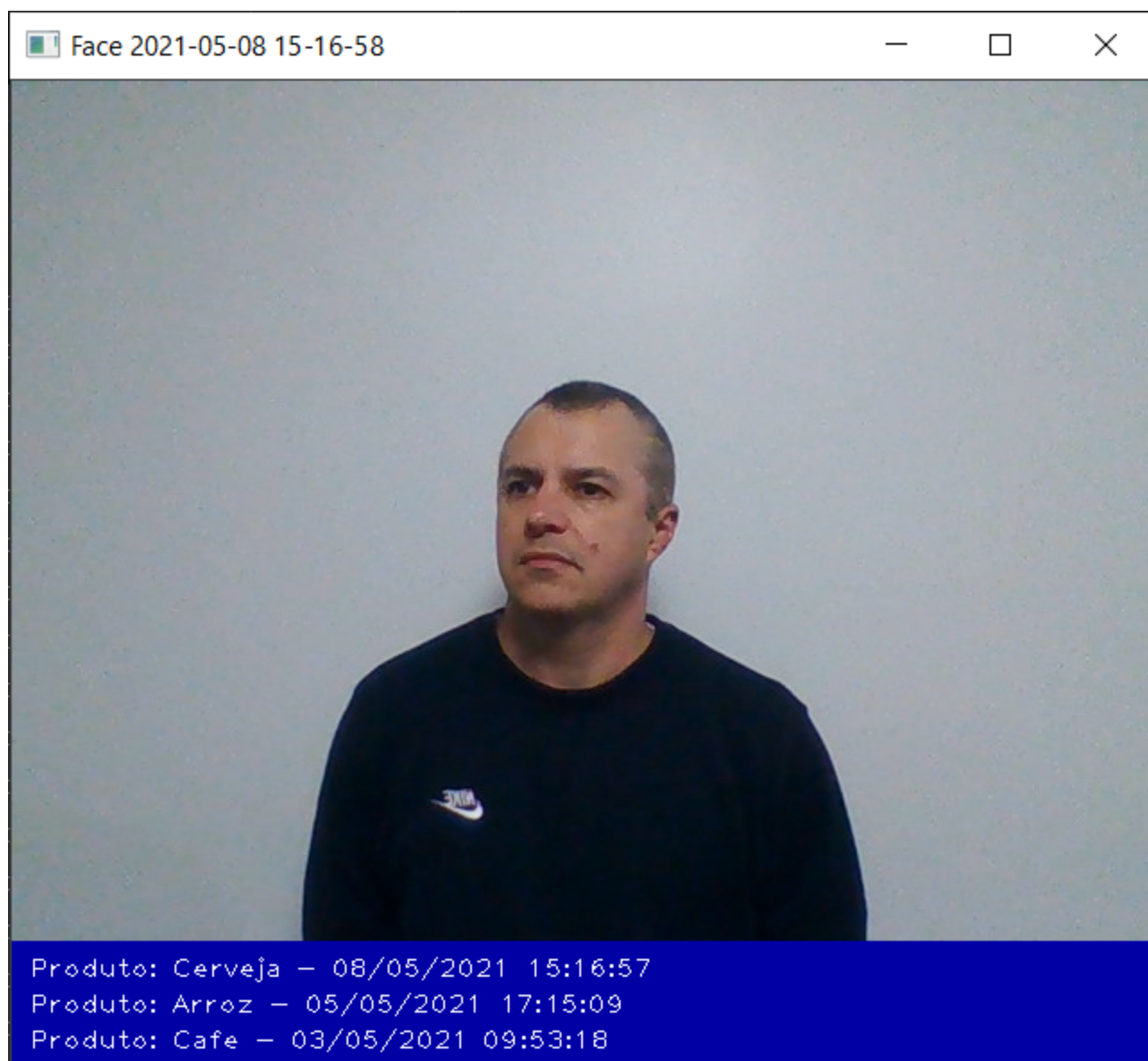


Figura 4. Exibição do registro da face, produtos de interesse associados e respectivas datas de visualização. Fonte: o autor

O reconhecimento do cliente é efetuado sem que o mesmo seja identificado ou rotulado na base de dados, ou seja, o cliente não é nominado. Nesse contexto, a recomendação apenas considera a exibição de seu rosto e a indicação do item de interesse. Computacionalmente, existe apenas o registro por correspondência facial e o armazenamento dessas características no histórico do sistema, conforme Tabela 2, preservando sua identidade e considerando o Art. 11, inciso 2^o, alínea c da LGPD [14], que permite a utilização de dados sensíveis, sem o fornecimento de consentimento do titular para a realização de estudos, garantida, sempre que possível, a anonimização dos dados pessoais.

Tabela 2. Arquivo com os dados da captura e a face codificada

```

Filename: Face 2021-05-08 15-16-20.dat
Product: feijao
Face: [-0.08790072, 0.07978898, 0.00612106, -0.00551964, -0.07435261,
      -0.02481468, -0.02918948, -0.13523559, 0.1721967, -0.00097598,
      0.24708468, -0.01082398, -0.28301871, -0.05968395, -0.00335517,
      0.07745624, -0.15307772, -0.06374702, -0.1143828, -0.11789727,
      0.07585815, 0.06259202, 0.03190638, 0.03194787, -0.12563734,
      -0.29657772, -0.15479407, -0.13536167, 0.04666714, -0.10529239,
      0.0647224, 0.08878054, -0.14254032, -0.12092495, 0.06081695,
      0.00187417, -0.02280075, 0.03142776, 0.16625805, 0.0014745,
      -0.09150626, 0.06799706, 0.06705318, 0.31104404, 0.16557792,
      0.12654546, 0.03379099, -0.13153413, 0.07311518, -0.26107949,
      0.10352275, 0.19118796, 0.16263165, 0.07836154, 0.17463289,
      -0.19301197, 0.04330783, 0.07486887, -0.25381082, 0.13362409,
      0.10162593, -0.00436864, -0.01744764, -0.05368726, 0.16294007,
      0.10060201, -0.10620803, -0.15587036, 0.12545082, -0.17797422,
      -0.0893226, 0.14009696, -0.15013151, -0.20643575, -0.36988792,
      -0.00180737, 0.41289395, 0.16629003, -0.23283269, -0.0225064,
      -0.0858716, -0.01777866, 0.04985348, -0.01007843, -0.09342203,
      -0.08358538, -0.12210594, 0.02115355, 0.13069859, -0.01087623,
      -0.05640429, 0.2535021, 0.05759512, -0.010099, 0.02484501,
      0.11287101, -0.12928395, -0.08187023, -0.13358793, -0.05432684,
      -0.00365055, -0.02887814, 0.00379148, 0.10962277, -0.18040314,
      0.23001957, -0.03126309, -0.05287267, -0.1289905, -0.02355306,
      -0.0813248, 0.07086912, 0.19806305, -0.32003406, 0.1977516,
      0.17854978, -0.00761799, 0.10904342, -0.04781976, 0.05761618,
      -0.09999564, -0.03691645, -0.15145488, -0.12758856, -0.02877541,
      -0.06366588, -0.00456149, 0.03114138]

```

Nesta versão, a aplicação manipula apenas uma face por vez. Desta forma, quando existir mais de uma pessoa na mesma cena, apenas o rosto mais próximo da câmera será considerado no monitoramento.

3.4 AMBIENTE DE TESTE

A ideia inicial para validar o sistema era instalar o mesmo em duas lojas físicas. Em virtude da pandemia e restrições associadas, optou-se por efetuar a validação em ambiente de laboratório, já que, para a captura e reconhecimento facial, a solução considera a identificação de diferentes pontos faciais e, com as pessoas usando máscara, não seria possível obter os pontos da face.

Para que seja cumprido o objetivo da aplicação, a câmera deve ser instalada de maneira que seu foco esteja em direção ao rosto dos clientes (Figura 5). No ambiente de teste utilizado, a câmera foi instalada entre as prateleiras superior e inferior, centralizada, como ilustrado pela Figura 6. No entanto, cabe ressaltar que a câmera pode ser fixada em qualquer posição, conforme a preferência do lojista.



Figura 5. Desenho da posição de instalação da câmera para monitoramento da prateleira. Fonte: o autor

Para a primeira versão da abordagem proposta, e para fins de validação, a aplicação considera um cadastro com o mapeamento pré-definido dos produtos da prateleira em monitoramento, conforme Tabela 3.

As prateleiras utilizadas no ambiente de teste (Figura 6) seguiram o mapeamento contido no arquivo de configuração (Tabela 3). Para representar os diferentes produtos expostos na prateleira do ambiente de teste, foram impressos e fixados rótulos em uma superfície branca, simulando o cenário com os produtos colocados lado a lado. Tanto a prateleira superior quanto a inferior possuíam um metro de comprimento.



Figura 6. Prateleira modelo utilizada no ambiente de experimento. Fonte: o autor

Tabela 3. Arquivo de configuração

[General]

url_cam = 0

active_debug = False

points_landmark = 68

seconds_to_interest = 3

tolerance_for_face_recognition = 60

view_last_products = 5

[Files]

shape_predictor_68_face_landmarks_dat = ./lib/models/shape_predictor_68_face_landmarks.dat

[ShelfAngle]

shelf1 = 15

shelf2 = 10

[ShelfLevel1 Left]

cerveja = 40

fralda = 10

[ShelfLevel1 Right]

fralda = -50

[ShelfLevel2Left]

arroz = 25

feijao = 25

[ShelfLevel2Right]

acucar = -25

cafe = -10

pipoca = -15

4. RESULTADOS E DISCUSSÕES

No presente Capítulo são apresentados os resultados e as discussões de uma avaliação preliminar realizada com a aplicação proposta neste trabalho. Os testes do estudo piloto foram conduzidos em ambiente de laboratório, considerando o uso da solução com câmera em um computador sem GPU - o que simula uma situação comum encontrada em lojas físicas.

4.1 RESULTADOS DO EXPERIMENTO

O estudo piloto foi conduzido com a participação de 15 voluntários, nove mulheres e seis homens (idade $48,5 \pm 13,8$, altura $172,4 \pm 7,3$), conforme Tabela 4. Cada pessoa tinha a tarefa de ficar em pé em frente à câmera, a uma distância de 70 cm, de forma a visualizar todos os produtos das prateleiras. A câmera estava posicionada a 1,55 m de altura. Utilizando o protocolo verbal *Think Aloud*, foi solicitado ao participante que, ao observar cada produto, era necessário expressar em voz alta o item que estava visualizando.

Tabela 4. Dados das pessoas participantes no experimento

ID	Gênero	Altura (m)	Idade
#01	Feminino	1,58	56
#02	Feminino	1,63	68
#03	Feminino	1,65	43
#04	Feminino	1,65	27
#05	Feminino	1,68	55
#06	Masculino	1,71	46
#07	Feminino	1,72	72
#08	Masculino	1,72	42
#09	Feminino	1,74	65
#10	Feminino	1,76	44
#11	Feminino	1,76	30
#12	Masculino	1,78	26
#13	Masculino	1,82	58
#14	Masculino	1,82	53
#15	Masculino	1,84	43

O participante deveria olhar, primeiramente, cada produto disposto na prateleira superior, na ordem da esquerda para a direita. Em seguida, o mesmo procedimento foi requisitado para a prateleira inferior. Em outro computador, o vídeo da câmera estava sendo gravado, para possibilitar a análise dos resultados. O objetivo deste teste foi avaliar se o sistema estava capturando todos os objetos de interesse, e se estava detectando e registrando cada participante no banco de dados.

Os dados deste experimento estão listados na Tabela 5. O percentual geral de produtos detectados corretamente foi de 74,3%. Para pessoas de altura menor que 1,72 m, a aplicação identificou o produto de interesse em 80,4% das vezes, enquanto que para pessoas maiores que esta altura o valor ficou em 67,3%. Homens mais altos que 1,81 m tiveram o menor percentual de identificação, com 46,4%, em virtude da configuração vertical utilizada para delimitar as prateleiras, o que fez com que a visualização da maioria dos itens não fosse detectada.

Em relação às prateleiras, a superior, com os produtos ocupando um espaço maior, teve um acerto de 96,7%. Já, na inferior, os produtos identificados representaram 65,3%. Produtos que ocupavam espaço de 25 cm ou menos (café e pipoca), foram detectados apenas em 46,7% das vezes, enquanto os itens que utilizaram espaços maiores tiveram um percentual de identificação de 85,3%. O produto localizado no extremo direito da prateleira inferior foi detectado em 53,3% das vezes, em virtude de ocupar um espaço menor que os demais e por estar numa posição com ângulo maior que 30° em relação ao observador - uma limitação da biblioteca DLib, descrita na Seção 4.2.

Tabela 5. Resultado da visualização dos produtos

ID	Prateleira Superior		Prateleira Inferior				
	Cerveja	Fralda	Arroz	Feijão	Açúcar	Café	Pipoca
#01	OK	OK	OK	OK	OK	OK	OK
#02	OK	OK	OK	OK	OK	OK	-
#03	OK	OK	OK	OK	OK	-	-
#04	-	OK	-	OK	OK	-	-
#05	OK	OK	OK	OK	OK	-	-
#06	OK	OK	OK	OK	OK	OK	OK
#07	OK	OK	OK	OK	OK	-	OK
#08	OK	OK	OK	OK	OK	-	OK
#09	OK	OK	OK	OK	OK	OK	OK
#10	OK	OK	OK	OK	OK	-	OK
#11	OK	OK	OK	OK	OK	OK	OK
#12	OK	OK	OK	OK	OK	OK	OK
#13	OK	OK	-	-	-	-	-
#14	OK	OK	-	-	-	-	-
#15	OK	OK	-	-	-	-	-
Taxa de Acerto:	93,3%	100,0%	73,3%	80,0%	80,0%	40,0%	53,3%

O experimento verificou, também, se as informações processadas pela aplicação estavam sendo registradas e exibidas corretamente ao operador. Primeiramente, a aplicação registrou somente as faces dos 15 participantes do experimento, única e exclusivamente, não gerando falsos positivos, tampouco falsos negativos. Já sobre a exibição dos produtos visualizados pelos participantes, a aplicação também funcionou corretamente, pois foram listados somente os itens identificados, na ordem em que foram visualizados, considerando data e horário, do mais recente para o mais antigo.

Outra situação avaliada foi o desempenho para efetuar o registro da face, que se mostrou rápida para a abordagem proposta. Esta operação consistia na codificação da face capturada, transformando suas características em um *array* de 128 posições. A seguir, era feita a verificação da similaridade desta face com as demais salvas anteriormente (e previamente carregadas na memória). O fator utilizado foi de 60% de similaridade, conforme recomendação da biblioteca *Face Recognition*.

As faces selecionadas nesta etapa foram armazenadas em um novo *array*, que continha os dados a serem exibidos para o lojista (data, hora e produto). A face capturada e codificada, que gerou esta operação, era salva em disco e adicionada neste *array*, o qual, então, era ordenado por data e hora, de forma decrescente. Foram obtidos os primeiros n elementos deste *array*, conforme a quantidade de produtos a serem exibidos para o lojista, definido no arquivo de configurações. Estes elementos continham os últimos produtos observados pela face que foi capturada, bem como a data e a hora da visualização.

De posse destas informações, a aplicação exibia uma janela para o operador, contendo a foto da pessoa e a lista com o nome, data e hora dos últimos produtos visualizados (Figura 4). O processo descrito demorou, em média, 1,7s para ser executado. A comparação, na situação relatada, foi feita com 500 registros. Uma vez que todas estas informações estão na memória, o crescimento exponencial de registros não aumenta o tempo de execução. Para certificar-se disso, foram feitos testes adicionais com 2500 e 5000 faces, e o tempo médio se manteve conforme citado.

Além dos testes de observação dos produtos e desempenho, também foram feitos testes de oclusão. A aplicação sempre identifica a face que está mais próxima à câmera, como mostra a Figura 7. Durante os testes, a aplicação se comportou como esperado.

Adicionalmente, considerou-se também casos em que ocorria a interrupção na identificação da face por parte do sistema como, por exemplo, algo ou alguém obstruindo temporariamente a face em monitoramento. Para estas situações, definiu-se previamente que o cronômetro que determina o tempo de observação do cliente seria reiniciado, e que não haveria registro. Adotou-se esse comportamento pois entendeu-se que tal situação poderia gerar um desvio de atenção do observador. Durante esses rápidos bloqueios nos testes, a aplicação seguiu as regras que haviam sido estabelecidas.

4.2 DISCUSSÕES

Baseado no estudo piloto, notou-se que a aplicação identificou corretamente os produtos em 74,3% das vezes. Objetos com espaço maior foram detectados facilmente (cerveja e fralda) e objetos com tamanho mediano tiveram um percentual de detecção de 85,3%.

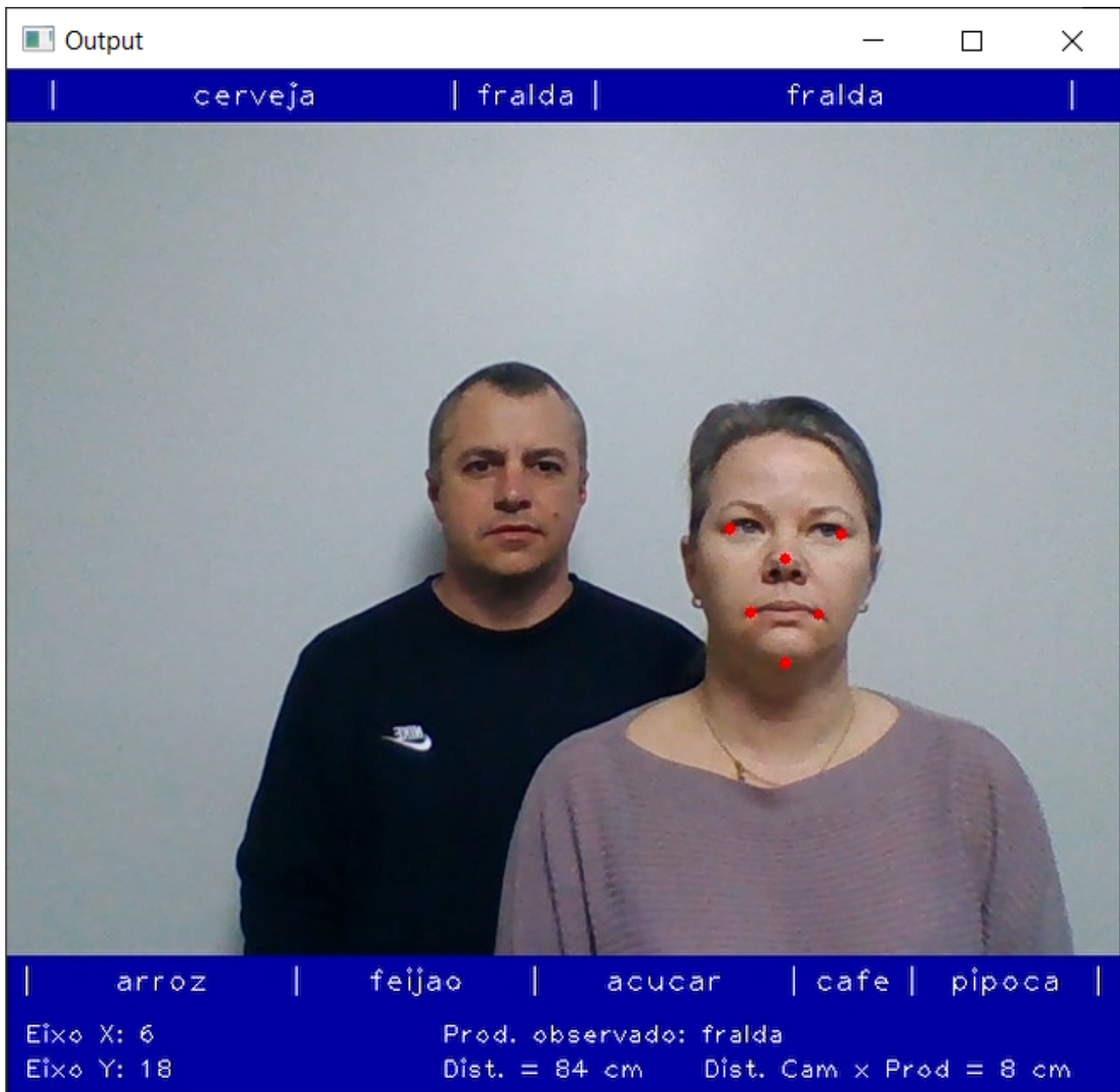


Figura 7. Exemplo de captura da face mais próxima à câmera. Fonte: o autor

A aplicação mostrou corretamente os itens visualizados e associados ao cliente, conforme esperado, seguindo as definições do arquivo de configuração. Adicionalmente, mostra o ângulo de visão da pessoa e o produto que está sendo observado, permitindo ao operador visualizar estas informações em tempo real no seu computador (Figura 8).

O tempo de resposta da aplicação foi satisfatório. Observou-se que o item que estava sendo observado era exibido antes mesmo do participante falar o mesmo em voz alta. Deve-se levar em conta, também, que o processo de identificação da face e do produto, gravação do arquivo em disco e a comparação com as faces previamente salvas, levou, em média, 1,7s para ser executado.

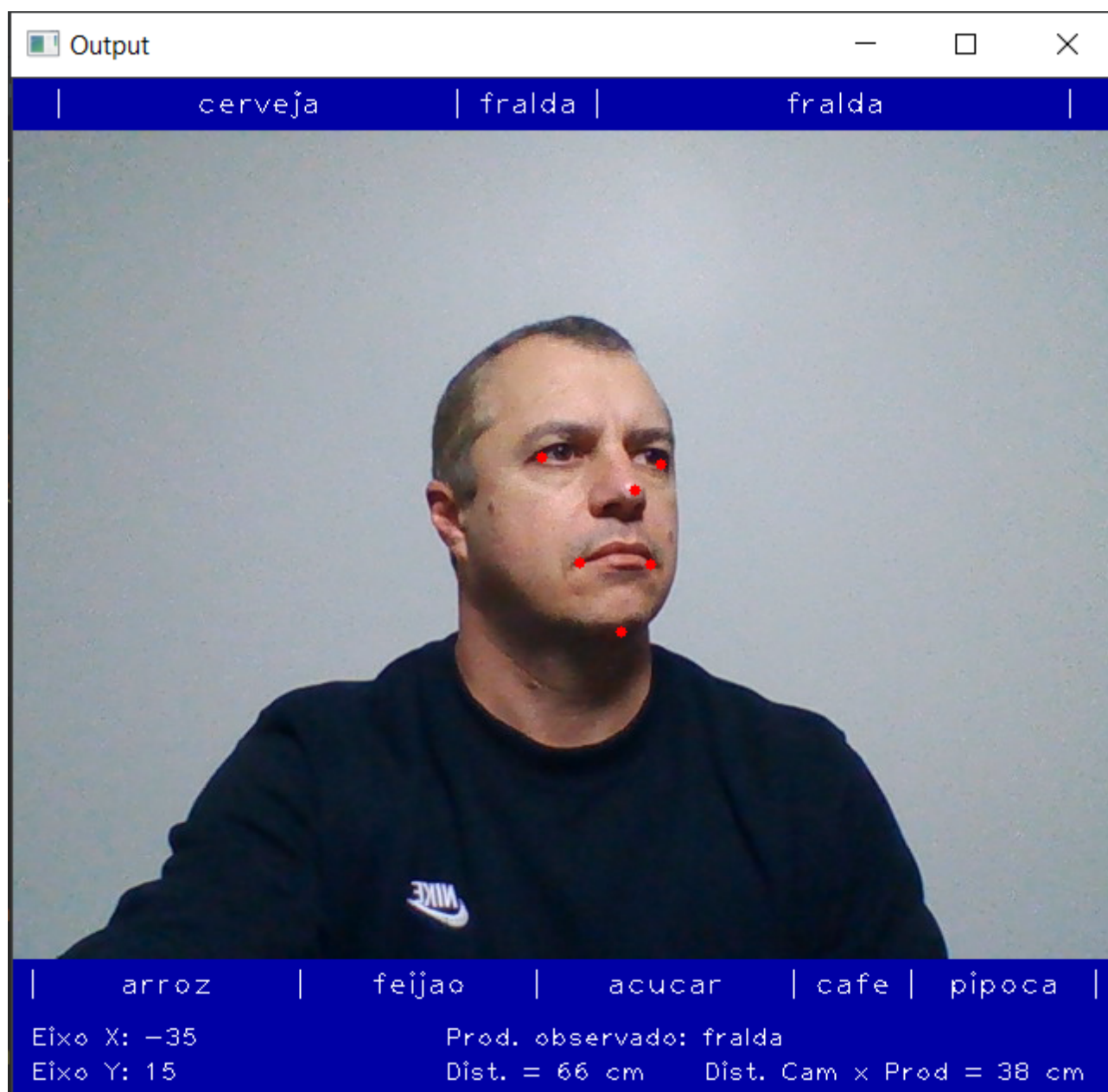


Figura 8. Aplicação mostrando as prateleiras configuradas e as informações sobre a face. Fonte: o autor

A identificação da face se mostrou precisa quando a distância entre a pessoa e a câmera estava entre 0,20 m e 1,50 m. Em distâncias maiores, o sistema não é capaz de identificar os pontos do rosto, que são utilizados para a definição do ângulo e, por consequência, o produto que está sendo observado. Considerando que o comprimento do braço de um adulto é menor que 1 m, a pessoa normalmente estará próxima da prateleira no momento da compra, analisando os produtos a adquirir. Desta forma, a distância que a aplicação identifica as faces é adequada para o cumprir seu objetivo.

Como pode ser observado na Tabela 4, houveram problemas com pessoas maiores de 1,81 m. Isso se deve à configuração do limite do ângulo vertical de cada prateleira. Suspeita-se que o mesmo problema poderia ocorrer com pessoas de estatura baixa (inferior

a 1,58 m) caso o ângulo vertical fosse configurado para identificar os produtos observados pelas pessoas com mais de 1,81 m. Para resolver esses problemas, pode-se avaliar o uso de um *offset*, de modo a fazer que a altura dos olhos das pessoas fique numa posição paralela em relação à câmera, tanto vertical quanto horizontalmente.

Dois produtos tiveram percentual de acerto menor no reconhecimento: *café* e *pipoca*. O *café* ocupa um espaço pequeno na prateleira e, por este motivo, a aplicação não conseguiu identificar claramente quando o mesmo estava sendo visualizado. Para solucionar o problema de itens com esta configuração faz-se necessário aprimorar o cálculo da distância lateral em relação à câmera. A não detecção da *pipoca* está associada a uma restrição da biblioteca Face Recognition, herdada da biblioteca DLib. Elas compartilham uma limitação no ângulo de visão, que varia de +30 a -30°. Desta forma, quando a pessoa está posicionada com o rosto além do limite de 30°, a face não é detectada e, com isso, o interesse no produto não é registrado. A utilização de mais de uma câmera, para monitorar a mesma prateleira, poderia sanar este problema.

Durante o estudo, observou-se um caso em que um dos participantes (ID #04) não movimentou muito a cabeça, mexendo apenas os olhos. Nesta situação, os produtos que estavam posicionados mais à esquerda (arroz) e mais à direita (café e pipoca) não foram identificados. Nesta situação, faz-se necessária a detecção da posição da íris para a correta captura do produto de interesse.

Com outro participante (ID #12), a face não foi identificada num primeiro momento, em virtude da barba. Após ele começar a falar os produtos que estava analisando, a aplicação passou a detectar o rosto, uma vez que os pontos relativos à boca são relevantes para a identificação da face.

Referente aos locais de instalação da câmera, se existir iluminação solar direta no rosto da pessoa, esta não interfere na identificação da face e no produto que está sendo observado. Porém, em locais escuros, a aplicação pode não detectar a face que está na imagem. Desta forma, o lojista deverá escolher um local com boa iluminação natural ou artificial para instalar a câmera. Outra maneira de resolver essa situação é adotar o uso de câmera com infravermelho.

Em síntese, pode-se destacar como uma das contribuições desse estudo o fato da abordagem proposta identificar corretamente os produtos de interesse de um cliente durante as suas compras, preservando sua identidade, sem identificação nominal, considerando a LGPD. Além disso, demonstrou-se que a aplicação é capaz de fornecer ao lojista o registro do comportamento e do hábito de compra de seus clientes. Esse fato pode ajudar o lojista na oferta de produtos similares com preços mais acessíveis, oferecer vantagens ao cliente para agregar valor de venda, melhorar a disposição de produtos nos pontos de venda, além de ajudar no treinamento de seus colaboradores.

Dentre as limitações da aplicação, podem ser citadas a necessidade da pessoa estar posicionada em frente à câmera e com o nariz apontando para o produto que está

visualizando, a dificuldade de identificar produtos que ocupam um espaço pequeno na prateleira e a obtenção da face em ambientes com luminosidade reduzida. Outro fator limitante foi a impossibilidade de realizar um teste em uma loja física, devido às restrições impostas pela pandemia da Covid-19.

5. CONCLUSÃO

Neste trabalho foi apresentada uma aplicação chamada **Recommender Vision**, um sistema de recomendação que utiliza visão computacional para fornecer sugestões de venda de produtos para os lojistas. Seu objetivo foi identificar os produtos de interesse de clientes em compra, buscando fornecer mais informações ao lojista a respeito dos hábitos dos seus consumidores e, desta forma, permitir que produtos similares e mais em conta possam ser oferecidos, com o intuito de aumentar o valor das vendas e, conseqüentemente, obter mais lucro.

Este dados podem se tornar relevantes ao lojista, uma vez que, conhecendo o comportamento dos clientes, pode, também, reorganizar os produtos nas prateleiras, fazendo com que os itens mais vendidos ou com maior margem fiquem em uma posição de destaque.

Com base no que foi descrito, pode-se observar que a aplicação identificou as pessoas corretamente, sem gerar falsos positivos, ao mesmo tempo que preservou a identidade das pessoas, já que não houve nomeação das mesmas em nenhum momento.

Em relação à associação correta dos produtos, sua eficiência média foi de 74,3% - o que permite ao lojista oferecer corretamente sete em cada 10 produtos de interesse dos clientes. Adicionalmente, permite que sejam visualizados, em tempo real, os produtos que os clientes estão observando, possibilitando identificar alguns hábitos que os mesmos possam ter, além de detectar possíveis situações onde os produtos podem ser realocados para facilitar a visualização dos consumidores.

O resultado da avaliação mostra que a tecnologia de reconhecimento facial pode melhorar os resultados de recomendação de forma eficaz e, portanto, ajudar os varejistas a aprimorar seu marketing e a abordagem aos clientes.

Como trabalhos futuros, existem funcionalidades que ainda precisam ser implementadas e melhoradas. É necessário efetuar o cálculo do *offset* da face, para posicioná-la no centro da câmera, tanto horizontal quanto verticalmente, e testar com uma câmera com abertura focal maior. Adicionalmente, levar em consideração a íris da pessoa, a qual pode estar observando um item diferente do que aquele que o nariz está "apontando". Também deve ser pesquisada uma forma de identificar, com maior precisão, itens que ocupam um espaço reduzido na prateleira, melhorando a assertividade da aplicação.

Adicionalmente, a interface deve ter a possibilidade de excluir as faces previamente identificadas, caso o cliente faça esta solicitação. Conseqüentemente, os produtos observados também serão eliminados, extinguindo, desta forma, o histórico de determinado cliente na aplicação.

Apesar do software ter se comportado satisfatoriamente no ambiente de laboratório, é importante que seja feito um teste em ambiente real para a identificação de outros pontos de melhoria na aplicação. Assim, pode-se considerar características do âmbito comercial diário para evoluir a solução, como o tratamento de faces em aglomerações de pessoas.

Ao se pensar em uma loja física, deve-se levar em conta as faces dos funcionários, as quais devem ser ignoradas pela aplicação, para que a mesma não os considere como clientes interessados nos produtos. De maneira análoga e de maior complexidade, deve-se avaliar também o funcionamento da aplicação em casos de irmãos muito parecidos ou gêmeos univitelinos.

Outra evolução necessária à aplicação é a identificação de mais de uma pessoa ao mesmo tempo, controlando o produto observado e o tempo de observação independentemente. Nesse âmbito, também é relevante comparar o desempenho da solução proposta usando algum modelo pré-treinado, como o FaceNet, sistema de reconhecimento facial que utiliza rede neural convolucional profunda. É importante também considerar testes com um grupo maior de pessoas, de forma a validar o fator de similaridade utilizado na comparação entre as faces codificadas.

Após o aperfeiçoamento do software, outra sugestão de trabalho futuro é o seu uso em uma solução embarcada. Isto pode facilitar sua replicação utilizando microcontroladores de baixo custo e baixo consumo de energia, como ESP32.

Por fim, pretende-se efetuar o devido registro de software do **Recommender Vision**, visando proteger os direitos autorais sobre a aplicação que foi produzida neste trabalho.

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] VAIDYA, N.; KHACHANE, A. Recommender systems-the need of the ecommerce era. In: IEEE. *2017 International Conference on Computing Methodologies and Communication (ICCMC)*. [S.l.], 2017. p. 100–104.
- [2] RICCI, F.; ROKACH, L.; SHAPIRA, B. Introduction to recommender systems handbook. In: *Recommender systems handbook*. [S.l.]: Springer, 2011. p. 1–35.
- [3] ALFIAN, G. et al. Customer behavior analysis using real-time data processing. *Asia Pacific Journal of Marketing and Logistics*, p. 265–290, 2019.
- [4] YU, L. et al. A content-based goods image recommendation system. *Multimedia Tools and Applications*, p. 4155–4169, 2018.
- [5] BATMAZ, Z. et al. A review on deep learning for recommender systems: Challenges and remedies. *Artificial Intelligence Review*, p. 1–37, 6 2018.
- [6] HERNANDEZ, D. A. M.; NALBACH, O.; WERTH, D. How computer vision provides physical retail with a better view on customers. In: IEEE. *2019 IEEE 21st Conference on Business Informatics (CBI)*. [S.l.], 2019. v. 1, p. 462–471.
- [7] ADOMAVICIUS, G.; TUZHILIN, A. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE Transactions On Knowledge and Data Engineering*, p. 734–749, 2005.
- [8] MORA, D. et al. An in-store recommender system leveraging the microsoft hololens. In: SPRINGER. *International Conference on Human-Computer Interaction*. [S.l.], 2020. p. 99–107.
- [9] NEVES, L. A. P.; Vieira Neto, H.; GONZAGA, A. (Ed.). *Avanços em Visão Computacional*. 1. ed. Curitiba, PR: Omnipax, 2012. 406 p. ISBN 978-85-64619-09-8.
- [10] PORTUGAL, I.; ALENCAR, P.; COWAN, D. The use of machine learning algorithms in recommender systems: A systematic review. *Expert Systems with Applications*, p. 205–227, 2018.
- [11] GHEISARI, M.; WANG, G.; BHUIYAN, M. Z. A. A survey on deep learning in big data. In: IEEE. *2017 IEEE International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC)*. [S.l.], 2017. v. 2, p. 173–180.
- [12] SINGH, S.; PRASAD, S. Techniques and challenges of face recognition: A critical review. *Procedia computer science*, Elsevier, v. 143, p. 536–543, 2018.

- [13] TARNOWSKA, K.; RAS, Z. W.; DANIEL, L. *Recommender System for Improving Customer Loyalty*. [S.l.]: Springer, 2020.
- [14] BRASIL. Lei nº 13.709, de 14 de agosto de 2018. Dispõe sobre a proteção de dados pessoais e altera a Lei nº 12.965, de 23 de abril de 2014 (Marco Civil da Internet). *Diário Oficial da República Federativa do Brasil*, Brasília, DF, 14 ago. 2018. Disponível em: <http://www.in.gov.br/materia/-/asset_publisher/Kujrw0TZC2Mb/content/id/36849373/do1-2018-08-15-lei-no-13-709-de-14-de-agosto-de-2018-36849337>. Acesso em: 03 Out. 2019.
- [15] WEITZEL, L.; OLIVEIRA, J. P. M. de. Sistemas de recomendação de informação em saúde baseado no perfil do usuário. *Journal of Health Informatics - Ano XX - Volume XX*, 2010.
- [16] TSIHRINTZIS, G. A.; VIRVOU, M.; JAIN, L. C. *Multimedia Services in Intelligent Environments: Recommendation Services*. Switzerland: Springer International Publishing, 2013. 185 p. ISBN 978-3-319-00375-7.
- [17] CAZELLA, S. C. et al. Recomendação de objetos de aprendizagem empregando filtragem colaborativa e competências. *XX Simpósio Brasileiro de Informática na Educação*, 2009.
- [18] LESKOVEC, J.; RAJARAMAN, A.; ULLMAN, J. D. *Mining of Massive Datasets*. 2. ed. New York, USA: Cambridge University Press, 2014. 459 p. ISBN 978-1-107-07723-2.
- [19] FERRETTO, L. R. *Um sistema de recomendação de atividades físicas para hipertensos no contexto de e-Health*. Dissertação (Mestrado) — Dissertação (Mestrado em Computação Aplicada) - Universidade de Passo Fundo, Passo Fundo, RS, Brasil, 2018. Disponível em: <<http://tede.upf.br/jspui/handle/tede/1465>>. Acesso em: 20 Set. 2019.
- [20] BARBIERI, L. E. *Um sistema de recomendação aplicado a produtos com baixo giro de estoque*. Dissertação (Mestrado) — Dissertação (Mestrado em Computação Aplicada) - Universidade de Passo Fundo, Passo Fundo, RS, Brasil, 2019.
- [21] VIVIAN, G. R. *Recomendação de carreira de pesquisadores: uma abordagem baseada em personalização, similaridade de perfil e reputação*. Dissertação (Mestrado) — Dissertação (Mestrado em Computação Aplicada) - Universidade de Passo Fundo, Passo Fundo, RS, Brasil, 2017. Disponível em: <<http://tede.upf.br/jspui/handle/tede/1425>>.
- [22] GONZALEZ, R. C.; WOODS, R. E. *Digital Image Processing*. 3. ed. Upper Saddle River, New Jersey, USA: Pearson Education, Inc., 2008. 977 p. ISBN 978-0-13-168728-8.
- [23] SAS. *History Of Computer Vision*. 2017. Disponível em: <https://www.sas.com/de_ch/insights/analytics/computer-vision.html>. Acesso em: 20 Set. 2019.

- [24] SURMA, G. *Image Classifier - Catsvs Dogs*. 2018. Disponível em: <<https://towardsdatascience.com/image-classifier-cats-vs-dogs-with-convolutional-neural-networks-cnns-and-google-colabs-4e9af21ae7a8>>. Acesso em: 20 Set. 2019.
- [25] POKHREL, S. *How Does Computer Understand Images?* 2018. Disponível em: <<https://towardsdatascience.com/how-does-computer-understand-images-c1566d4537bf>>. Acesso em: 20 Set. 2019.
- [26] SZELISKI, R. *Computer Vision: Algorithms and Applications*. Ithaca, New York, USA: Springer, 2011. 833 p. ISBN 978-1-84882-934-3.
- [27] MENA, I. *Verbete Draft: o que é Reconhecimento Facial*. 2018. Disponível em: <<https://projetodraft.com/verbete-draft-o-que-e-reconhecimento-facial/>>. Acesso em: 20 Set. 2019.
- [28] GRANATYR, J.; SILVA, L. da. *Curso: Reconhecimento de faces e de Objetos com Python e DLib*. 2018. Disponível em: <<https://www.udemy.com/course/reconhecimento-de-faces-e-de-objetos-com-python-e-dlib>>. Acesso em: 14 Set. 2019.
- [29] DALAL, N.; TRIGGS, B. Histograms of oriented gradients for human detection. In: *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Volume 1 - Volume 01*. Washington, DC, USA: IEEE Computer Society, 2005. (CVPR '05), p. 886–893. ISBN 0-7695-2372-2. Disponível em: <<http://dx.doi.org/10.1109/CVPR.2005.177>>.
- [30] CRUZ, J. E. C. *Reconhecimento de objetos em imagens orbitais com o uso de abordagens do tipo descritor-classificador*. Dissertação (Mestrado) — Instituto Nacional de Pesquisas Espaciais - INPE, São José dos Campos, 2014. Disponível em: <<https://repositorio.ufsc.br/bitstream/handle/123456789/160734/337662.pdf>>. Acesso em: 20 Set. 2019.
- [31] INTEL. *Histogram of Oriented Gradients (HOG) Descriptor*. 2018. Disponível em: <<https://software.intel.com/en-us/ipp-dev-reference-histogram-of-oriented-gradients-hog-descriptor>>. Acesso em: 20 Set. 2019.
- [32] JOSÉ, I. *Facial mapping (landmarks) with Dlib + python*. 2018. Disponível em: <<https://towardsdatascience.com/facial-mapping-landmarks-with-dlib-python-160abcf7d672>>. Acesso em: 20 Set. 2019.
- [33] ARAPAKIS, I. et al. Integrating facial expressions into user profiling for the improvement of a multimodal recommender system. In: IEEE. *2009 IEEE International Conference on Multimedia and Expo*. [S.l.], 2009. p. 1440–1443.

- [34] SEBE, N.; COHEN, I.; HUANG, T. S. Multimodal emotion recognition. In: _____. *Handbook of Pattern Recognition and Computer Vision*. [s.n.], 2005. p. 387–409. Disponível em: <https://www.worldscientific.com/doi/abs/10.1142/9789812775320_0021>.
- [35] LUCEY, P. et al. The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression. In: *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops*. [S.l.: s.n.], 2010. p. 94–101.
- [36] WU, C.-C.; ZENG, Y.-C.; SHIH, M.-J. Enhancing retailer marketing with an facial recognition integrated recommender system. *IEEE International Conference on Consumer Electronics-Taiwan*, p. 25–26, 2015.
- [37] SIGHTCORP. *DeepSight Toolkit*. 2013. Disponível em: <<https://sightcorp.com/deepsight-toolkit/>>. Acesso em: 12 Dez. 2020.
- [38] CABRAL, E. L. L. *Preliminares Matemáticos: transformação de coordenadas*. 2021. Disponível em: <[http://sites.poli.usp.br/p/eduardo.cabral/Preliminares Matemáticos.pdf](http://sites.poli.usp.br/p/eduardo.cabral/Preliminares%20Matemáticos.pdf)>. Acesso em: 04 Jun. 2021.
- [39] ROSEBROCK, A. *Find distance from camera to object/marker using Python and OpenCV*. 2015. Disponível em: <<https://www.pyimagesearch.com/2015/01/19/find-distance-camera-objectmarker-using-python-opencv/>>. Acesso em: 04 Jun. 2021.
- [40] GIDLÖF, K. et al. Using eye tracking to trace a cognitive process: Gaze behaviour during decision making in a natural environment. *Journal of Eye Movement Research*, v. 6, n. 1, Apr. 2013. Disponível em: <<https://bop.unibe.ch/JEMR/article/view/2351>>.



UPF

UNIVERSIDADE
DE PASSO FUNDO

UPF Campus I - BR 285, São José
Passo Fundo - RS - CEP: 99052-900
(54) 3316 7000 - www.upf.br