



**UNIVERSIDADE ESTADUAL DA PARAÍBA  
CAMPUS I – CAMPINA GRANDE  
CENTRO DE CIÊNCIAS E TECNOLOGIA  
CURSO DE CIÊNCIA DA COMPUTAÇÃO**

**TALLES HENRIQUE SOUSA GOMES**

**UTILIZAÇÃO DE SINALIZAÇÃO DIGITAL PARA RECOMENDAÇÃO DE  
PRODUTOS BASEADOS NAS CARACTERÍSTICAS EXTRAÍDAS DE PESSOAS  
PRESENTES NO AMBIENTE**

**CAMPINA GRANDE  
2018**

**TALLES HENRIQUE SOUSA GOMES**

**SISTEMA DE SINALIZAÇÃO DIGITAL PARA EXIBIÇÃO DE ANÚNCIOS  
BASEADOS NAS CARACTERÍSTICAS EXTRAÍDAS DE PESSOAS PRESENTES  
NO AMBIENTE**

Trabalho de Conclusão de Curso da  
Universidade Estadual da Paraíba, como  
requisito parcial à obtenção do título de  
Graduado em Ciência da Computação.

**Área de concentração:** Visão  
computacional.

**Orientador:** Prof. Dr. Frederico Moreira  
Bublitz.

**CAMPINA GRANDE  
2018**

É expressamente proibido a comercialização deste documento, tanto na forma impressa como eletrônica. Sua reprodução total ou parcial é permitida exclusivamente para fins acadêmicos e científicos, desde que na reprodução figure a identificação do autor, título, instituição e ano do trabalho.

G633s Gomes, Talles Henrique Sousa.  
Sistema de sinalização digital para exibição de anúncios baseados nas características extraídas de pessoas presentes no ambiente [manuscrito] : / Talles Henrique Sousa Gomes. - 2018.  
49 p. : il. colorido.  
Digitado.  
Trabalho de Conclusão de Curso (Graduação em Computação) - Universidade Estadual da Paraíba, Centro de Ciências e Tecnologia, 2018.  
"Orientação : Prof. Dr. Frederico Moreira Bublitz, Departamento de Computação - CCT."  
1. Visão computacional. 2. Varejo inteligente. 3. Sinalização digital. 4. Sistemas de recomendação.  
21. ed. CDD 005.3

TALLES HENRIQUE SOUSA GOMES

**SISTEMA DE SINALIZAÇÃO DIGITAL PARA EXIBIÇÃO DE  
ANÚNCIOS BASEADOS NAS CARACTERÍSTICAS  
EXTRAÍDAS DE PESSOAS PRESENTES NO AMBIENTE**


Trabalho de Conclusão de Curso de Graduação  
em Ciência da Computação da Universidade  
Estadual da Paraíba, como requisito à  
obtenção do título de Bacharel em Ciência da  
Computação.

Aprovada em 21 de Junho de 2018.



---

Prof. Dr. Frederico Márcia Bublitz (DC – UEPB)  
Orientador



---

Prof. Me. Adrian Lívio Vasconcelos Guedes (DC – UEPB)  
Examinador



---

Prof. Dr. Wellington Candeia de Araújo (DC – UEPB)  
Examinador



## **AGRADECIMENTOS**

A Deus, pelo seu grande amor e por sempre guiar minha vida por caminhos que, aos meus olhos, parecia impossível.

A toda a minha família, em especial aos meus pais pelo amor, confiança e pelo suporte constante nesta longa jornada.

A minha esposa Daniella, pelo amor, auxílio e compreensão nos momentos difíceis, como quando ainda namorávamos que tinha que passar meses sem visita-la na cidade onde residia para me dedicar aos estudos.

Ao meu avô Adauto Maurício do Nascimento (*in memorian*), que sempre me incentivava a estudar e se alegrava com minhas conquistas.

A todos os meus amigos do curso de computação (UEPB), não vou citar os nomes para não correr o risco de esquecer-me de nenhum deles, mas todos não hesitavam em ajudar quando era necessário.

A todos os professores do curso de Computação (UEPB) por transmitirem o conhecimento necessário para que eu alcançasse a graduação.

Ao professor Dr. Frederico Moreira Bublitz, pelo apoio concedido, pela confiança e pelas oportunidades a mim concedidas.

## RESUMO

Sistemas que selecionam e apresentam anúncios ou outro tipo de informação que coincide com o interesse do usuário presente, mediante análise prévia de informações obtidas, são chamados Sistemas de Recomendação. Esses softwares estão presentes em sites de compras e em sites que cedem espaços para divulgação de uma marca ou produto. Porém existe necessidade de empresas físicas adaptarem-se às preferências dos consumidores buscando um diferencial competitivo num mercado dinâmico, complexo e concorrido. Empresas tendem a utilizar tecnologias que ofereçam uma experiência diferenciada para os clientes, se enquadrando no conceito de Varejo Inteligente. Nesse sentido, o Marketing, que consiste em uma série de ações sistemáticas visando a adequação das capacidades de uma empresa aos interesses dos consumidores, tem sido amplamente usado. Existe atualmente uma grande quantidade de ambientes, tais como universidades, rodoviárias, aeroportos e eventos em que painéis digitais (e.g., televisores LCD e dispositivos de sinalização digital) são usados para trazer informações às pessoas e promover produtos e serviços. Embora estes dispositivos possam atingir uma grande quantidade de consumidores, os anunciantes contam com pouca informação das pessoas que estão presentes no ambiente, inviabilizando que os anúncios exibidos sejam mais relevantes para essas pessoas, ou dificultando o feedback em relação à audiência da propaganda. Diante disso, é proposto neste trabalho o desenvolvimento de um modelo e protótipo de software para viabilizar a circulação de anúncios contextualizados por meio de painéis digitais em ambientes públicos. Com isso, micro e pequenas empresas do poderão ter acesso ao público-alvo de seus produtos, melhorando sua capacidade competitiva por meio de uma comunicação adequada com o mercado. Além disso, lojas podem utilizar o sistema para realizar recomendações de produtos e serviços direcionados para o respectivo público-alvo.

**Palavras-Chave:** Visão computacional. Varejo inteligente, Sinalização digital. Reconhecimento de face, Estimação de idade, Classificação de gênero.

## ABSTRACT

Systems that select and present ads or other information that coincides with the interest of the present user, through prior analysis of information obtained, are called Recommendation Systems. These softwares are present on shopping sites and on sites that give you spaces to promote a brand or product. However, there is a need for physical companies to adapt to the preferences of consumers seeking a competitive advantage in a dynamic, complex and competitive market. Companies tend to use technologies that offer a differentiated experience for the customers, if it falls within the concept of Intelligent Retail. In this sense, Marketing, which consists of a series of systematic actions aimed at adapting the capabilities of a company to the interests of consumers, has been widely used. There are currently a large number of environments, such as universities, roads, airports and events where digital panels (eg, LCD TVs and digital signage devices) are used to bring information to people and promote products and services. While these devices can reach a large number of consumers, advertisers have little information on the people in the room, making the ads that are most relevant to them difficult, or making it difficult to get feedback on the advertising audience. Therefore, it is proposed in this work the development of a model and software prototype to enable the circulation of contextualized ads through digital panels in public environments. With this, micro and small businesses can have access to the target audience of their products, improving their competitive capacity through proper communication with the market. In addition, stores can use the system to make recommendations for products and services targeted to their target audience.

**Keywords:** Computer vision. Digital signage. Smart Retail. Face recognition. Age estimation. Gender classification.

## LISTA DE ILUSTRAÇÕES

|   |    |
|---|----|
| Figura 1: Exemplo de recursos de retângulo .....  | 19 |
| Figura 2: Exemplos de pontos onde são calculados os valores da imagem integral .....  | 19 |
| Figura 3: Recursos selecionados pelo AdaBoost .....   | 20 |
| Figura 4: Etapas para reconhecimento de gênero .....  | 23 |
| Figura 5: LBP básico (Bouchirika et al., 2015) .....  | 24 |
| Figura 6: Percentagem de classificação correta por níveis de faixas etárias .....   | 24 |
| Figura 7: cinco áreas-chaves do varejo inteligente (Grewal, Roggeveen e Nordfält, 2017) ..  | 26 |
| Figura 8: Faixas etárias definidas .....  | 29 |
| Figura 9: Gêneros definidos .....   | 29 |
| Figura 10: Etapas do funcionamento da aplicação .....   | 31 |
| Figura 11: Alinhamento da posição dos olhos feitas pelo algoritmo .....   | 33 |
| Figura 12: Trecho do código em Python responsável pela rotação e alinhamento das<br>imagens com base na posição dos olhos .....                                 | 34 |
| Figura 13: Nomes das imagens de acordo com o padrão definido para diferenciação entre<br>sexo e idade .....   | 35 |
| Figura 14: Trecho da codificação C++ onde é incluída a biblioteca OpenCV no projeto. ....   | 36 |
| Figura 15: Treinando a máquina através das imagens do banco de imagens .....  | 37 |
| Figura 16: Classificadores haar disponíveis na biblioteca OpenCV. ....  | 38 |
| Figura 17: Trecho do código responsável pela detecção de faces no frame capturado pela<br>câmera. ....  | 39 |
| Figura 18: Trecho do código que faz a previsão das características dos rostos detectados. 39  |    |
| Figura 19: Programa detectando o rosto e classificando o sexo e a faixa etária, neste caso<br>uma pessoa do sexo masculino e com idade entre 20 e 36 anos. .... | 40 |
| Figura 20: Estrutura das pastas onde são armazenados os anúncios. ....  | 41 |
| Figura 21: Trecho da codificação onde é apresentado o anúncio para o expectador. ....   | 41 |
| Figura 22: TV de 19" utilizada para exibição do anúncio. Em cima da TV está a câmera<br>utilizada para capturar as imagens dos clientes. ....                   | 42 |
| Figura 23: Equipamento posicionado em cima do balcão de atendimento, local visível para<br>todos os clientes presentes. ....                                    | 43 |
| Figura 24: Cliente visualizando o anúncio. ....   | 44 |

## LISTA DE ABREVIATURAS E SIGLAS

LED - Light Emitting Diode

LCD - Liquid Crystal Display

TV - Televisão

BSD - Berkeley Software Distribution

LBP - Local binary patterns

HOG - Histogram of Oriented Gradients

FV – Fisher Vectors

IDE - Integrated Development Environment

CSV - Comma-Separated Values

XML - Extensible Markup Language

JPG - Joint Photographic Group

PNG - Portable Network Graphics

WMV – Windows Media Video

AVI - Audio Video Interleave

## SUMÁRIO

|  |    |
|--|----|
| 1 – INTRODUÇÃO .....   | 9  |
| 1.1 - MOTIVAÇÃO E CARACTERIZAÇÃO DO PROBLEMA .....   | 10 |
| 1.2 - OBJETIVO GERAL.....  | 11 |
| 1.3 - OBJETIVOS ESPECÍFICOS .....  | 11 |
| 2 – FUNDAMENTAÇÃO TEÓRICA .....  | 13 |
| 2.1 – SINALIZAÇÃO DIGITAL .....  | 13 |
| 2.2 – VISÃO COMPUTACIONAL .....  | 13 |
| 2.3 – VAREJO INTELIGENTE.....  | 15 |
| 2.4 – OpenCV .....   | 16 |
| 3 – TRABALHOS RELACIONADOS.....  | 17 |
| 3.2 – VISÃO COMPUTACIONAL .....  | 17 |
| 3.2.1 - Detecção De Rosto.....   | 17 |
| 3.2.2- Classificação de gênero e idade .....   | 21 |
| 3.3 – SINALIZAÇÃO DIGITAL NO COMÉRCIO VAREJISTA .....  | 25 |
| 4 - ESPECIFICAÇÃO .....  | 27 |
| 6 - PROCESSOS ASSOCIADOS AO DESENVOLVIMENTO .....  | 31 |
| 7 - REQUISITOS DO SISTEMA.....   | 32 |
| 8 - DESENVOLVIMENTO .....  | 32 |
| 8.1 – TRATAMENTO DAS IMAGENS DO BANCO DE IMAGENS .....   | 32 |
| 8.2 – TREINAMENTO DE MÁQUINA .....   | 36 |
| 8.3 – DETECÇÃO DE ROSTOS .....   | 37 |
| 8.4 – CLASSIFICAÇÃO DE GÊNERO E IDADE .....  | 39 |
| 8.5 – SELEÇÃO E EXIBIÇÃO DO ANÚNCIO.....   | 40 |
| 9 – IMPLANTANDO O PROJETO EM UM ESTABELECIMENTO COMÉRCIAL .....  | 41 |
| 10 – CONCLUSÃO .....   | 44 |
| 10.1 – TRABALHOS FUTUROS .....   | 45 |
| BIBLIOGRAFIA.....  | 47 |
| Site da Microsoft. Disponível em: < <a href="https://www.microsoft.com/en-us/research/research-area/computer-vision/">https://www.microsoft.com/en-us/research/research-area/computer-vision/</a> >. Acesso em: 2 Março 2018. .... | 47 |

## 1 – INTRODUÇÃO

A sinalização digital é definida como a exibição de informações usando vários tipos de mídia digital em *displays* que usam a tecnologia de diodo emissor de luz (LED), de cristais líquidos (LCD) ou projetores (LAWYER, 2015). É comum a presença da sinalização digital em espaços públicos como rodoviárias, aeroportos, no comércio varejista, e em diversos lugares públicos espalhados pela cidade. Porém o que vemos hoje são apenas *displays* que “empurram” conteúdos para os transeuntes, sem levar em consideração a preferência de quem está observando (DAVIES, CLINCH, *et al.*, 2014).

A geração Z, mencionada por Priporas, Stylos e Fotiadis (2017), está cada vez mais habituada com as tecnologias de dispositivos pessoais, como computadores e smartphones, com acesso à rede mundial de computadores. Esta geração se depara constantemente com sistemas de recomendação de produtos, seja em páginas da *web*, nos smartphones ou em algum outro dispositivo, é possível perceber recomendações que, na sua maioria, se adequa aos interesses do usuário.

Esta numerosa forma de fazer recomendações através de diversos dispositivos é conhecida como propaganda pervasiva. Segundo Bublitz (2013), A propaganda pervasiva se destaca das outras formas de veiculação de anúncios por permitir a entrega de anúncios sensíveis ao contexto dos consumidores.

Em um futuro próximo, as lojas físicas deverão fazer uso de sistemas de recomendações, assim como as lojas virtuais, seja dentro do ambiente da loja ou em um espaço público, para oferecer conteúdos personalizados a partir de características extraídas dos clientes (DAVIES *et al.*, 2014). A utilização dessa tecnologia se enquadra no conceito de varejo inteligente. (PIPORAS, STYLOS e FOTIADIS, 2017) (QUINTANA *et al.*, 2016 ) (WILLEMS *et al.*, 2016) (PANTANO e TIMMERMANS, 2014)

O varejo inteligente é um termo usado para descrever um conjunto de tecnologias inteligentes que são projetadas para dar ao consumidor uma experiência maior, mais rápida, mais segura e mais inteligente ao fazer compras. (QUINTANA *et al.*, 2016 )

Segundo Eleonora Pantano (2014) A ideia emergente de varejo inteligente refletiria uma ideia particular do varejo, onde empresas e consumidores usam a

tecnologia para reinventar e reforçar seu papel na nova economia de serviços, melhorando a qualidade de suas experiências de compras. Para as empresas é cada vez mais importante interagir com os seus clientes, compreendendo seus requisitos e suas necessidades (HSU, 2017).

Utilizando os conceitos definidos até aqui, este trabalho propõe um modelo e um protótipo de um *software* que se utiliza da sinalização digital em conjunto com a propaganda pervasiva para que empresas varejistas possam atingir o seu público alvo através de anúncios recomendados de acordo com o contexto dos clientes, e assim, adquirindo a característica de varejo inteligente.

## **1.1 - MOTIVAÇÃO E CARACTERIZAÇÃO DO PROBLEMA**

A geração Z está fortemente conectada à internet e bastante focada na inovação. Hoje muitas pessoas dispõem de dispositivos que as fazem interagirem com serviços online, especialmente o comércio online. Devido a facilidade e atratividade das compras *online*, se faz necessário que os varejistas se adaptem às novas características dos consumidores para se manterem competitivos (PRIPORAS, STYLOS e FOTIADIS, 2017).

Atualmente são realizados bastantes estudos buscando desenvolver tecnologias que visam tornar o varejo mais inteligente, e, conseqüentemente, mais atrativo para os clientes. Varejistas estão cada vez mais adquirindo e utilizando tecnologias para melhorar as experiências de compra, pois existe a necessidade das empresas buscarem se adequar aos consumidores tornando um diferencial neste mercado tão competitivo. (BUBLITZ 2013)

Em entrevistas realizadas por Priporas, Stylos e Fotiadis (2017), os autores puderam identificar que os clientes reconhecem a necessidade de tecnologias de informação inteligente dentro do comércio varejista, as quais poderiam ajuda-los a localizar ofertas e produtos que precisam, bem como combinar as suas necessidades com a oferta mais apropriada.

Para Davies et al. (2014) este desejo por parte dos consumidores por conteúdo personalizado reflete precisamente a influência causada por sistemas de recomendações que entregam conteúdos personalizados na web, o que corrobora com Zhitomirsky, Geffet e Blau apud Priporas, Stylos e Fotiadis (2017) quando



afirmam que os consumidores se tornaram cada vez mais dependentes das tecnologias.

Com o intuito de contribuir com o varejo inteligente, e melhorar a experiência dos clientes, a presente pesquisa vem propor soluções simples que com um baixo investimento, empresas possam adquirir características de varejo inteligente, levando seus clientes a experiências de compras diferentes. Para alcançar essa experiência agradável, devem ser fornecidas informações valiosas para auxiliar os clientes em suas tomadas de decisões na hora de efetuar compras, apresentando informações condizentes com o seu perfil, possibilitando a oferta de conteúdos personalizados (QUINTANA et al., 2016). Tal sistema beneficia tanto os consumidores quanto os comércios que utilizam a tecnologia (Bublitz, 2013).

Além disso, a tecnologia proposta também pode ser utilizada como uma forma de promover as micro e pequenas empresas através de propagandas direcionadas ao seu público alvo em painéis instalados em espaços públicos.

## **1.2 - OBJETIVO GERAL**

O objetivo geral deste trabalho é fazer com que empresas, em especial o varejo, tenham acesso a tecnologias de sinalização digital, onde, a partir daí, são adquiridas características que são comuns no comércio *on line*, como a realização de recomendação de ofertas de acordo com a análise do cliente, visando aumentar os lucros e tornando essas empresas mais competitivas e mais interativas ao atingir o seu público definido.

## **1.3 - OBJETIVOS ESPECÍFICOS**

O intuito dessa pesquisa é propor um sistema que utiliza a sinalização digital de forma inteligente para que empresas de pequeno porte possam ter acesso à tecnologias que melhorem a experiência de compra dos usuários e também busque alcançar maiores lucros com a possibilidade de promover produtos e serviços tendo como requisito para implementação do sistema um computador com um *display* (uma TV de LED, por exemplo) e uma câmera conectados.

O *software* faz recomendações, através do *display*, com base nas características de pessoas presentes no ambiente através da extração de características das pessoas presentes através das câmeras de vídeo. Com as imagens obtidas através das câmeras, um algoritmo irá analisá-las, buscando encontrar rostos de pessoas, e, feito isso, realizar uma estimativa de idade e uma classificação de gênero de cada rosto detectado.

Com as informações estimadas pelo algoritmo, painéis localizados dentro das lojas ou em espaços públicos, exibirão ofertas, propagandas ou outro tipo de anúncio que poderá ser direcionado para o seu público alvo, fazendo com que o que está sendo apresentado no painel seja relevante para a maioria dos presentes no ambiente.

No capítulo 2 dessa monografia, busca-se fazer uma abordagem da utilização de sinalização digital e visão computacional definindo conceitos e utilidades. Também explica a evolução do varejo inteligente, suas principais características e a utilização do *marketing* dentro do varejo. A partir daí, será apresentada formas de introduzir tecnologias que busquem alavancar as vendas com o uso de tecnologias que criam uma nova experiência de compras para os clientes tornando as empresas mais competitivas e mais atrativas do ponto de vista do consumidor.

No capítulo 3 são apresentados os algoritmos utilizados para aprendizagem de máquina, detecção de rostos, estimação de idade e classificação de gênero em trabalhos relacionados. Unidos esses algoritmos se tornam uma poderosa ferramenta para extração e análise de características do público presente diante de uma câmera, e que serão utilizados para construção do *software* proposto especificado no Capítulo 4.

O capítulo seguinte descreve os processos utilizados para ao desenvolvimento deste, seguido pela definição dos requisitos apresentado no capítulo 7. Em seguida, o capítulo 8 vai mostrar o processo de desenvolvimento e codificação do *software*. O projeto foi codificado utilizando a linguagem C++ juntamente com a utilização da biblioteca OpenCV. Por fim, o nono capítulo apresenta testes simulados em um ambiente real, um comércio farmacêutico.

Alguns resultados e conclusões obtidas através deste trabalho serão apresentados nos capítulos finais.

## 2 – FUNDAMENTAÇÃO TEÓRICA

### 2.1 – SINALIZAÇÃO DIGITAL

A sinalização busca passar uma mensagem, para a pessoa que a vê, através de imagens. Placas de trânsito, por exemplo, são compostas por desenhos ou símbolos que tem o intuito de transmitir informações relevantes para os condutores de veículos. Outra forma bastante comum de sinalização é a fixação de placas de anúncios ou *outdoors* em lugares públicos, que buscam promover uma marca ou uma campanha governamental, ou expor qualquer outro tipo de informação que desperte o interesse do espectador. A utilização da sinalização vem desde os primórdios da humanidade, e segue inserida no cotidiano da sociedade.

Nos dias atuais, com o avanço da tecnologia, podemos visualizar anúncios sendo exibidos em grandes centros comerciais, nas ruas das cidades e em outros espaços públicos, utilizando dispositivos que dispõem de telas permitindo a exposição de informações através de painéis digitais. Daí surgiu o conceito de Sinalização Digital. Para Lawier:

*“[...] a exibição de informações usando vários tipos de mídia digital em displays que usam a tecnologia de diodo emissor de luz (LED) ou de cristais líquidos (LCD), mas também usando projetores e telas. Os componentes do sistema de sinalização digital incluem uma tela, um player de mídia (algum tipo de computador) ou um sistema de distribuição e um sistema de gerenciamento de conteúdo.” (Lawier, 2015)*

Para Hsu (2017) a Sinalização Digital consiste na utilização de telas para que a atenção das pessoas seja atraída pelo som, pela visão e pelas animações divertidas.

Televisores, projetores, celulares e computadores estão entre os dispositivos utilizados para exibição da informação na sinalização digital (Hsu, 2017).

A grande vantagem na utilização desses *displays* é que seu conteúdo pode ser atualizado com facilidade e muito mais frequentemente que as opções de sinalização estática (Lawyer, 2015).

### 2.2 – VISÃO COMPUTACIONAL

A visão computacional busca obter entendimento de alto nível a partir da análise de imagens digitais, sejam elas dinâmicas (vídeo) ou estáticas (fotografia). O intuito da visão computacional é automatizar tarefas que o sistema visual humano pode fazer utilizando algoritmos, como algoritmos de aprendizagem, computação

gráfica e reconhecimento de padrões, para alcançar a compreensão visual autônoma (Borji, 2017).

Larry Roberts é citado como o pai da visão computacional. Em seu trabalho, no ano de 1960 no MIT, realizou um estudo sobre as possibilidades de extrair informações geométricas 3D a partir de imagens em perspectiva 2D.(Huang, 2017)

O trabalho de visão computacional veio ser aperfeiçoado com os estudos de David Marr em 1978, que adotou uma abordagem diferente, primeiro aplicava algoritmos de processamento de imagem para segmentar as bordas dos objetos a partir da qual extraía um esboço da cena utilizando o estéreo binocular (Huang, 2017), que é a comparação entre imagens diferentes do mesmo objeto. Em estudos posteriores, Marr passou a utilizar técnicas de alto nível para reconhecimento de imagens como análise estrutural e aprendizagem de máquina.

Huang (2017) destaca que a visão computacional não só utiliza captura e processamento de imagens, mas também faz a utilização de sensores que identificam se um objeto está se afastando ou se aproximando de seu veículo, por exemplo. Este paradigma é chamado de “Visão intencional”, o que implica que os algoritmos podem obter informações através de sensores que o homem obtém pela visão.

Shu (2017), em seu trabalho, apresenta a visão computacional como a área que tem por objetivo fazer o computador “ver”, ou seja, adquirir características humanas relacionadas à visão, como o reconhecimento de objetos, reconhecimento de formas, reconhecimento rostos, entre outras características que a visão humana consegue definir.

O conceito de visão computacional definido por Shu (2017) assemelha-se ao que escreveu Huang (2017), que acrescenta em sua pesquisa que a visão computacional, utilizando o ponto de vista da ciência biológica, visa desenvolver modelos computacionais semelhantes ao sistema visual humano. Huang (2017) e Borji (2017) escrevem que, do ponto de vista da engenharia, a visão computacional buscam aplicar suas teorias e modelos para construir sistemas autônomos que consigam desempenhar funções que são comuns aos seres humanos quando utilizam a visão.

A área de visão computacional é multidisciplinar e são bastante utilizados em projetos ambiciosos como desenvolvimento de robôs e de veículos autônomos, como carros, caminhões, aviões e foguetes, por exemplo. A aprendizagem de

máquina e a inteligência artificial são pilares para o desenvolvimento de *softwares* (<https://www.microsoft.com/en-us/research/research-area/computer-vision/>, acessado em 02 de março de 2018) que se utilizam da visão por computador para realizarem algumas tarefas como reconhecimento facial e de objetos, biometria e reconhecimento de padrões, possibilitando o desenvolvimento de softwares automatizados na área médica, na área de segurança com detecção de movimentos, na indústria, utilizando a visão computacional para inspeção de máquinas, como também em aplicações de veículos aéreos, sistemas de prevenção de desastres, robótica, e podendo até ser utilizada na agricultura (Borji, 2017; Huang, 2017; <https://www.microsoft.com/en-us/research/research-area/computer-vision/>, acessado em 02 de março de 2018; Al-Kaff et al., 2017; Yogameena e Nagananthini, 2017; Zion, 2012). Podemos ver que é bastante vasta a área que pode utilizar a visão computacional para automatizar processos, visto que em diversas atividades manuais feitas pelo homem é necessária a utilização da visão.

O objetivo final da visão por computador, segundo Borji (2017) “é a compreensão da imagem, a capacidade não só de recuperar a estrutura da imagem, mas também de saber o que ela representa”.

Mas, apesar do grande investimento realizado na evolução dos algoritmos da visão por computador, é importante destacar que, em termos de desempenho, ainda encontram-se distantes da capacidade e da complexidade da visão humana.

### **2.3 – VAREJO INTELIGENTE**

É cada vez mais importante determinar como o varejo pode interagir com os clientes e compreender os seus requisitos (Willems et al., 2016). O comércio eletrônico atrai muitos compradores pela facilidade e por possuir recursos que atraem os clientes. No entanto, apesar dos muitos méritos do comércio eletrônico, é improvável que as configurações tradicionais do varejo desapareçam (Zhang et al., 2010 apud Willems et al., 2016) em vez disso, o varejo deve inovar e proporcionar recursos que ofereçam recursos para tornar as compras mais atrativas para os clientes.

O conceito de varejo inteligente foi introduzido por Pantano e Timermans (2014), eles se referiam ao uso de tecnologias inteligentes no varejo que proporcionassem uma experiência melhor, maior, mais rápida, mais segura e mais inteligente aos clientes na realização de compras, (Willems et al., 2016) (Quintana et

al., 2016) proporcionando também um retorno para empresa visando aumentar lucros para suas lojas.

Para alcançar essa experiência citada, devem-se fornecer informações valiosas para ajudar os clientes a cumprir seus objetivos através da introdução de novas formas de entender o movimento dos clientes, as intenções e seu perfil de compras e oferecer mais serviços personalizados (Willems et al., 2016).

## 2.4 – OpenCV

Existe uma biblioteca bastante utilizada em trabalhos que envolvem visão computacional chamada OpenCV (Open Source Computer Vision Library). O OpenCV possui licença BSD, ou seja, é gratuita para uso acadêmico e comercial. Ela possui integração com as linguagens C++, Python e Java. Os sistemas Linux, Windows, Mac OS, IOS e Android suportam esta ferramenta. O OpenCV foi todo escrito utilizando a linguagem C/C++ otimizado. A biblioteca pode usar o processamento de mais de um núcleo. (<https://opencv.org/>, acessado em 11 de abril de 2018)

A biblioteca OpenCV é bastante popular no campo da visão computacional. Seu desenvolvimento teve início pela Intel em 1999. A biblioteca concentra seu foco no processamento de imagens em tempo real e inclui implementações de algoritmos de visão computacional mais recentes e que não possuem patentes. Em 2008 Willow Garage assumiu o suporte do OpenCV ([https://docs.opencv.org/2.4/modules/contrib/doc/face\\_rec/face\\_rec\\_tutorial.html#face-recognition](https://docs.opencv.org/2.4/modules/contrib/doc/face_rec/face_rec_tutorial.html#face-recognition), Acessado em 11 de Abril de 2018).

Esta ferramenta é uma biblioteca de software de visão computacional e de aprendizagem de máquina em código aberto capaz de fornecer uma infra-estrutura comum para as aplicações que utilizam visão computacional. Na biblioteca há mais de 2500 algoritmos otimizados.

*Esses algoritmos podem ser usados para detectar e reconhecer rostos, identificar objetos, classificar ações humanas em vídeos, rastrear movimentos de câmera, rastrear objetos em movimento, extrair modelos 3D de objetos, produzir nuvens de pontos 3D a partir de câmeras estéreo, unir imagens para produzir alta resolução imagem de uma cena inteira, encontrar imagens semelhantes de um banco de dados de imagens, remover olhos vermelhos de imagens tiradas usando flash, acompanhar movimentos dos olhos, reconhecer paisagens e estabelecer marcadores para sobrepô-lo com realidade aumentada, etc. (<https://opencv.org/about.html>, Acessado em 11 de abril de 2018)*



Além disso, possui uma documentação detalhada, tutoriais e comunidades que dão suporte para os desenvolvedores, tanto iniciantes quanto experientes. A biblioteca é bastante utilizada em empresas, instituições de pesquisa e órgãos governamentais.

Dentre as empresas que utilizam a biblioteca OpenCV estão o Google, Yahoo, Microsoft, Intel, IBM, Sony, Honda e Toyota, além de muitas startups como a Applied Minds, VideoSurf e Zeitera. (<https://opencv.org/about.html>, Acessado em 11 de abril de 2018).

### **3 – TRABALHOS RELACIONADOS**

#### **3.2 – VISÃO COMPUTACIONAL**

##### **3.2.1 - Detecção De Rosto**

Para os seres humanos é extremamente fácil detectar um rosto de qualquer pessoa que esteja em seu campo de visão. Estudos revelam que crianças de um a três dias de idade são capazes de detectar rostos e distinguir entre rostos conhecidos. Isso porque nosso cérebro possui células nervosas especializadas, que respondem a características locais específicas de uma cena, como bordas, ângulo ou movimentos ([https://docs.opencv.org/2.4/modules/contrib/doc/facerec/facerec\\_tutorial.html](https://docs.opencv.org/2.4/modules/contrib/doc/facerec/facerec_tutorial.html) Acessado em 18-03-2018).

Um dos recursos mais utilizados da visão computacional é a detecção de rostos em imagens, estática ou dinâmica, pois a partir de detecção é que se podem extrair informações relevantes para auxiliar no reconhecimento, na classificação de idade, de gênero, estado emocional, entre outras características extraídas a partir de análises (Bouchrika et al., 2015).

Uma quantidade substancial de informações e características que podem ser deduzidas a partir da imagem do rosto humano, incluindo estado emocional, etnia, idade, gênero e outras características que são extraídas pelos algoritmos da visão computacional. (Bouchrika et al., 2015)

Para resolver o problema de detecção de rostos na computação, é sugerido o algoritmo de Viola-Jones (Pavlov et al., 2013). Segundo Zhang et al. (2016), O

algoritmo mais utilizado para detectar rostos é o algoritmo de Viola e Jones devido à sua facilidade de implementação e por possuir uma alta precisão, e também por possibilitar a detecção em tempo real.

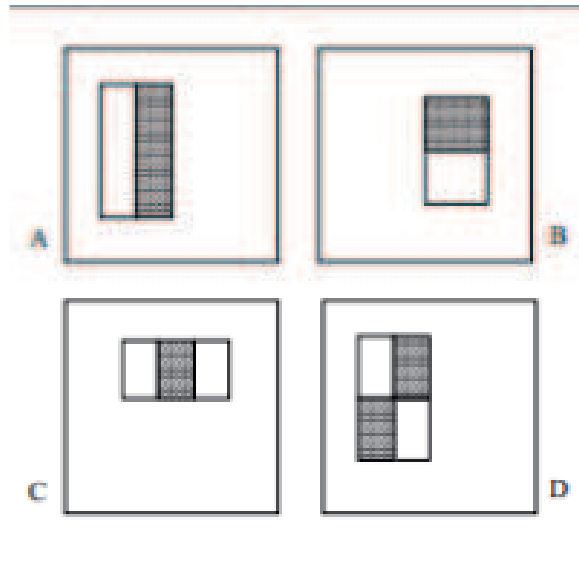
Viola e Jones (2001) propuseram um algoritmo de aprendizagem de máquina para detecção de objetos visuais com capacidade de processamento de imagem de forma muito rápida, o que permite a utilização em sistemas em tempo real.

O algoritmo é baseado em recursos, que funciona muito mais rápido do que sistemas baseados em pixels, alcançando a marca de 0,7 segundos para detectar em uma imagem de 385 por 288 pixels. O desempenho se dá pelo uso do AdaBoost que fornece um algoritmo de aprendizagem e de forte desempenho, e pela utilização de um método para combinar classificadores em uma estrutura de cascata que aumenta drasticamente a velocidade do detector, se concentrando em determinadas regiões da imagem. O maior processamento é dedicado a essas áreas promissoras da imagem (Viola e Jones, 2001).

O algoritmo utiliza três tipos de recursos. O primeiro é o valor de uma característica de dois retângulos que é a diferença entre a soma dos pixels dentro de duas regiões retangulares. O segundo recurso é o valor da característica de três retângulos que calcula a soma dentro dos dois retângulos mais externos subtraídos da soma em um retângulo central. O terceiro e último recurso, é um retângulo de quatro, que calcula a diferença entre os pares de retângulos diagonais (Viola e Jones, 2001).



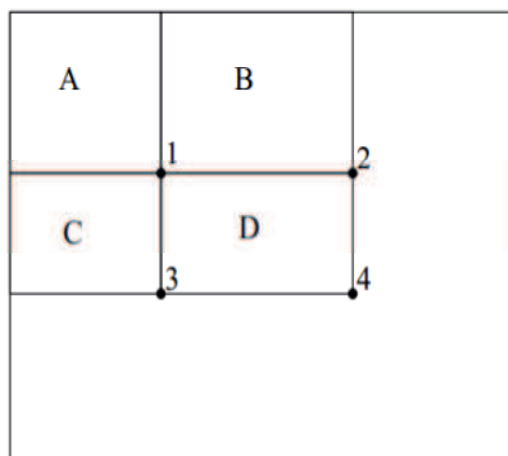
Figura 1: Exemplo de recursos de retângulo



Esses retângulos de recursos “podem ser calculados muito rapidamente usando uma representação intermediária para a imagem que chamamos de imagem integral. A imagem integral no local  $x,y$  contém a soma dos *pixels* acima e à esquerda de  $x,y$ ” (Viola e Jones, 2001). Com isso:

*“[...]Qualquer soma retangular pode ser calculada em quatro referências de matriz. Claramente, a diferença entre duas somas retangulares pode ser calculada em oito referências. Uma vez que os recursos de dois retângulos definidos acima envolvem somas retangulares adjacentes, eles podem ser computados em seis referências de matrizes, oito no caso dos recursos de três retângulos e nove para recursos de quatro retângulos.” (Viola e Jones, 2001)*

Figura 2: Exemplos de pontos onde são calculados os valores da imagem integral



Na Figura 2, temos que o valor da imagem integral no ponto 1 é a soma de pixels do retângulo A. O valor na localização 2 é a soma dos pixels do retângulo A com a soma de *pixels* do retângulo B, na localização 3 é somado os *pixels* de A e C, e no ponto 4 é a soma de todos os retângulos, A, B, C e D, a soma dentro de D pode ser Calculada com essas quatro referências de matriz, ou seja,  $4 + 1 - 2 + 3$ .

Com um conjunto de benefícios e um conjunto de competência de imagens positivas e recusas, qualquer número de abordagens de aprendizado de utensílio poderia ser usado para treinar uma função de classificação. No nosso sistema, uma modificação do AdaBoost é usada tanto para selecionar um conversado conjunto de recursos como para instrumentar o classificador.(Viola e Jones, 2001).

Experimentos realizados por Viola e Jones (2001) demonstraram que um classificador de rosto frontal construído com 200 características produz uma taxa de detecção de 95% com uma taxa de falso positivo de 1 em 14084. Para a tarefa de detecção de rosto, os recursos do retângulo inicial escolhidos pelo AdaBoost são destacados e facilmente interpretados. O primeiro recurso selecionado parece focar a propriedade de que a território dos olhos é constantemente mais escura que a espaço do nariz e das bochechas. Esse recurso é relativamente grande em comparação com a sub-janela de detecção e deve ser reduzido ao tamanho e lugar do rosto. A segunda característica selecionada depende da propriedade de que os olhos são mais escuros do que a ponta do nariz.

**Figura 3: Recursos selecionados pelo AdaBoost**



Viola e Jones (2001) exemplificam o funcionamento da seleção de recursos pelo AdaBoost com a imagem acima. Os dois recursos selecionados pelo AdaBoost

são mostrados na figura acima na linha superior e depois sobrepostos em um rosto de aprendizagem típico na linha de baixo. A primeira característica mede a discordância de intensidade entre o espaço dos olhos e um espaço através das bochechas superiores. O recurso capitaliza a observação de que a área dos olhos é normalmente mais escura do que as bochechas. A segunda característica confronta as intensidades nas zonas dos olhos com a intensidade na ponta do nariz.

Um detector de rosto extremamente veloz como o mencionado acima, permite a utilização em aplicações que requerem uma análise rápida de detecção, como em sistemas que funcionam em tempo real que apenas faz a detecção, como para sistemas que utilizam o reconhecimento de rostos para um posterior processamento e extração de características das imagens.

“A qualidade do passo de detecção de rosto é fundamental para o resultado final de todo o sistema, pois as imprecisões na determinação da posição do rosto podem levar a decisões erradas no estágio de reconhecimento” (Pavlov et al., 2013)..

### **3.2.2- Classificação de gênero e idade**

O sistema de visão humano, além de possuir uma grande habilidade em reconhecer rostos, também possui a capacidade de reconhecer características próprias de cada face. Ao olhar para uma pessoa, o ser humano consegue reconhecer características como sexo ou estimar a idade da pessoa a partir da aparência dela, ou até mesmo perceber em expressões faciais demonstração característica de sentimentos, como medo, tristeza ou alegria (Bouchrika et al., 2015) (Sai, Wang e Teoh, 2014). Com a possibilidade de detectar o rosto das pessoas nas imagens, estudos surgiram voltados para desenvolvimento de algoritmos que realizasse a análise da face com o objetivo de extrair características assim como fazem os seres humanos.

A idade e o gênero das pessoas desempenham papéis importantes nas interações sociais (Levi e Hassner, 2015) e podem ser reconhecidos espontaneamente a partir de traços faciais (Bouchirika et al., 2015). Um rosto humano transmite muitas informações que permite a outros humanos decifrar facilmente, influenciando diretamente na comunicação diária. Isso inclui o

reconhecimento da identidade de uma pessoa, bem como gênero, emoção e a faixa etária (Sai, Wang e Teoh, 2014).

Classificação de gênero está se tornando um passo fundamental em muitas aplicações onde há interação humana (Santacangelo, Farinella e Battiato, 2015). Mas, para um computador, reconhecer o gênero a partir de imagens faciais é uma tarefa desafiadora. A classificação de gênero é um desafio porque a face das pessoas possuem uma ampla forma de variações devido a atributos faciais ou fatores ambientais. Dentre os atributos faciais que causam complicações estão a idade, etnia e a maquiagem. Os fatores Ambientais incluem-se as condições de iluminação, a pose da cabeça, oclusão facial, movimento da câmera, ângulo de vista, desordem de fundo e a qualidade do dispositivo de captura de imagem.

Estimar a idade com precisão é desafiador porque a variação do envelhecimento facial é específica para cada indivíduo e é determinada pela genética da pessoa e muitos fatores externos como cuidados com a pele, exposição ao sol, clima, gênero e estilo de vida (Sai, Wang e Teoh, 2014).

Ainda sobre os fatores que dificultam a classificação correta de idade, Bouchirika et al. cita em seu artigo que:

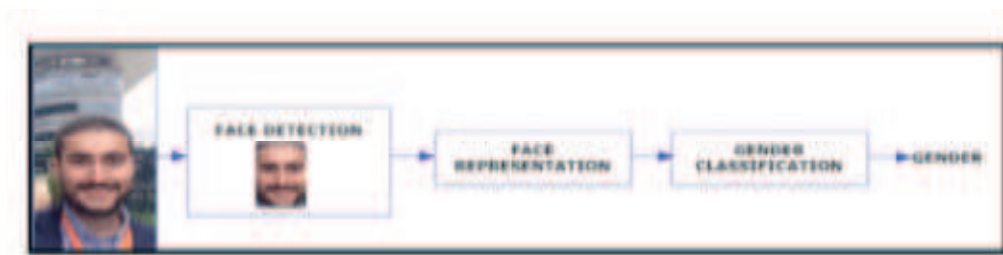
*“O progresso do envelhecimento é incontrolável e irreversível, de modo que diferentes pessoas sofram uma taxa de envelhecimento diferente, dirigida não só pelo gene humano, mas também por muitos outros fatores externos que influenciam o processo de envelhecimento, incluindo condições de saúde, estilo de vida e ambiente de trabalho”. (Bouchrika et al., 2015)*

Ou seja, fatores genéticos ou de rotinas diárias influenciam diretamente na fisionomia da face da pessoa, podendo confundir tanto o sistema de visão humano quanto o sistema de visão por computador.

Grande parte dos métodos de reconhecimento de gênero de alto desempenho envolve aprendizado de máquina e seguem quatro fases: detecção de rosto, pré-processamento de imagem facial, procura de recursos e classificação. (Zhang et al., 2016).

Bouchirika et al. (2015) diz que os sistemas de reconhecimento projetados são compostos por duas fases: Extração de recursos e classificação, porém na extração de recursos, está incluída a fase de detecção de rostos e pré-processamento das imagens citado no trabalho de Zhang et al. (2016).

Figura 4: Etapas para reconhecimento de gênero



Diversos algoritmos de aprendizagem de máquina podem ser aplicados, podemos citar dentre elas algoritmos LBP, HOG, EIPD, Fisher Vectors, entre outros. (Santarcangelo, Farinella e Battiato, 2015) (Zhang et al., 2016). Todos possuem bom desempenho para reconhecimento tanto de faces como objetos.

Zhang et al. (2016) utilizou em seu artigo o algoritmo *Fisher Vectors* para fazer o reconhecimento de gênero. Ele explica que esse algoritmo faz o reconhecimento de características faciais que são codificadas a partir de características de baixo nível em imagens faciais. Um *Vetor Fisher* (FV) é um vetor codificado que aplica *kernels Fisher* em vocabulários visuais onde as palavras visuais são representadas por meio de um modelo gaussiano de mistura. A função *kernel* de *Fisher* é derivada de um modelo de probabilidade generativa e fornece um mecanismo genérico que combina as vantagens das abordagens generativas e discriminativas.

Simonyan et al. (2013) apud Zhang (2016) afirma que o reconhecimento facial utilizando *Fisher Vectors* provaram ser um excelente método de codificação.

Em 1996 o LBP foi introduzido pela primeira vez para análise de textura por Ojala et al. em seu editorial “um estudo comparativo de medidas de textura com classificação baseada em distribuição em destaque”. O LBP pode ser processado de forma eficiente e rápida em uma única imagem que oferece recursos de reconhecimento facial mesmo para imagens de baixa resolução. (Bouchirika et al., 2016)

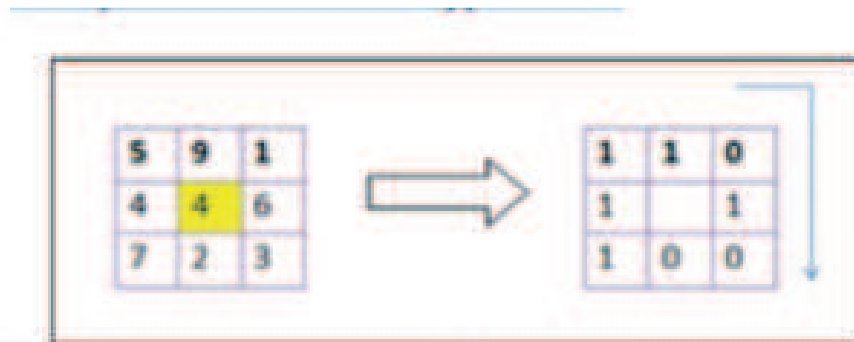
O algoritmo Local Binary Patterns (LBP) apresentou ótimos resultados comparado com os de outros algoritmos para classificação de gênero e estimativa de idade, especialmente quando utilizado com variações (Simonyan et al., 2013 apud Zhang et al., 2016) (Levi e Hassner, 2015).

*Descritor LBP permite realizar a comparação espacial de vizinhança local de um pixel e cria rótulos que são agregados em histogramas. Descritores LBP são robustos para iluminação e às variações de rotação, e são úteis para*

*capturar texturas (barba, por exemplo) (Santacangelo, Farinella e Battiato, 2015).*

O LBP define os pixels de uma imagem qualquer ao limitar cada número dos pixels vizinhos contra o pixel central dentro de uma matriz 3x3. A principal desvantagem do operador binário local básico é a sua pequena área de vizinhança de (3x3) pelo qual pode ignorar ou ignorar características proeminentes para estruturas maiores. A limitação 3x3 resulta em uma série de 1 ou 0 consecutivo. A imagem abaixo apresenta um exemplo. Ao ler na mesma direção da seta azul, é formulada um número binário que é convertido em um número decimal, isto é, um rótulo onde o número 11010011 é convertido para seu representante decimal que é 211. O histograma de 256 binários de rótulos resultantes são calculados e empregados como descritor de textura para aplicações baseadas na face. (Bouchirika et al., 2015).

Figura 5: LBP básico (Bouchirika et al., 2015)



Em testes realizados por Bouchirika et al. (2015) chegaram a alcançar uma taxa de classificação correta de 96,1% ao classificar com sucesso entre dois grupos etários, de 18 à 43 anos e de 44 à 100 anos. Esse primeiro grupo com duas faixas etárias eles classificaram como *level 1*. O *level 2* possui 4 faixas etárias e o *level 3* possui 8 faixas etárias. A imagem abaixo apresenta as faixas etárias por *level* e a taxa de classificação correta alcançada.

Figura 6: Percentagem de classificação correta por níveis de faixas etárias

| Level 1 |       | Level 2 |       | Level 3 |       |
|---------|-------|---------|-------|---------|-------|
| 18-43   | 96.1% | 18-32   | 93.3% | 18-24   | 91.5% |
|         |       | 33-43   |       | 25-32   |       |
| 44-100  | 96.1% | 44-56   | 97.2% | 33-37   | 92.1% |
|         |       |         |       | 38-43   |       |
|         |       | 57-100  |       | 44-48   | 90.7% |
|         |       |         |       | 49-56   |       |
|         |       | 57-72   | 94.6% |         |       |
|         |       | 73-100  |       |         |       |

Em todos os métodos para detecção e classificação se faz necessário a realizar o treinamento do software. Esse treinamento concede ao software a capacidade de reconhecer padrões que definem a diferença entre sexo, idade, rosto de pessoas ou qualquer outro objeto visível. Para realizar a treinamento da maquina é necessária a utilização de um banco de imagens, onde o algoritmo ira realizar a análise e extrair características para serem comparadas com imagens adquiridas posteriormente. Porém é possível notar que a maioria dos algoritmos utilizados para detecção e classificação consegue obter melhores resultados com o alinhamento dos rostos do banco de imagem treinado (Rahman et al., 2013).

Existe uma ampla variedade de métodos publicados na literatura, o que mostra que não existe uma solução única para classificação de gênero e estimativa de idade (Rahman et al., 2013).

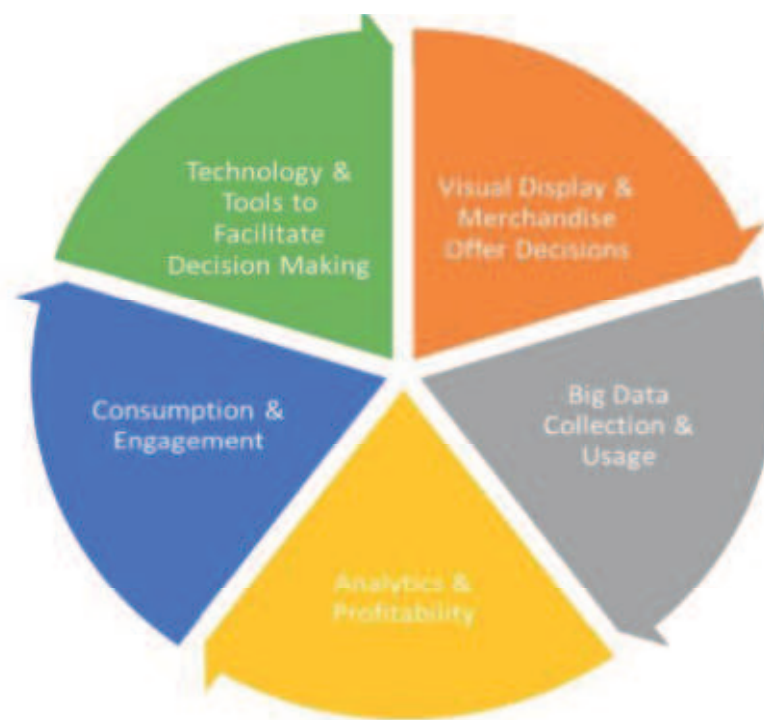
### **3.3 – SINALIZAÇÃO DIGITAL NO COMÉRCIO VAREJISTA**

Roggeveen, Norfalt e Grewal (2016) analisaram a utilização de sinalização digital dentro do comércio varejista e o seu impacto nas vendas e na influência das decisões dos usuários. Os autores observaram que as vendas nos hipermercados são melhoradas quando há a utilização de sinalização digital, e confirmaram através de um estudo de acompanhamento que o aumento nas vendas continuam 5 meses após a sua implantação.

O artigo redigido por Grewal, Roggeveen e Nordfalt (2017) centra-se no futuro do varejo, destacando cinco áreas-chave que estão impulsionando o campo do varejo inteligente: tecnologia e ferramentas para facilitar a tomada de decisões, exibição visual e decisões de oferta de mercadorias, consumo e engajamento, grandes coleta de uso de dados e a análise de rentabilidade.



Figura 7: cinco áreas-chaves do varejo inteligente (Grewal, Roggeveen e Nordfält, 2017)



Com a utilização desse modelo de varejo Os varejistas podem se conectar com seu os clientes, fornecendo-lhes informações direcionadas que podem levar a um envolvimento profundo do cliente, permitindo que os consumidores tomem decisões mais bem informadas sobre quais produtos comprar ou quais serviços consumirem (Grewal, Roggeveen e Nordfalt, 2017).

Em entrevistas realizadas por Piporas, Stylos e Fotiadis (2017), os consumidores solicitaram tecnologias de informação inteligentes aprimoradas que os auxiliassem a encontrar ofertas de produtos que eles precisam, bem como combinar suas necessidades com a oferta mais apropriada. Isso também é indicado nas descobertas empíricas de Kallweit, Spreer e Toporowski (2014) apud Piporas, Stylos e Fotiadis (2017), que propuseram que são necessárias informações que sejam relevantes para atender as necessidades do cliente.

Todas as aplicações de varejo inteligente que se baseiam no comportamento do cliente, receberam um grande interesse do comércio varejista e também de instituições governamentais por causa dos benefícios obtidos em melhor gerir e responder ao interesse dos usuários (Santarcangelo, Farinella e Battiato, 2015).

Dennis et al. (2012) apud Roggeveen, Norfalt e Grewal(2016) mostrou em sua pesquisa que um centro comercial utilizando Sinalização Digital obteve



comportamento de abordagem mais elevado (visitas, gastos, itens comprados, compras realizadas no tempo) do que outro centro comercial que não estava utilizando esta tecnologia. Um dos motivos apontados para um desempenho melhor de comércios com Sinalização digital é que os clientes consideram que as redes de Sinalização Digital são esteticamente agradáveis e fornecem uma imagem mais moderna do ambiente (Newman, Denis e Zaman, 2006 apud Roggeveen, Norfalt e Grewal, 2016) o que reflete sobre os gastos do consumidor, na frequência das visitas e no tempo gasto na loja (Dennis et al., 2012 apud Roggeveen, Norfalt e Grewal, 2016)

Dennis et al. (2014) analisou os fatores que podem ser inseridos nas mensagens exibidas por painéis de Sinalização Digital que influenciam os consumidores no momento da compra. Para ele, a sinalização digital funciona evocando experiências específicas - um intelectual sensorial-afetivo ou de ajuda para a decisão esteticamente agradável - que então afetam positivamente os comportamentos de "abordagem" dos compradores, direta e indiretamente, através das atitudes.

Os estudos sobre a sinalização digital mostram que a sinalização digital, quando posicionada de forma correta em um lugar visível, transmitem a mensagem de forma eficiente (Motamedi et al., 2017) atraindo a atenção do público que tem o *display* em seu campo de visão.

#### **4 - ESPECIFICAÇÃO**

O projeto foi dividido em módulos que possuem funções específicas. Esses módulos são capazes de realizar o treinamento da máquina, a detecção e classificação de rostos, e por fim, a seleção e exibição do anúncio.

Para que o computador consiga realizar classificação através da identificação de características comuns a cada tipo de pessoa, é necessário realizar o treinamento da máquina. De uma forma clara, o treinamento da máquina nada mais é do que "ensinar" a máquina as características que identificam do sexo masculino e feminino como também características próprias de cada faixa etária.

Ao iniciar a aplicação, o algoritmo realiza o treinamento da máquina utilizando as imagens do banco de imagens salvas no disco rígido, essas imagens foram nomeadas de acordo com o sexo e a idade. Após realizar o treinamento a

máquina é capaz reconhecer diferenças das características por sexo e faixa etária. Essas características ficam salvas na memória RAM do computador.

O computador utilizado para desenvolvimento e testes foi um notebook Samsung modelo RV411 com 4Gb de memória RAM, HD SSD de 120Gb e com processador Intel Core i3 2.53GHz.

Foi utilizada parte das imagens do banco de imagens disponibilizada pela AT&T (<http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html>), Esse banco de imagens contém um conjunto de imagens de rosto realizadas entre abril de 1992 e abril de 1994 no laboratório. O banco de dados foi utilizado no contexto de um projeto de reconhecimento facial realizado em colaboração com o Grupo de Discurso, Visão e Robótica do Departamento de Engenharia da Universidade de Cambridge. Esse banco dispõe de imagens de rostos de várias pessoas de ambos os sexos e de várias idades. Muitas delas são fotos das mesmas pessoas desde a infância até à juventude ou à vida adulta.

Para obter um melhor resultado foi analisada manualmente algumas imagem do banco de imagens AT&T. Foi observado que algumas imagens possuíam baixa resolução, e outras estavam danificadas por se tratarem de digitalizações de fotos antigas, algumas até em preto e branco, que estavam com qualidade ruim. Todas as imagens danificadas ou de baixa resolução foram excluídas para evitar erros do algoritmo de treinamento de máquina e, posteriormente, queda na precisão da classificação de gênero e de faixa etária.

Além de utilizar imagens deste banco de imagens, também foi criado um banco de imagens próprio, utilizando pesquisas através da ferramenta Google Images (<https://www.google.com.br/imghp?hl=pt-BR&tab=wi>), coletamos imagens de pessoas de todas as idades e de ambos os sexos. No total foram tratadas 650 imagens obtidas no banco de imagens da AT&T e no google images.

Devido à dificuldade em classificar a idade precisa das pessoas, foi proposta a utilização de classificação dividida por faixas etárias, que faz com que aumente a precisão de acerto. Outro ponto observado foi que crianças tem uma mudança de característica em menor tempo. Em um intervalo de doze anos, há uma mudança considerável nas características físicas das crianças, em especial da face, e também nas características psicológicas, por exemplo, é possível observar no dia-a-dia que os desenhos que uma criança de cinco anos aprecia, crianças de doze anos não tem interesse. Essa mudança é menor em adultos, considerando o mesmo intervalo de

tempo. Uma pessoa com quarenta anos possuem características não muito diferentes de pessoas com cinquenta anos.

O projeto possui uma divisão da faixa etária da seguinte forma:

**Figura 8: Faixas etárias definidas**

### Faixa 1

- de 0 a 2 anos

### Faixa 2

- de 3 a 6 anos

### Faixa 3

- de 7 a 12 anos

### Faixa 4

- de 13 a 19 anos

### Faixa 5

- de 20 a 36 anos

### Faixa 6

- de 37 a 65 anos

### Faixa 7

- Maior que 65 anos

E o sexo foi dividido em dois tipos:

**Figura 9: Gêneros definidos**

### Sexo

- Masculino
- Feminino

Agora que a máquina já consegue “entender” as diferentes características entre homens e mulheres, e diferentes idades, dá-se continuidade ao ciclo de tarefas a serem realizadas pela aplicação. Primeiramente, ela captura imagens em tempo real através de uma câmera instalada em uma posição que consiga capturar uma imagem nítida das pessoas presentes no ambiente em que ele foi instalado. Feito isso, procura identificar rostos, uma vez que um ou vários rostos são identificados, o algoritmo recorta a região da imagem onde o rosto identificado está posicionado e envia para a segunda etapa que é a preparação da imagem.

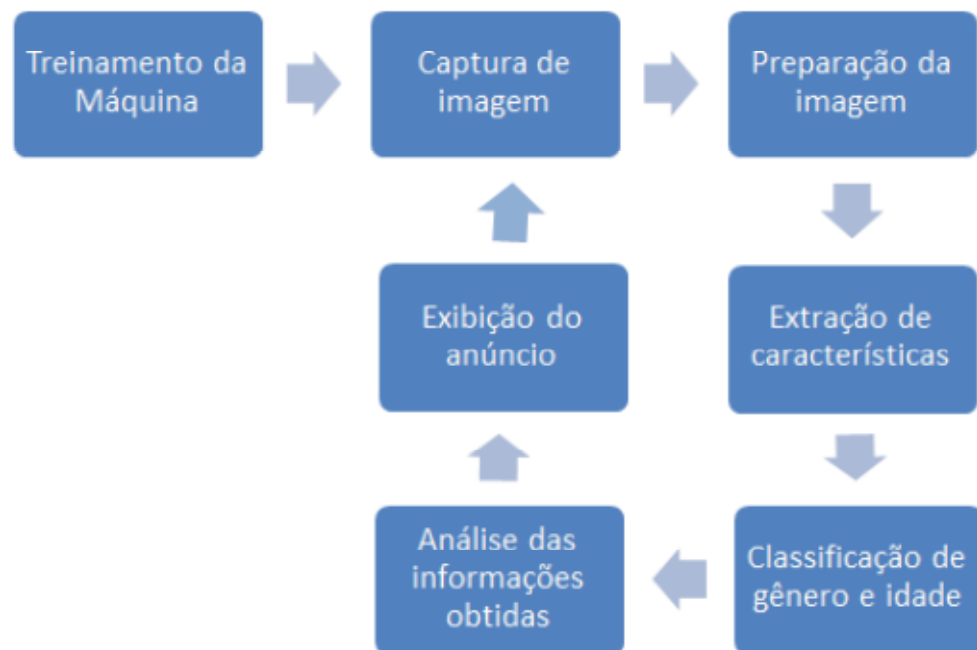
Na Preparação da Imagem, o algoritmo utiliza técnicas de computação gráfica para adiciona um filtro preto e branco às imagens, e também realiza uma equalização na imagem para melhorar sua qualidade. O filtro preto e branco se faz necessário para a extração de características utilizando o método de classificação LBP, que é o passo seguinte.

Após a extração de características, o programa faz a classificação de gênero e idade da imagem obtida através da câmera, comparando-as com as características aprendidas inicialmente com o treinamento da máquina.

Com as informações sobre as classificações de sexo e idade no conjunto de pessoas que foram identificados na imagem, chega ao momento de analisar a característica predominante. O resultado dessa análise é passada para a etapa de exibição do anúncio, que seleciona e exibe em um painel o anúncio com base nas características predominantes das pessoas presentes no espaço público.

Após exibir o anúncio por um período determinado, o algoritmo reinicia o ciclo a partir da captura de imagens através da câmera. O ciclo de funcionamento do programa é representado na imagem abaixo.

Figura 10: Etapas do funcionamento da aplicação



## 6 - PROCESSOS ASSOCIADOS AO DESENVOLVIMENTO

Após adquirir conhecimento sobre os temas que envolvia o projeto, foi decidido que o mesmo seria desenvolvido de forma incremental. O desenvolvimento incremental se baseia na ideia de desenvolver uma pequena parte da aplicação inicial, e incrementar novas funcionalidades aos poucos, adequando ao que já foi implementado anteriormente (Sommerville, 2011). Essa metodologia proporcionou a aquisição de conhecimento de forma também incremental.

De início foi desenvolvido um sistema que detectava o rosto das pessoas através de imagens capturadas por uma câmera acoplada em um computador. Em seguida foi incrementada a capacidade de detectar vários rostos em uma imagem só. O seguinte incremento foi a capacidade de distinguir o sexo das pessoas na imagem. O próximo passo foi adicionar a capacidade de classificar o sexo das pessoas detectadas. E, por fim, a possibilidade de exibir anúncios através da análise das características obtidas.

Para gestão do projeto foi utilizada a metodologia *Scrum* por ser uma metodologia de processo de desenvolvimento ágil que permite que o trabalho seja

dividido em *sprints* que duram de duas a quatro semanas, onde em cada *sprint* é planejado o que será desenvolvido tomando como base um *backlog* previamente definido. As *sprints* deste projeto duravam quatro semanas. A cada início de *Sprint* era planejado o que iria ser desenvolvido durante o período citado.

Como o software funcionava por etapas, como mostra a Figura 10 mostrada anteriormente, foi necessária a utilização do padrão de arquitetura Duto e Filtro, pois o projeto apresenta um início e um fim do processamento onde em cada etapa há uma saída de dados que serão utilizados nos processos seguintes (Somerville, 2011).

## 7 - REQUISITOS DO SISTEMA

O *software* é autônomo, ou seja, depois de configurado pela pessoa responsável por gerenciá-lo, o mesmo funcionará sozinho da seguinte maneira: Após instalado, o anunciante deverá inserir um *pendrive* contendo os vídeos ou fotos com o anúncio que se deseja. após montagem da estrutura de forma correta em um ambiente público, a aplicação deverá executar em tempo real, detectando pessoas que entrem no campo de visão da câmera e extraíndo informações sobre as características das pessoas de forma que seja possível apresentar um anúncio, que estará no dispositivo móvel, previamente configurada para as características obtidas. O software deve guardar informações sobre quantas vezes cada anúncio foi exibido.

O *software* deverá ser instalado em computadores com sistema operacional Windows, permitindo que anunciantes possam usar seus computadores pessoais para implementação e utilização.

## 8 - DESENVOLVIMENTO

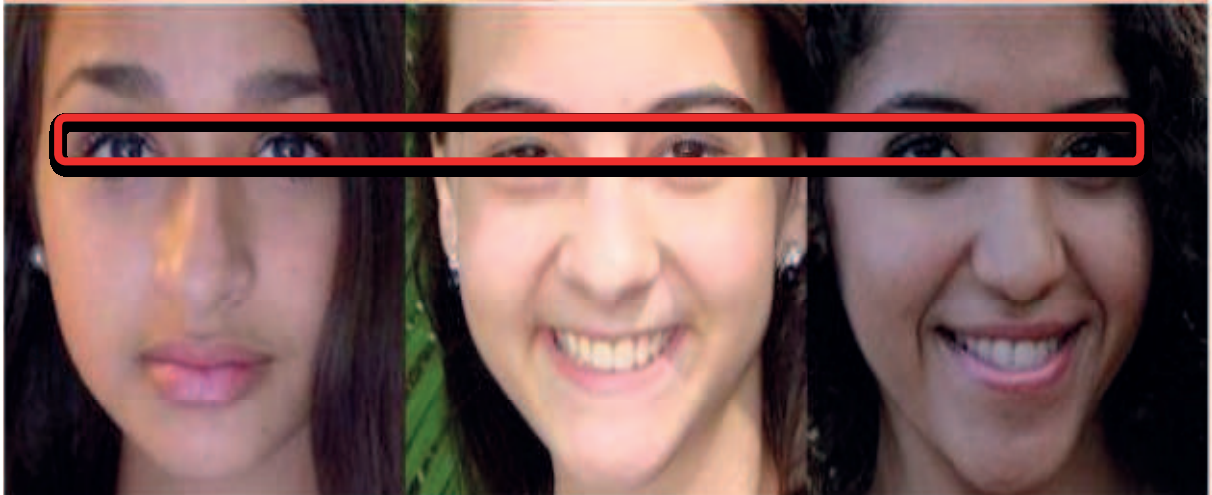
O programa foi todo codificado utilizando a linguagem de programação C++ utilizando a IDE (do inglês *Integrated Development Environment* ou Ambiente de Desenvolvimento Integrado) Visual Studio 2015, desenvolvido pela Microsoft.

### 8.1 – TRATAMENTO DAS IMAGENS DO BANCO DE IMAGENS

Para melhorar a precisão da classificação, foi necessário realizar ajustes nas imagens do banco de imagens. Rahman et al. (2013) propõe em seu trabalho o alinhamento de rostos das imagens tomando como referência os olhos, todas as

imagens devem estar com os olhos na mesma posição relativa, como mostra a imagem abaixo.

**Figura 11: Alinhamento da posição dos olhos feitas pelo algoritmo**



Observe pela linha vermelha que os olhos estão na mesma altura e em uma posição relativa em cada imagem.

O código de alinhamento dos olhos foi desenvolvido por Philipp Wagner (2012) utilizando a linguagem de programação Python, que foi adaptado para aumentar a quantidade de fotos que o algoritmo ajustava por vez. O código recebe como entrada a posição, em pixel, do olho esquerdo e do olho direito e realiza cálculos para rotacionar, se necessário, a imagem de forma que os olhos fiquem alinhados. Para realizar ajustes na codificação que modifica as imagens foi utilizada a IDE Python, que permite o desenvolvimento utilizando a codificação de linguagem de programação de mesmo nome.



Figura 12: Trecho do código em Python responsável pela rotação e alinhamento das imagens com base na posição dos olhos

```

import sys, math, Image

def Distance(p1,p2):
    dx = p2[0] - p1[0]
    dy = p2[1] - p1[1]
    return math.sqrt(dx*dx+dy*dy)

def ScaleRotateTranslate(image, angle, center = None, new_center = None, scale = None, resample=Image.BICUBIC):
    if (scale is None) and (center is None):
        return image.rotate(angle=angle, resample=resample)
    nx,ny = x,y = center
    sx=sy=1.0
    if new_center:
        (nx,ny) = new_center
    if scale:
        (sx,sy) = (scale, scale)
    cosine = math.cos(angle)
    sine = math.sin(angle)
    a = cosine/sx
    b = sine/sx
    c = x-nx*a-ny*b
    d = -sine/sy
    e = cosine/sy
    f = y-nx*d-ny*e
    return image.transform(image.size, Image.AFFINE, (a,b,c,d,e,f), resample=resample)

```

Como dito anteriormente, é necessário que o software realize a aprendizagem de máquina, tendo com base o banco de imagens. Para facilitar o programa a encontrar as imagens, foi criado um arquivo no formato .csv no qual contém todos os nomes das imagens disponíveis no banco.

CSV, (Comma Separated Values) é um formato de arquivo simples, que faz o agrupamento de informações de texto em planilhas para troca de dados entre bancos de dados ou entre aplicativos. Cada linha de um arquivo .csv representa uma linha em uma planilha, e cada célula é separada por um caractere. (<https://ajuda.rdstation.com.br/hc/pt-br/articles/205623999-Como-eu-crio-um-arquivo-CSV->, acessado em 20 de março de 2018).

Os arquivos de imagens foram nomeados de acordo com o sexo e a idade das pessoas fotografadas. A nomenclatura seguia a seguinte estrutura:

### **IdSexoldade.jpg**

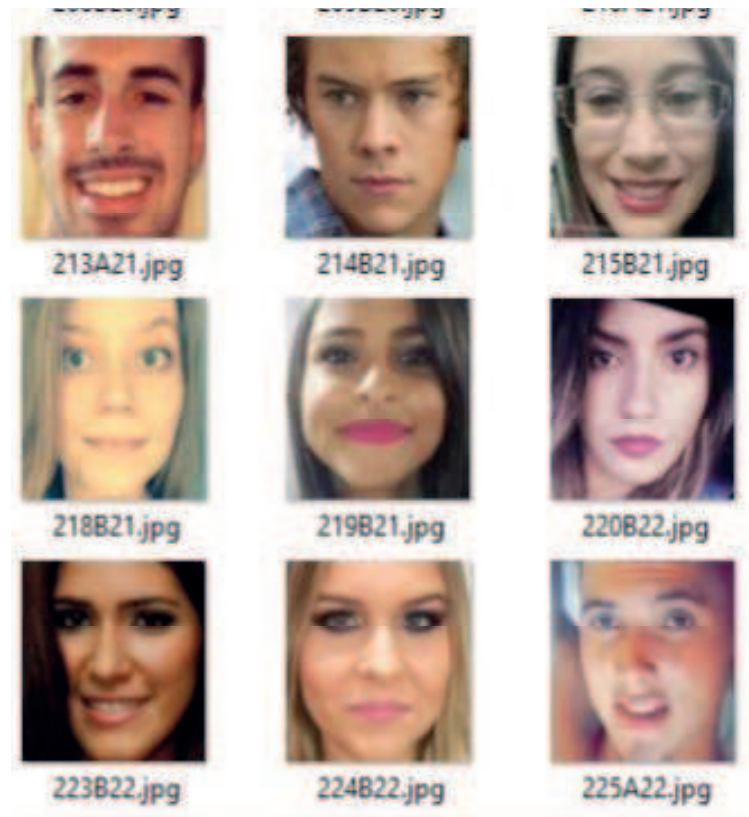
**Id**– Numero de identificação da imagem;

**Sexo** – sexo da pessoa identificada. Masculino ou feminino;



**Idade** – Idade correspondente ao rosto da pessoa na imagem.

Figura 13: Nomes das imagens de acordo com o padrão definido para diferenciação entre sexo e idade



O programa cria um vetor de *strings* onde são armazenados os nomes das imagens que são lidos do arquivo CSV. O vetor será utilizado como um apontador para os arquivos contidos no banco de imagens para que o módulo de treinamento possa buscar e analisar a imagem. Também são criados mais dois vetores que armazenam informações sobre o sexo e a idade da imagem na posição respectiva do vetor que contém o nome das imagens.

Para armazenar informações sobre o sexo, foi criado um vetor booleano, onde pra cada imagem inserida no vetor de nomes dos arquivos de imagem, era inserido um valor na mesma posição do vetor de sexo, *true* caso a imagem for um homem e *false* caso a imagem for uma mulher. A diferenciação entre homem ou mulher era feita através da análise do nome da imagem, A letra A não nome da imagem indicava que a imagem corresponde a uma pessoa do sexo masculino, e a letra B corresponde a uma pessoa do sexo feminino.

Assim como foi feito com o vetor que armazena o sexo das pessoas, foi feito também com o vetor de idade. A cada imagem inserida em uma posição no

vetor de *strings* com os nomes das imagens, também era inserida a idade em um vetor de inteiros, na mesma posição referente ao vetor de imagens, a idade correspondente.

Esse processo torna mais ágil o treinamento de máquina, já que todas as informações necessárias já estarão armazenadas nos vetores.

## 8.2 – TREINAMENTO DE MÁQUINA

Para o treinamento da máquina, é utilizado o algoritmo LBP, que está disponível na biblioteca OpenCV. Ela permite a criação de modelos que são treinados a partir da análise das imagens. Para criação de modelos dois parâmetros são necessários, o primeiro é o endereço onde se encontra a imagem, que obtemos com o vetor de strings que contém esta informação, e o segundo parâmetro é chamado flag, que consiste na característica que é atribuída a esta imagem, o sexo por exemplo, pra isso é utilizado o vetor booleano que indica o sexo.

Figura 14: Trecho da codificação C++ onde é incluída a biblioteca OpenCV no projeto.

```
// ClassificacaofaixaIdade.cpp : Defines the entry poi

#include "stdafx.h"
#include "opencv2/core/core.hpp"
#include "opencv2/contrib/contrib.hpp"
#include "opencv2/highgui/highgui.hpp"
#include "opencv2/imgproc/imgproc.hpp"
#include "ShowVideo.h"
#include "LoadCSV.h"
#include "AnalisaPredicao.h"

#include <iostream>
#include <fstream>
#include <sstream>

using namespace cv;
using namespace std;
```

Ao passar as imagens e as informações sobre o sexo e a faixa etária referente, o algoritmo extrai informações sobre características que são coincidentes em todas as imagens que contém pessoas do sexo masculino, por exemplo. É aqui

onde o software consegue entender a diferença entre os sexos das pessoas e as diferenças que são visíveis em cada faixa etária.

Figura 15: Treinando a máquina através das imagens do banco de imagens

```
// cria um FaceRecognizer e treina o reconhecimento por sexo, idade e ID,  
//a partir das imagens carregadas do banco de imagens:  
modelIdade = createLBPHFaceRecognizer(1, 3, 6, 6, 20.0);  
cout << "modelo idade criado!" << endl;  
  
modelSexo = createLBPHFaceRecognizer(1, 3, 6, 6, 30.0);  
cout << "modelo sexo criado!" << endl;  
  
modelID = createLBPHFaceRecognizer(1, 3, 6, 6, 15.0);  
cout << "modelo ID criado!" << endl;  
  
//realizando aprendizagem de máquina  
cout << "treinando Idade" << endl;  
modelIdade->train(images, labelsIdade);  
  
cout << "Treinando sexo" << endl;  
modelSexo->train(images, labelsSexo);  
  
cout << "treinando ID" << endl;  
modelID->train(images, labelsID);
```

Através do treinamento a máquina obtém a capacidade de, ao detectar rostos, realizar a análise e classificar as pessoas de acordo com as características que foram aprendidas utilizando as imagens do banco de imagens.

### 8.3 – DETECÇÃO DE ROSTOS

Viola e Jones (2001) propôs um algoritmo para detecção e reconhecimento que pode ser utilizado para detectar rostos. A biblioteca OpenCV dispõe deste algoritmo que pode ser usado para detectar rostos. O algoritmo proposto por Viola e Jones possui classificadores em cascata baseados em recursos *Haar* que são utilizados para agilizar o processo de detecção. Esses classificadores são arquivos XML contendo informações que descrevem as características necessárias para que o algoritmo possa detectar os rostos das pessoas de forma rápida.

Figura 16: Classificadores haar disponíveis na biblioteca OpenCV.

|                                      |                  |               |
|--------------------------------------|------------------|---------------|
| haarcascade_eye.xml                  | 13/11/2013 06:44 | Documento XML |
| haarcascade_eye_tree_eyeglasses.xml  | 13/11/2013 06:44 | Documento XML |
| haarcascade_frontalface_alt.xml      | 13/11/2013 06:44 | Documento XML |
| haarcascade_frontalface_alt_tree.xml | 13/11/2013 06:44 | Documento XML |
| haarcascade_frontalface_alt2.xml     | 13/11/2013 06:44 | Documento XML |
| haarcascade_frontalface_default.xml  | 13/11/2013 06:44 | Documento XML |
| haarcascade_fullbody.xml             | 13/11/2013 06:44 | Documento XML |
| haarcascade_lefteye_2splits.xml      | 13/11/2013 06:44 | Documento XML |
| haarcascade_lowerbody.xml            | 13/11/2013 06:44 | Documento XML |
| haarcascade_mcs_eyepair_big.xml      | 13/11/2013 06:44 | Documento XML |
| haarcascade_mcs_eyepair_small.xml    | 13/11/2013 06:44 | Documento XML |
| haarcascade_mcs_leftear.xml          | 13/11/2013 06:44 | Documento XML |
| haarcascade_mcs_lefteye.xml          | 13/11/2013 06:44 | Documento XML |
| haarcascade_mcs_mouth.xml            | 13/11/2013 06:44 | Documento XML |
| haarcascade_mcs_nose.xml             | 13/11/2013 06:44 | Documento XML |
| haarcascade_mcs_rightear.xml         | 13/11/2013 06:44 | Documento XML |
| haarcascade_mcs_righteye.xml         | 13/11/2013 06:44 | Documento XML |

A biblioteca dispõe de vários classificadores que são utilizados para detecção de boca, olhos, nariz, orelha, entre outros. O classificador utilizado neste projeto foi o `haarcascade_frontalface_alt_tree.xml`. Este classificador foi escolhido porque em testes de mesa foi o que menos detectou falsos positivos.

O programa desenvolvido captura um frame do vídeo a cada segundo para analisar e detectar se existem, no frame, rostos detectáveis que possam ser analisados posteriormente. É possível capturar vários rostos em um mesmo frame.

Após detectar os rostos no frame, o software faz um tratamento de cada um cortando a região do rosto e convertendo a imagem para escala de cinza, o que aumenta a precisão dos classificadores LBP utilizados mais afrente, e passa a face detectada para os módulos de classificação de sexo e faixa etária para poderem ser analisados.



Figura 17: Trecho do código responsável pela detecção de faces no frame capturado pela câmera.

```

for (;;) {
    cap >> frame;
    numeroDeFrame++;

    // converte o frame corrente em escala de cinza:
    cvtColor(frame , gray, CV_BGR2GRAY);

    // Procura faces no frame
    haar_cascade.detectMultiScale(frame, faces, 1.1, 2, 0 | CV_HAAR_SCALE_IMAGE, Size(30, 100));

    //procura as faces nos frames capturados
    for (int i = 0; i < faces.size(); i++) {
        // processa cada face encontrada:
        Rect face_i = faces[i];
        // Crop the face from the image. So simple with OpenCV C++:
        Mat face = gray(face_i);
        //redimensiona a imagem para o padrão das imagens salvas no banco de imagens
        Mat face_resized;
        cv::resize(face, face_resized, Size(200, 220), 1.0, 1.0, INTER_CUBIC);
    }
}

```

#### 8.4 – CLASSIFICAÇÃO DE GÊNERO E IDADE

O treinamento realizado ao iniciar o software cria modelos para serem comparados na etapa de classificação de gênero e idade. Os modelos foram criados utilizando as imagens do banco de imagens contendo fotos de homens e mulheres de várias idades.

Ao receber os rostos detectados, o algoritmo LBP cria os rótulos associados à imagem recebida e fazem uma agregação a um histograma (Santacangelo, Farinella e Battiato, 2015). Esses rótulos das imagens são comparados aos modelos criados anteriormente. Ao analisar os modelos o algoritmo faz uma comparação entre os rótulos obtidos da imagem detectada com as características que foram “aprendidas” no treinamento com o banco de imagens, após isso, retorna como resultado a característica que mais se assemelha.

Figura 18: Trecho do código que faz a previsão das características dos rostos detectados.

```

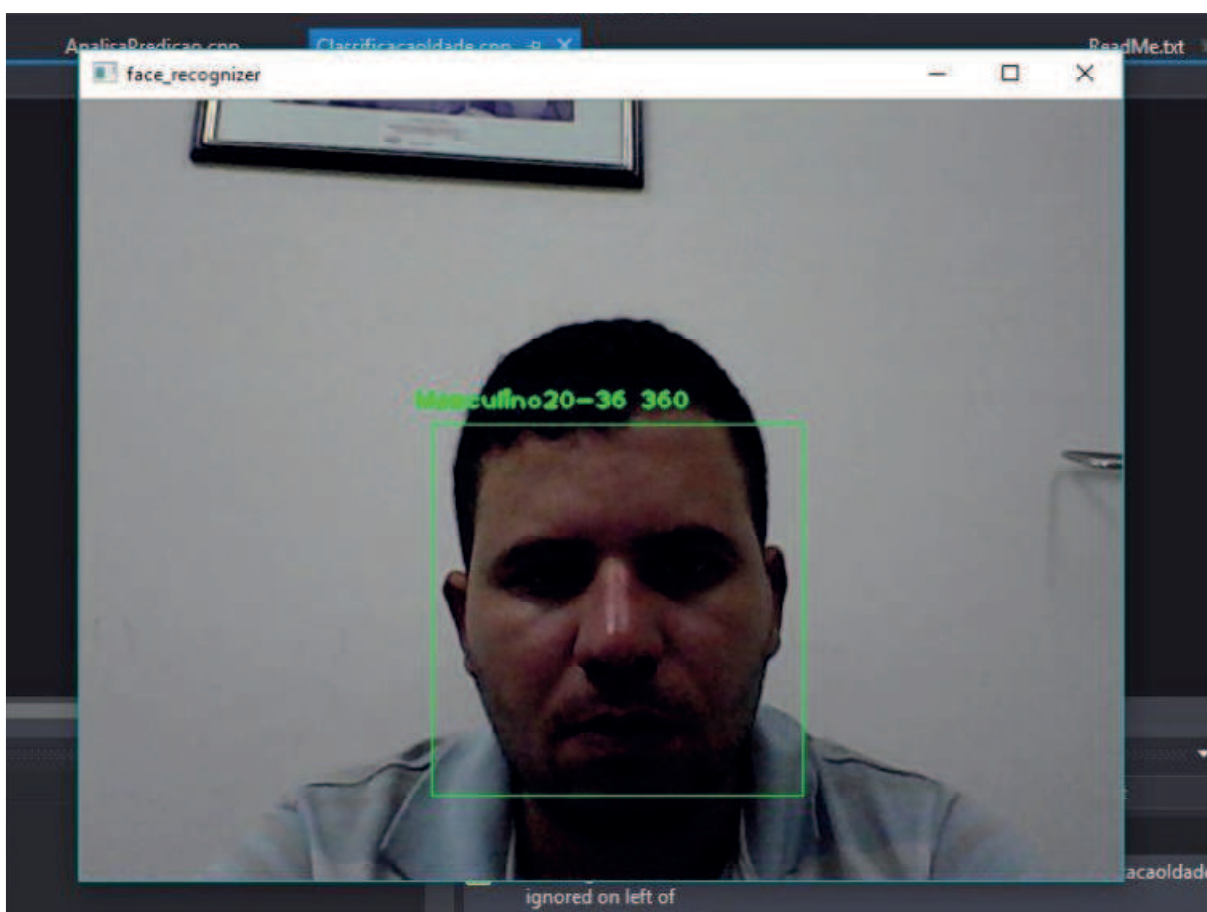
Mat face_resized;
cv::resize(face, face_resized, Size(200, 220), 1.0, 1.0, INTER_CUBIC);

// Realiza uma predeterminação das características do rosto extraído na imagem.
predictionIdade = modelIdade->predict(face_resized);
predictionSexo = modelSexo->predict(face_resized);
predictionID = modelID->predict(face_resized);

```

Após fazer a análise, as classificações realizadas são armazenadas em uma variável que é enviada para a seleção do vídeo que fará a escolha do anúncio baseado no sexo e na faixa etária encontrada.

Figura 19: Programa detectando o rosto e classificando o sexo e a faixa etária, neste caso uma pessoa do sexo masculino e com idade entre 20 e 36 anos.

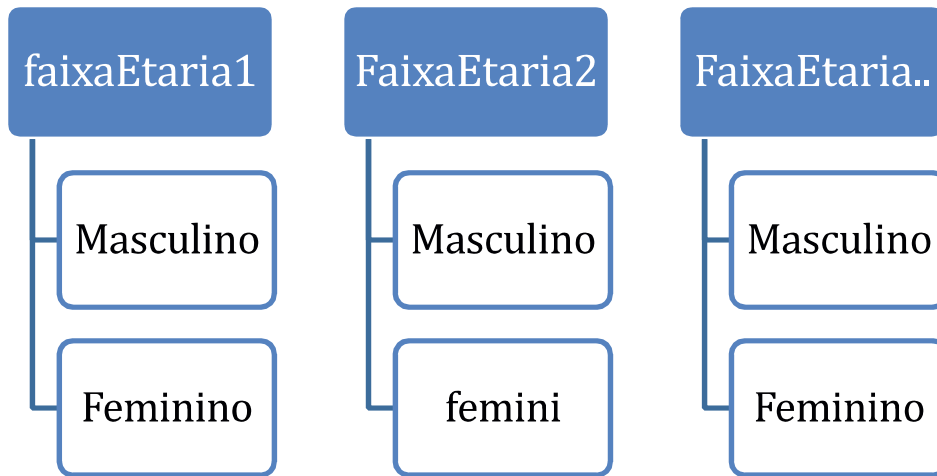


## 8.5 – SELEÇÃO E EXIBIÇÃO DO ANÚNCIO

Todos os anúncios criados devem estar nos formatos de imagem JPG ou PNG, e os arquivos de vídeo devem estar nos formatos MP4, WMV, AVI ou OGG. Anúncios que contém apenas a imagem estática do que deve ser exibido, deverá ser informada por quanto tempo a imagem ficará à mostra. Já os vídeos, executarão completos antes de exibir outros.

Os anúncios devem ser separados em pastas de acordo com a faixa etária e o sexo. O anunciante é quem deverá definir para qual público alvo é direcionado cada anúncio. As pastas seguem a seguinte hierarquia dentro da pasta anúncio:

Figura 20: Estrutura das pastas onde são armazenados os anúncios.



Na exibição, o anúncio é sorteado e é criada uma variável com o nome do anúncio que será incrementada toda vez que o anúncio for exibido para que seja armazenada posteriormente a informação de quantas vezes o anúncio foi exibido.

Figura 21: Trecho da codificação onde é apresentado o anúncio para o expectador.

```

namedWindow("Video", CV_WINDOW_NORMAL);
setWindowProperty("Video", CV_WND_PROP_FULLSCREEN, CV_WINDOW_NORMAL);

while (1)
{
    Mat frame;

    // lê um novo frame para o vídeo
    bool bSuccess = cap.read(frame);

    //se não obteve sucesso, para o vídeo.
    if (!bSuccess)
    {
        break;
    }

    //mostra o vídeo
    imshow("Video", frame);

    if (waitKey(30) == 27) |
    {

```

## 9 – IMPLANTANDO O PROJETO EM UM ESTABELECIMENTO COMÉRCIAL

Após o desenvolvimento do projeto, foi possível realizar testes em um ambiente público para analisar o comportamento do software e também a reação das pessoas presentes dentro do ambiente.



O projeto foi implantado dentro de um comércio de medicamentos. Segundo o dono do estabelecimento, alguns produtos ofertados pela drogaria passavam despercebidos pelas pessoas, pois, em geral, o público vai até uma drogaria para comprar remédios, porém, muitas drogarias vendem diversos tipos de produtos, como produtos de higiene, cosméticos e outros.

A drogaria fica situada no bairro das Malvinas, em Campina Grande. Seus clientes são em geral pessoas de renda média-baixa.

Os equipamentos utilizados foram um *notebook*, que era responsável pelo processamento do programa, uma *webcam*, para coletar imagens do ambiente, e uma TV LCD de 19" que ficaria localizada no balcão da farmácia para exibir o anúncio aos clientes, conforme mostrado nas imagens abaixo.

**Figura 22: TV de 19" utilizada para exibição do anúncio. Em cima da TV está a câmera utilizada para capturar as imagens dos clientes.**



Figura 23: Equipamento posicionado em cima do balcão de atendimento, local visível para todos os clientes presentes.

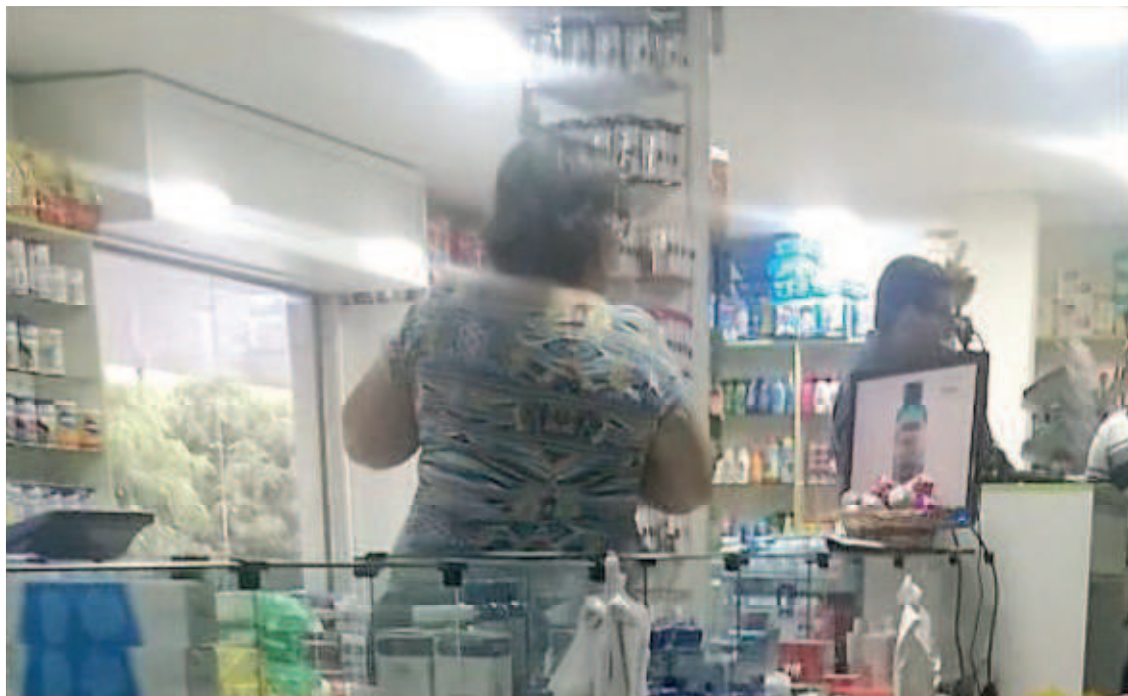


Percebido que as pessoas permaneciam aguardando por um tempo no balcão de atendimento, foi visto juntamente com o dono do estabelecimento que seria o melhor local para fazer os anúncios. Também ficaram definidos quais seriam os produtos que iriam ser promovidos. Após a definição de quais ofertas anunciar, foi criado vídeos apresentando a imagem do produto e o preço. Os vídeos tinham 30 segundos de duração, e foram criados utilizando um *software* de edição chamado Sony Vegas.

Dentre os produtos apresentados para o público estavam escovas de dente infantis, para o público de faixa etária menor que 12 anos, produtos removedores de cravos e espinhas, para o público mais jovem, produtos à base de ômega 3 para pessoas de mais idade, dentre vários outros itens selecionados pelo dono do estabelecimento que deu um direcionamento para quais faixas etárias e sexos se destinariam o anúncio do produto.

Visto que a tela ficava preta antes de exibir o próximo anúncio, foi adicionada no software, na etapa de exibição do anúncio, uma função capaz de mostrar a logomarca da loja, ou seja, sempre que um anúncio era encerrado, iniciava a exibição da marca da empresa.

Figura 24: Cliente visualizando o anúncio.



## 10 – CONCLUSÃO

Por meio da pesquisa realizada foi possível encontrar soluções que tornassem acessível para micro e pequenas empresas de varejo o título de varejo inteligente. Porém é possível observar que, não só o varejo, mas outras áreas podem se utilizar da recomendação de anúncios direcionados ao seu público alvo.

Os trabalhos permitiram a utilização de tecnologias que estão em crescimento no mercado, como a visão computacional e a aprendizagem de máquina.

Os testes realizados na drogaria permitiu a observação de alguns problemas. Um deles, já citado anteriormente, era a falta de uma imagem significativa no intervalo entre um vídeo e outro, que foi resolvida incluindo uma função que insere a logomarca no intervalo entre os anúncios.

Outro problema detectado foi o da iluminação, que, dependendo da posição da câmera, ofusca as imagens, dificultando a detecção dos rostos dos clientes presentes. Isso também foi observado no trabalho de Battiato, Cavallaro e Distante (2016) quando citaram em seu trabalho que “a variabilidade do ambiente tem um

impacto nas imagens a serem processadas, que podem ser de má qualidade devido à distância entre o usuário e a câmera, bem como problemas como oclusão e mudanças de pose e iluminação” (Battiato, Cavallaro e Distante, 2016).

Nos testes realizados não foi possível fazer a análise se a implantação do sistema gerou aumento nas vendas dos produtos anunciados devido ao pouco tempo em que foi instalado o software, porém, o dono do estabelecimento, ficou bastante entusiasmado com a ideia e, segundo ele, o sistema poderia ajuda-lo a aumentar as vendas e principalmente a aumentar a saída de produtos que tem pouca saída devido à falta de exposição visível.

## **10.1 – TRABALHOS FUTUROS**

Este software é bastante promissor, e possui uma vasta utilidade dentro do varejo e do marketing. Visto isto, diversos trabalhos futuros são elencados abaixo visando melhorar o sistema e analisar seus resultados.

Implantar em mais de um estabelecimento e analisar o comportamento das vendas após um período que varia de 2 a 6 meses.

Adicionar ao software a capacidade de rastrear a retina do cliente, buscando determinar se o mesmo encontra-se observando o anúncio e calculando por quanto tempo observa.

Adicionar capacidade de detecção de novas características do usuário, como emoção, atenção, altura e outras características que permitam ao sistema melhorar as recomendações de produtos selecionados.

Portar o software para um dispositivo embarcado que possa ser conectada ao painel onde é exibido o anúncio para os clientes. A utilização de dispositivos embarcados evitará a utilização de um computador de mesa e economizará espaço.

Tornar possível o reconhecimento de pessoas, que ao retornar ao estabelecimento, o software poderá fazer recomendações diferentes e poderá armazenar informações sobre quantas vezes aquela pessoa visitou o estabelecimento.

Ligar o software em rede para que possa haver comunicação e troca de informação coletadas pelo sistema entre diferentes dispositivos em diferentes lugares.

Discussão ética e um estudo sobre o uso das informações capturadas através das câmeras, pois existe uma questão de privacidade das pessoas presentes no ambiente que não deve ser violada.



## BIBLIOGRAFIA

**Site da Microsoft. Disponível em:** <<https://www.microsoft.com/en-us/research/research-area/computer-vision/>>. **Acesso em: 2 Março 2018.**

**Documentação da OpenCV.** Disponível em: <[https://docs.opencv.org/2.4/modules/contrib/doc/facerec/facerec\\_tutorial.html#face-recognition](https://docs.opencv.org/2.4/modules/contrib/doc/facerec/facerec_tutorial.html#face-recognition)>. **Acesso em: 2018 Abril 11.**

AL-KAFF, A. et al. Survey of Computer Vision Algorithms and Applications for Unmanned Aerial Vehicles. **Expert Systems with Applications**, 2017.

BATTIATO, S.; CAVALLARO, A.; DISTANTE, C. Special issue on “Video analytics for audience measurement in retail. **Pattern Recognition Letters**, 2016.

BOUCHRIKA, I. et al. Age Estimation from Facial Images based on Hierarchical Feature Selection. **16th international conference on Sciences and Techniques of Automatic control**, Monastir, Tunisia, Dezembro 2015.

BUBLITZ, F. M. **Abordagem Multiagente para a Propaganda.** Tese de Doutorado, Universidade Federal de Campina Grande - UFCG. Campina Grande. 2013.

COMO eu crio um arquivo CSV? **RD Station.** Disponível em: <<https://ajuda.rdstation.com.br/hc/pt-br/articles/205623999-Como-eu-crio-um-arquivo-CSV->>. **Acesso em: 20 Março 2018.**

DAVIES, N. et al. Personalisation and Privacy in Future Pervasive Display. **ACM**, 2014.

HSU, W.-Y. Interactive e-billboard combined with guide system: Using. **Telematics and Informatics**, Taiwan, 2017.

HUANG, T. **Computer Vision: Evolution and Promise**, 2018.

LAWYER, R. S. Student Driven Digital Signage. **ACM**, 2015.

LEVI, G.; HASSNER, T. Age and Gender Classification using Convolutional Neural Networks, 2015.

MOTAMEDI, A. et al. Signage visibility analysis and optimization system using BIM-enabled. **virtual reality (VR) environments** , 2017.

OPENCV. **Site do OpenCV**. Disponível em: <<https://opencv.org/>>. Acesso em: 2018 Abril 2018.

PANTANO, E.; TIMMERMANS, H. What is Smart for Retailing? **Procedia Environmental Sciences**, 2014.

PAVLOV, V. et al. Application for Video Analysis Based on Machine Learning and Computer Vision Algorithms. **IEEE**, 2013.

PERSPECTIVE, N. R. I. C. V. A. Negative results in computer vision: A perspective. **Image and Vision Computing**, Orlando, FL., USA, 2018.

PRIPORAS, C.-V.; STYLOS, N.; FOTIADIS, A. K. Generation Z consumers' expectations of interactions in smart. **retailing: A future agenda**, 2017.

QUINTANA, M. et al. Improving retail efficiency through sensing technologies: A survey. **Pattern Recognition Letters**, Madrid, Spain, 17 Junho 2016.

RAHMAN, H. et al. An Automatic Face Detection and Gender, 2013.

ROGGEVEEN, A. L.; NORDFÄLT, J.; GREWAL, D. Do Digital Displays Enhance Sales? Role of Retail Format and, 2015.

SAI, P.-K.; WANG, J.-G.; TEOH, E.-K. Facial age range estimation with extreme learning machines. **Neurocomputing**, 2015.

SHU, C. Introduction to Computer Vision, 2008. Disponível em: <[http://people.scs.carleton.ca/~c\\_shu/Courses/comp4900d/notes/lect1\\_intro\\_2up.pdf](http://people.scs.carleton.ca/~c_shu/Courses/comp4900d/notes/lect1_intro_2up.pdf)>. Acesso em: 20 mar. 2018.

SOMMERVILLE, I. **Engenharia de Software**. 9ª Edição. ed. São Paulo: Pearson, 2011.



THE Database of Faces. **AT&T Laboratories Cambridge**. Disponível em: <<http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html>>. Acesso em: 15 Fevereiro 2017.

VIOLA, P.; JONES, M. Rapid Object Detection using a Boosted Cascade of Simple. **Accepted Conference on Computer Vision and Pattern Recognition**, 2001.

VITO SANTARCANGELO, G. M. F. S. B. Gender Recognition: Methods, Datasets and Results, 2015.

WILLEMS, K. et al. The path-to-purchase is paved with digital opportunities: An inventory of. **Technological Forecasting & Social Change**, 2016.

YOGAMEENA, B.; NAGANANTHINI, C. Computer vision based crowd disaster avoidance system: A survey. **International Journal of Disaster Risk Reduction**, Junho 2017.

ZHANG, W. et al. Gender and gaze gesture recognition for human-computer interaction. **Computer Vision and Image Understanding**, 2016.

ZION, B. The use of computer vision technologies in aquaculture – A review. **Computers and Electronics in Agriculture**, 2012.