

Estudo da Relação Entre o Gênero de Espectadores e o Conteúdo de Vídeos

Study of the Relationship Between Spectators Gender and Video Content

Ítalo de Pontes Oliveira^{1, 2, 3}

Eanes Torres Pereira^{1, 4}

Herman Martins Gomes^{1, 5}

Data de submissão: 15/09/2015, Data de aceite: 10/05/2016

Resumo: A sinalização digital pode ser vista como uma forma de publicidade exibida painéis posicionados em locais com circulação ou presença de pessoas, a qual pode ser incrementada com métodos de visão computacional para oferecer recursos de medição da audiência. Neste contexto, o presente trabalho apresenta a avaliação experimental de uma abordagem para estimar, a partir de vídeos, o perfil de espectadores relacionando o gênero e os conteúdos exibidos. A abordagem foi implementada utilizando o detector de faces proposto por Viola-Jones e um classificador de gênero com base em características *Fisherfaces*, validados utilizando a base de faces FERET. O foco central deste trabalho foi o levantamento e teste experimental de hipóteses relativas à associação entre o gênero e categorias de conteúdos exibidos em sinalização digital. A análise realizada a partir de dados obtidos de um ambiente de cantina, em que os espectadores, monitorados por uma câmera de vídeo, assistiam a conteúdos selecionados em uma TV enquanto se alimentavam, mostrou que existe diferença estatística na atenção dada por homens e mulheres em função do tipo de conteúdo. Mais especificamente, (i) homens assistiram mais aos anúncios do que as mulheres; e (ii), os homens e as mulheres se concentraram mais aos anúncios do tipo esportivo do que aos anúncios do tipo jornalístico.

Palavras-chave: sinalização digital, detecção de faces, classificação de gênero

¹Laboratório de Percepção Computacional (LPC), Departamento de Sistemas e Computação (DSC), Centro de Engenharia Elétrica e Informática (CEEI), Universidade Federal de Campina Grande (UFCG), R. Aprígio Veloso, 882 - Bairro universitário, Campina Grande - PB, 58429-900.

²Bolsista do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq)

³{italooliveira@copin.ufcg.edu.br}

⁴{eanes@computacao.ufcg.edu.br}

⁵{hmg@dsc.ufcg.edu.br}

Abstract: Digital signage can be viewed as a form of advertising exhibited in displays positioned in spaces with transit or presence of people which can be enhanced with computer vision methods to offer resources of measuring the audience. In this context, the present work shows the experimental evaluation of an approach for estimating, from videos, the profile of viewers by relating gender and the displayed content. The approach was implemented using a face detector proposed by Viola-Jones and a genre classifier based on Fisherfaces features, which were validated using the FERET face dataset. The central focus of this study was to formulate and to experimentally test hypotheses regarding the association between gender and content categories displayed in digital signage. The analysis of data obtained from a cafeteria environment where viewers, being monitored by a video camera, watched selected contents on a TV whilefeeding, showed that there is an statistical difference in the attention of men and women with regards to the type of displayed contents. More specifically, (i) men payed more attention to news programs than women; and (ii), men and women focused more on sports contents than on news contents.

Keywords: digital signage, face detection, gender classification

1 Introdução

A sinalização digital expõe informações como notícias urgentes, previsão do tempo, entretenimento, etc. para captação da atenção das pessoas, empregando telas de LCD ou plasma de alta resolução, com seus conteúdos atualizados em tempo real [15]. Tem sido utilizada para diversas finalidades, tais como: (i) *marketing*, na promoção e lançamento de produtos e serviços; (ii) anúncios, vendendo tempo de transmissão para fornecedores ou terceiros; (iii) levar informação aos frequentadores de um determinado ambiente; (iv) *infotainment*, cujo propósito é o entretenimento para reduzir a percepção do tempo de espera; dentre outros. Os contextos de aplicação são diversos, incluindo bancos, aeroportos, academias, igrejas, hospitais/consultórios médicos, feiras, etc.

Grandes empresas utilizam e investem em sinalização digital para a divulgação de marcas ou mercadorias, ao mesmo tempo que desenvolvem novas tecnologias que promovem melhor experiência do usuário, provendo maior interatividade, a exemplo de Google[®] [20] e Intel[®] [4]. Padrões também tem sido definidos, como o *Point of Purchase Advertising International* (POPAI) [15], apoiado por empresas como Cisco[®] e LG[®].

O problema de negócio ao qual este trabalho está associado consiste em conhecer o perfil dos espectadores e usar essa informação para otimizar a audiência, visando maximizar o *Return On Investment* (ROI) dos anunciantes. Câmeras podem ser utilizadas para realizar o e monitoramento dos espectadores. A Intel[®], por exemplo, desenvolveu um Arcabouço

de Publicidade Inteligente (API) integrando Análise de Observadores Anônimos (AOA) e mineração de dados, para transmitir em tempo real o anúncio mais adequado ao público presente, provendo interatividade de acordo com o perfil dos espectadores [18].

Want e Schilit [20], consideram que a interatividade poderá ser uma peça chave na sinalização digital visto que, estima-se que o mercado receberá US\$ 5 bilhões em investimentos até 2016 e até então, nenhum modelo de interatividade está bem concebido. Um desafio relevante desta área é o desenvolvimento de métodos para traçar o perfil dos espectadores a partir da análise de vídeo por meio de técnicas como a detecção de faces, a classificação de gênero e idade, dentre outros. O propósito é lançar anúncios personalizados para atingir o público-alvo de maneira mais efetiva e ajustar os conteúdos exibidos em tempo real.

Tratar os consumidores de forma personalizada é motivo de sucesso para algumas empresas, como é o caso das empresas de *e-commerce* que utilizam os *cookies* dos navegadores para personalizar o atendimento dos clientes. A Marisa[®], contratou a empresa AlmapBBDO[®] para desenvolver um aplicativo com anúncios interativos voltado para o público masculino, com versões para Android e iPad [2]. Em um mês, o *e-commerce* da Marisa[®] obteve oito mil novos consumidores e a AlmapBBDO[®] recebeu o prêmio leão de bronze no Festival de Cannes 2014 na categoria de publicidade multimídia. A partir dos fatos relatados, pode-se observar que combinar interatividade e conhecimento do perfil dos clientes/consumidores/espectadores tem forte potencial em aplicações comerciais.

Inserido no contexto apresentado, o presente trabalho apresenta um sistema para análise de atenção dos espectadores, e um estudo de caso em um cenário real. A atenção é analisada indiretamente a partir de evidência proveniente de um detector de faces. Neste trabalho, o perfil dos espectadores é mensurado com respeito à variável gênero, estimada a partir de um classificador treinado com imagens de faces. A principais contribuições do trabalho são as seguintes:

- Avaliação experimental de um sistema para análise da atenção dos espectadores fundamentado em código aberto (*Open Source Computer Vision Library (OpenCV)*⁶), formado por um detector de faces (Viola-Jones) [19] e um classificador de gênero (Fisher-faces) [14]. Foram avaliados o impacto do tamanho do conjunto de treinamento no tempo de processamento e na acurácia do classificador de gênero.
- Estudo de caso envolvendo a análise da atenção de espectadores, a qual foi realizada transmitindo anúncios esportivos e jornalísticos em uma cantina, em que foram testadas hipóteses sobre a preferência a categorias de vídeos em relação ao gênero dos espectadores.

⁶Disponível em <http://opencv.org>, acessado em abril de 2016.

As conclusões deste trabalho, a partir de uma análise estatística sobre as estimativas de gênero obtidas durante a exibição dos vídeos, indicam que existe diferença estatística na atenção dada por homens e mulheres em função do tipo de conteúdo. A análise mostrou que: (i) homens assistiram mais aos anúncios do que as mulheres; e (ii), os homens e as mulheres se concentraram mais aos anúncios do tipo esportivo do que aos anúncios do tipo jornalístico.

O artigo está organizado como descrito a seguir. A Seção 2 apresenta uma discussão sobre os trabalhos relacionados. As questões de pesquisa e a definição das hipóteses são feitas na Seção 3 e os materiais e métodos utilizados são descritos na Seção 4. Na Seção 5, é apresentado um estudo sobre o tempo de processamento para o treinamento do classificador de gênero. Uma análise sobre a acurácia do classificador se encontra na Seção 6. A análise dos vídeos capturados na catina é apresentada na Seção 7. Finalmente, as conclusões do trabalho e propostas para trabalhos futuros são apresentadas na Seção 8.

2 Trabalhos Relacionados

No trabalho de Ravnik [9] foi utilizada uma câmera monocular para segmentação dos espectadores. O algoritmo para detecção de faces foi implementado utilizando a biblioteca de código aberto OpenCV, facilitando a replicação do experimento. Além disso, foram comparados dois tipos de câmera: uma *webcam* logitech Pro 9000 e uma câmera Kinect RGB-D. Foi observado que para distâncias acima de quatro metros, o classificador de gênero obteve melhores resultados com os vídeos da *webcam*. No estudo de caso realizado no presente trabalho foi utilizada uma *webcam* do mesmo modelo, uma vez que os códigos implementados no OpenCV obtiveram melhores resultados nesta câmera.

O sistema proposto em [9] foi aplicado a cenário real de sinalização digital [10], localizado em uma boutique de roupas em Ljubljana, capital da Eslovênia. A atenção dada pelos espectadores foi correlacionada ao tipo de conteúdo transmitido, ao grupo de idade e ao tempo de atenção. Entretanto, não foi apresentada qualquer análise estatística quanto à normalidade dos dados e outros parâmetros que eventualmente causariam ameaças à validade. Similarmente à pesquisa reportada em [9], no presente trabalho também foram realizadas análises da atenção dos espectadores, entretanto, foi realizada uma validação mais robusta do sistema que gera os dados experimentais, fato que aumentou a consistência das inferências. Em [11], o tempo de atenção dos espectadores foi correlacionado ao gênero, sendo verificado que o tempo médio de atenção dos homens (1,2s) foi substancialmente maior do que o tempo médio das mulheres (0,4s), resultado alinhado ao que foi observado no presente trabalho (ver Seção 7).

No trabalho de Yin *et. al.* [5], foram estimados o tempo médio de visualização e o número de pessoas que assistiram a diferentes tipos de anúncios, com o propósito de definir o perfil dos espectadores. Naquele trabalho, observou-se uma relação inversa quanto ao tempo

de atenção e o número de pessoas que assistiram às notícias globais e locais, i.e., enquanto as notícias locais atraíram um número maior de espectadores do que notícias globais, as notícias globais receberam tempo de atenção do telespectador inversamente proporcional à quantidade de espectadores presentes. Para outros tipos de anúncios, os tempos de atenção não apresentaram tais variações. Entretanto, estas análises se limitaram ao tipo de conteúdo transmitido e ao tempo de visualização, não sendo apresentadas considerações demográficas, tais como relacionar o conteúdo ao gênero (que foi realizado no presente trabalho) ou à idade do espectador.

O presente trabalho preenche lacunas nas pesquisas supracitadas ao apresentar uma validação experimental do classificador de gênero, que fornecerá os dados necessários para análise da atenção dos espectadores. Esse tipo de validação não é observada em outros trabalhos relacionados [9], [10], [11]. Além disso, a abordagem foi implementada utilizando código-fonte aberto, permitindo replicação das análises feitas. Por fim, correlacionar o gênero do espectador ao conteúdo tem se mostrado relevante para que as empresas atinjam seu público-alvo com maior eficácia [5].

3 Questões de Pesquisa e Definição de Hipóteses

Antes de realizar análises da atenção dos espectadores, realizou-se uma validação do classificador de gênero, o qual gerou os dados de entrada para as inferências. Neste sentido, foram formuladas questões de pesquisa associadas à relação entre a quantidade de imagens no conjunto de treinamento e o tempo de processamento, e à acurácia [16] do classificador de gênero.

Foram realizados testes paramétricos para determinar se o tempo de processamento para os diferentes conjuntos de treinamento obedecem a uma distribuição normal e para determinar se os tempos de processamento para os diferentes conjuntos de treinamento são iguais. A hipótese formulada é que, quanto maior o conjunto de treinamento, maior será o tempo de processamento. Por fim, foi testado se, ao aumentar o número de imagens no conjunto de treinamento, ocorre um aumento na acurácia do classificador de gênero. Desta forma, foi realizado o cálculo dos índices de correlação entre as medidas de acordo com o tipo de distribuição previamente concebido.

Mais especificamente, as seguintes questões de pesquisa foram formuladas:

Q₁: Os tempos para treinar classificadores de gênero, com conjuntos de treinamento de diferentes tamanhos, são iguais?

Q₂: Os tempos de treinamento obedecem a uma distribuição normal?

Q₃: O tamanho do conjunto de treinamento afeta a acurácia do classificador de gê-

nero?

Na Tabela 1, podem ser observados os testes de hipóteses realizados neste trabalho com respeito ao classificador de gênero. O superconjunto C contém conjuntos de treinamento. Ou seja, c_i é o conjunto de treinamento i , $N(\mu, \sigma^2)$ indica uma distribuição normal, com média μ e variância $\sigma^2 > 0$. A função $t(c_i)$ retorna o tempo de processamento para o conjunto de treinamento c_i . Por fim, a função $corr()$ retorna a correlação com a acurácia definida por ACC.

Tabela 1. Matriz de hipóteses considerando o tempo de processamento e o conjunto de treinamento.

Hipóteses Nulas	Hipóteses Alternativas
$H_{1_1} : \forall c_{i,j} \exists C_{i \neq j} t(c_i) = t(c_j)$	$H_{1_2} : \forall c_{i,j} \exists C_{i \neq j} t(c_i) \neq t(c_j)$
$H_{2_1} : \forall c \exists C f(t(c)) = N(\mu, \sigma^2)$	$H_{2_2} : \forall c \exists C f(t(c)) \neq N(\mu, \sigma^2)$
$H_{3_1} : corr(ACC, Tam) \geq 0,7$	$H_{3_2} : corr(ACC, Tam) < 0,7$

A questão de pesquisa central desta investigação, no contexto da análise da atenção dos expectadores, é a seguinte:

Q₄: Homens e mulheres prestam atenção igualmente a anúncios jornalísticos e esportivos?

As variáveis envolvidas no experimento projetado para responder à pergunta acima são apresentadas na Tabela 2. O experimento considera dois fatores com dois níveis cada. Isto é, os gêneros possíveis são masculino e feminino, enquanto que o tipo de anúncio transmitido pode ser esportivo ou jornalístico. Sendo assim, a base das análises consistirá na combinação deste fatores.

Tabela 2. Variáveis analisadas.

Fator	Níveis
Gênero	Masculino
	Feminino
Tipo de anúncio	Esportivo
	Jornalístico

Por fim, na Tabela 3, pode ser observada uma matriz de hipóteses em que na coluna da esquerda são apresentadas as hipóteses nulas e, na coluna da direita as hipóteses alternativas associadas à questão de pesquisa **Q₄**. A variável μ representa a média da quantidade de espectadores, os índices M e F rotulam o gênero em Masculino ou Feminino, respectivamente.

Enquanto os sobrescritos e e j , indicam o tipo do conteúdo transmitido em vídeos esportivos e jornalísticos, respectivamente. Por exemplo, μ_M^e é a média de homens detectados que assistiram ao anúncio esportivo.

Tabela 3. Matriz de hipóteses considerando o gênero do telespectador e o tipo do conteúdo transmitido.

Hipóteses Nulas	Hipóteses Alternativas
$H_{3_1}: \mu_M^j = \mu_M^e$	$H_{3_2}: \mu_M^j \neq \mu_M^e$
$H_{4_1}: \mu_F^j = \mu_F^e$	$H_{4_2}: \mu_F^j \neq \mu_F^e$

4 Materiais e Métodos

O estudo de campo do sistema de sinalização digital, para análise da atenção dos espectadores correlacionado o gênero ao tipo de conteúdo, foi realizado em uma cantina localizada no campus sede da Universidade Federal de Campina Grande, Paraíba, a qual possui mais de sete mil discentes e mais de 800 docentes vinculados. Para sinalização digital, foi utilizado um televisor LG 42" com resolução *Full HD* (1920×1080) e formato de tela 16:9. Tal equipamento está de acordo com o padrão definido no POPAI [8] para a exibição de vídeos na sinalização digital. A câmera utilizada para captura dos espectadores foi a Logitech HD Pro *Webcam C920* com gravação *Full HD* codec H.264 e 24 fps. Os algoritmos foram desenvolvidos na linguagem de programação C/C++, usando a biblioteca OpenCV 2.4.8 com *GNU Compiler Collection (GCC)* versão 4.8.2 executados em uma máquina Intel® Core™ i7 3770 3,40GHz com 8GB DDR3 de memória principal e placa de vídeo nVIDIA 210/1GB com GNU/Linux Ubuntu 14.04 LTS 64 bits.

Na Figura 1, é apresentado o ambiente do experimento. Na Subfigura 1(a), pode ser observada a estrutura montada para transmissão da sinalização digital e a captura do vídeo dos espectadores e a Subfigura 1(b) representa a planta baixa do estabelecimento comercial (~45m²), em que o equipamento de transmissão da sinalização é destacado em vermelho, situado frontalmente aos espectadores que utilizam o local.

Na Figura 2, pode ser observado o diagrama de fluxo que representa o funcionamento lógico da aplicação desenvolvida para realizar a análise da atenção dos espectadores. Os vídeos capturados das pessoas no estabelecimento comercial são processados quadro-a-quadro, as faces dos vídeos são detectadas automaticamente pelo OpenCV usando o algoritmo de Viola e Jones [19] e então equalizadas, pois segundo Khan *et. al.* [12] o classificador de gênero apresenta melhores resultados após a equalização dos histogramas. Em seguida, o classificador de gênero previamente treinado, recebe a face por parâmetro para determinação do gênero. Após isso, o contador de gênero é atualizado, i.e., as quantidades das faces detectadas

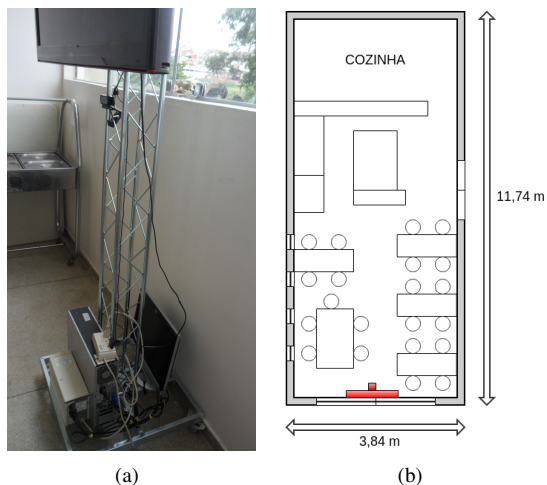


Figura 1. Estrutura e cenário da realização do experimento: (a) equipamento utilizado na sinalização digital; (b) planta baixa do local.

para cada gênero no quadro são armazenadas para posterior análise.

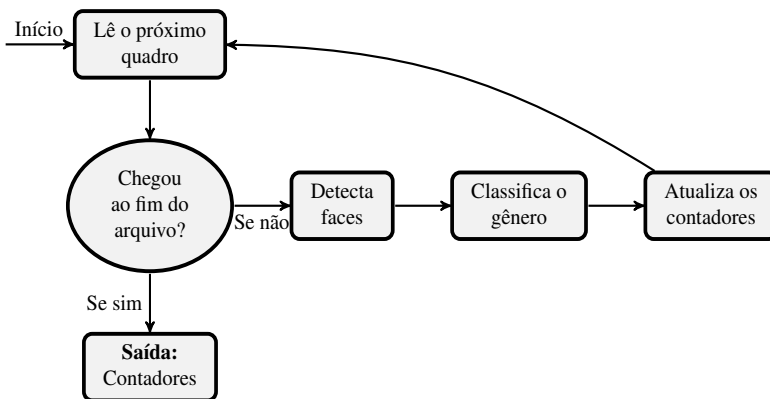


Figura 2. Diagrama de fluxo da aplicação.

4.1 Considerações Éticas

O experimento está de acordo com a Resolução nº 466 do Conselho Nacional de Saúde ⁷, que define diretrizes e normas regulamentadoras de pesquisas envolvendo seres humanos. Os procedimentos adotados nesta pesquisa asseguraram a confidencialidade, a privacidade e a proteção da imagem dos espectadores.

5 Tempo de Treinamento do Classificador de Gênero

Para validar o classificador de gênero foi utilizada a base de dados FERET, em que 200 faces masculinas e 200 faces femininas foram selecionadas aleatoriamente para compor o grupo de imagens de teste. Foi utilizado o classificador de gênero Fisherfaces [14], disponível na biblioteca OpenCV. Alterou-se o tamanho do conjunto de treinamento para observar a variação da acurácia e do tempo de processamento para os seguintes cenários: $C = \{200, 250, 300, 350, 400, 450, 500, 550\}$, em que cada valor representa a quantidade de imagens no conjunto de treinamento. Para se determinar o intervalo de tempo, foi utilizada a estrutura de dados `timespec` do cabeçalho `time.h` da linguagem GNU C com precisão de nanosegundos, como pode ser observado no seguinte trecho de código:

```
1. struct timespec inicio={0,0}, fim={0,0};
2. clock_gettime( CLOCK_MONOTONIC, &inicio );
3. // <Colocar aqui> o trecho de código a ter
   // o tempo de processamento medido
4. clock_gettime( CLOCK_MONOTONIC, &fim );
5. double tempo_de_execucao;
6. tempo_de_execucao <<
   (((double) fim.tv_sec + 1.0e-9*fim.tv_nsec)
  - (double) inicio.tv_sec + 1.0e-9*inicio.tv_nsec));
```

em que a variável `tempo_de_execucao` armazena o tempo decorrido entre o início e o final da execução do algoritmo.

Na Figura 3, pode ser observado o tempo médio de processamento considerando diferentes tamanhos do conjunto de treinamento, com intervalos de confiança de 95% calculados com a função `t.test()$conf.int` da linguagem R. Assim, pode-se afirmar que houve diferença estatística para treinar com diferentes quantidades de imagens no grupo de treinamento, i.e., quanto mais imagens na base de treinamento, maior o tempo de processamento.

⁷ Disponível em: http://bvsmms.saude.gov.br/bvsm/saudelegis/cns/2013/res0466_12_12_2012.html, acessado em abril de 2016.

Como consequência, pode-se rejeitar a hipótese nula $H_{3_1}: \mu_M^j = \mu_M^e$, ou seja, os tempos de treinamento para diferentes quantidades de imagens foram diferentes.

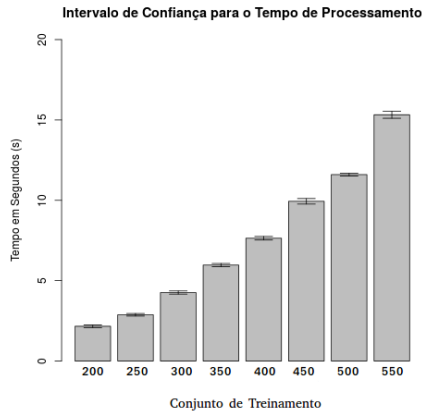


Figura 3. Intervalo de confiança de 95% para o tempo de treinamento do classificador de gênero, em função dos diferentes tamanhos de conjunto de treinamento.

Na Tabela 4, podem ser observados os resultados dos testes de normalidade Anderson-Darling [1] e Shapiro-Wilk [13], os quais evidenciam que os dados são normais, i.e., os valores obtidos foram menores do que 0,05. Por esta razão não é possível rejeitar a hipótese nula $H_{0_1}: \forall t(c_i) := N(\mu, \sigma^2)$. A normalidade indica se os dados obedecem a uma distribuição normal, fato que ajuda a determinar quais os métodos estatísticos mais adequados ao problema.

Tabela 4. Teste de normalidade para os diferentes tamanhos de conjunto de treinamento do classificador de gênero.

Treinamento	Anderson-Darling	Shapiro-Wilk
200	1,2e-05	5,4e-06
250	2,2e-05	6,5e-06
300	7,8e-03	1,5e-05
350	3,9e-15	4,7e-13
400	2,2e-16	6,4e-16
450	6,2e-15	1,2e-10
500	1,8e-02	8,6e-03
550	2,2e-16	2,2e-16

Em seguida, foi utilizada a função `levene.test()` do pacote “car” da linguagem R para teste de homocedasticidade, conforme proposto por John Fox [6], com intuito de verificar se os dados são estatisticamente iguais. Na Tabela 5, os resultados indicam que os dados são diferentes entre si, pois o *Pr* foi menor que o *F-Value* com sete graus de liberdade, rejeitando-se, desta forma, a hipótese nula $H_{0_1}: \forall t(c_i) := N(\mu, \sigma^2)$. Ou seja, o tempo de processamento foi diferente para as diferentes quantidades de imagens no conjunto de treinamento.

Tabela 5. Teste de homocedasticidade para o tempo de processamento do classificador de gênero.

Graus de Liberdade	F-Value	Pr(> F)
7	19,9	2,2e-16

6 Análise Sobre a Acurácia do Classificador de Gênero

Neste experimento, como no anterior, foram alterados os tamanhos dos conjuntos de treinamento do classificador, mas agora o propósito foi avaliar o impacto dos diferentes tamanhos na acurácia (ACC) do classificador. A Equação 1 [16] define formalmente a acurácia.

$$ACC = \frac{VP + VN}{P + N}, \tag{1}$$

em que, *VP* corresponde à taxa de verdadeiros positivos. Convencionamos que o gênero masculino seria a classe positiva e o feminino, a negativa. Desta forma, *VP* corresponde à quantidade de homens classificados corretamente, *VN* corresponde à taxa de verdadeiros negativos, i.e., corresponde à quantidade de mulheres classificadas corretamente, *P* é a quantidade total de homens no conjunto de teste e *N* a quantidade total de mulheres do conjunto de teste.

A Figura 4 exibe a acurácia para os diferentes conjuntos de treinamento. Pode-se observar que há pouca variação estatística entre eles. Para verificar a variância da acurácia, foi realizado o `levene.test()`. Como pode ser observado na Tabela 6, o valor calculado para a homocedasticidade foi menor que o *F-Value*. Portanto, pode-se rejeitar a hipótese nula $H_{1_1}: \forall t(c_i) = t(\mu(C))$. Ou seja, existe diferença estatística para o valor da acurácia obtida utilizando-se diferentes tamanhos de conjunto de treinamento, assim como existe para o tempo de processamento (discutido na seção anterior). Por essa razão, faz-se necessário calcular a correlação entre a quantidade de imagens e a acurácia do classificador, com o objetivo de observar se ao se acrescentar imagens no conjunto de treinamento, ocorre aumento da acurácia ou vice-versa.

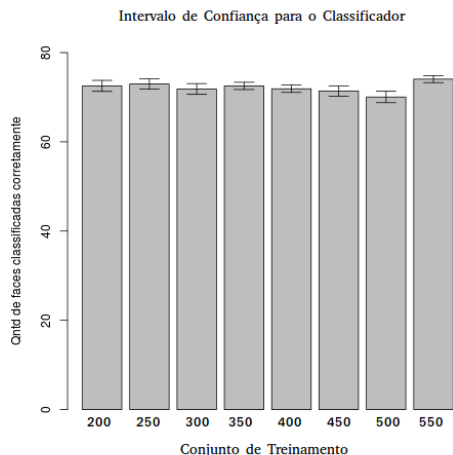


Figura 4. Intervalo de confiança de 95% para a acurácia do classificador de gênero.

Tabela 6. Teste de homocedasticidade para a acurácia.

Graus de Liberdade	F-Value	Pr(> F)
7	4,8	2,318e-05

Para aferir se há relação entre a acurácia do classificador de gênero e a quantidade de imagens no conjunto de treinamento, podem ser utilizados os coeficientes de correlação. As seguintes medidas foram utilizadas como indicadores: Coeficiente de Correlação Linear de Pearson (CCLP) [7] e o Coeficiente de Correlação dos Postos de Spearman (CCPS) [17]. O CCLP permite identificar a linearidade entre dois conjuntos de medidas e assim, quanto maior o valor de CCLP, significa que há mais correlação entre o conjunto de treinamento e a acurácia. O CCLP entre dois conjuntos de medidas $A = \{a_k\}$ e $B = \{b_k\} \mid k = 1, 2, \dots, K$ é dado por

$$CCLP = \frac{\sum_{k=1}^K (a_k - \bar{x}_a) \cdot (b_k - \bar{x}_b)}{\sqrt{\sum_{k=1}^K (a_k - \bar{x}_a)^2} \cdot \sqrt{\sum_{k=1}^K (b_k - \bar{x}_b)^2}}, \tag{2}$$

em que \bar{x}_a e \bar{x}_b significam as médias amostrais das medidas dos conjuntos A e B , respectivamente.

O CCPS indica a dependência estatística entre dois conjuntos ordenados de medidas

Tabela 7. Correlação entre quantidade de imagens no conjunto de treinamento e a acurácia do classificador de gênero.

Qntd. de Imagens	CCLP	CCPS
200	-0,47	-0,55
250	-0,48	-0,57
300	-0,21	-0,31
350	-0,32	-0,47
400	-0,20	-0,41
450	-0,37	-0,49
500	0,10	-0,25
550	0,01	-0,30

A e B , sendo matematicamente expresso por

$$CCPS = 1 - \frac{6 \cdot \sum_{k=1}^K d_k^2}{K(K^2 - 1)} \quad (3)$$

em que K é o número de medidas e d_k é a diferença entre os postos das medidas ordenadas $a_k \in A$ e $b_k \in B$.

Na Figura 5, pode-se observar um gráfico de dispersão que relaciona a quantidade de imagens no conjunto de treinamento e a acurácia do classificador. Nota-se uma fraca tendência no canto superior direito de todos os gráficos, sugerindo que há pouca relação no aumento do número de imagens com a acurácia do classificador. Os resultados dos coeficientes de correlação de Pearson e Spearman para os dados apresentados na Figura 5 são mostrados na Tabela 7, indicando que não há forte correlação (positiva ou negativa) entre as medidas. Portanto, rejeita-se a hipótese nula $H_{2,1}: \forall c_i \exists | \text{cor}(\text{ACC}) | \geq 0,7$, em que o tamanho do conjunto de treinamento possui correlação fraca com a acurácia. Mais especificamente, o aumento no tamanho do conjunto de treinamento tem baixo impacto na acurácia do classificador de gênero testado.

7 Aplicando o Classificador de Gênero na Sinalização Digital

O classificador de gênero validado foi aplicado em um cenário real de sinalização digital, em que vídeos do tipo esportivo e jornalístico foram exibidos em uma cantina da Universidade Federal de Campina Grande, Paraíba, enquanto que vídeos dos espectadores assistindo aos anúncios eram capturados. Tais vídeos, foram posteriormente processados com a aplicação do detector de faces e do classificador de gênero. Os vídeos dos espectadores foram capturados em dias alternados e no mesmo período do dia. Os resultados obtidos pelo

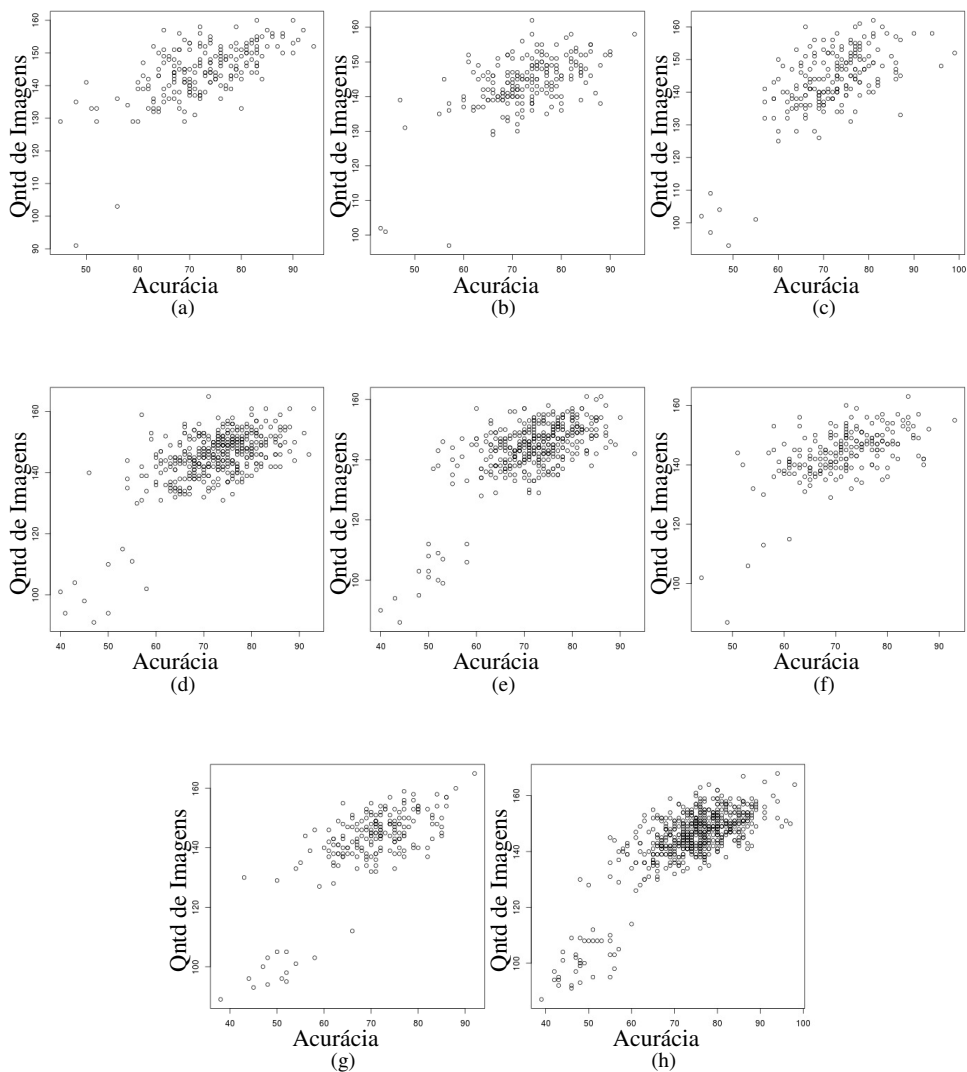


Figura 5. Relação entre a quantidade de imagens no conjunto de treinamento e a acurácia do classificador, utilizando diferentes quantidades de imagens no conjunto de treinamento: (a) 200; (b) 250; (c) 300; (d) 350; (e) 400; (f) 450; (g) 500; (h) 550 imagens, respectivamente.

classificador de gênero nos vídeos coletados são discutidos nesta seção.

Na Figura 6 relaciona-se a quantidade média de pessoas observando a sinalização digital ao tipo de conteúdo que está sendo transmitido, com respectivos Intervalos de Confiança (IC) de 95% no teste *t* de *student* [3]. Desta maneira, pode-se observar que as pessoas se mantêm mais concentradas nas transmissões do tipo esportiva 6 (a), isto é evidenciado com o menor intervalo de confiança nos gráficos, indicando que as pessoas se mantêm menos dispersas. O oposto é observado no gráfico das transmissões jornalísticas 6 (b), em que o intervalo de confiança é maior, conseqüentemente, os espectadores se concentraram menos nestes tipos de anúncios. Tais observações corroboram com as análises feitas por Ravnik e Solina [9], em que o público masculino se concentrou mais nos anúncios do que o público feminino, entretanto, o tipo de conteúdo não foi especificado naquele trabalho.

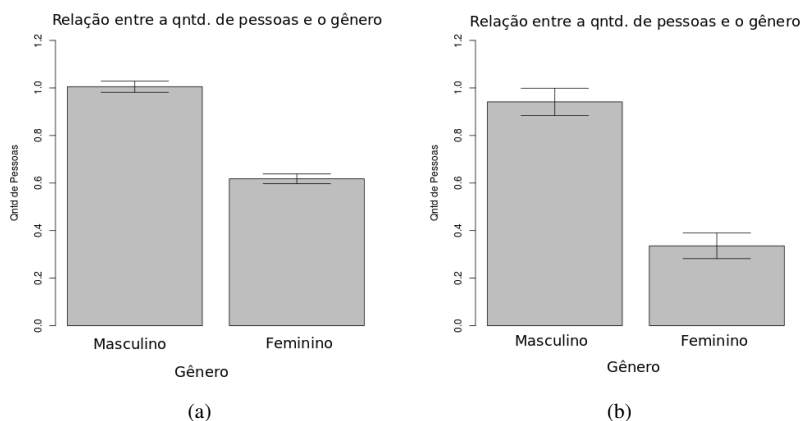


Figura 6. Transmissões de diferentes tipos de vídeo transmitidos: (a) esportivos; (b) jornalísticos

Na Figura 7, pode ser observado que os homens se concentraram mais nos anúncios esportivos do que nos jornalísticos devido ao menor intervalo de confiança [3]. Além disso, o estudo mostrou que homens atendem, de forma estatisticamente igual, aos dois tipos de anúncio. Enquanto isso, as mulheres se concentram mais em anúncios esportivos e, proporcionalmente, menos em anúncios jornalísticos. A partir das observações anteriores, não se pode rejeitar a hipótese nula $H_{3_1}: \mu_M^j = \mu_M^e$, pois há interseção nos intervalos de confiança. Conseqüentemente, não se pode afirmar a qual conteúdo os homens são mais frequentes. Quanto à hipótese nula $H_{4_1}: \mu_F^j = \mu_F^e$, ela pode ser rejeitada em favor da hipótese alternativa $H_{4_2}: \mu_F^j \neq \mu_F^e$, ou seja, pode-se afirmar que as mulheres ocorrem em maior frequência nos anúncios do tipo esportivo do que nos anúncios do tipo jornalístico. As constatações anteriores se

alinham ao que foi observado previamente em [10] e em [11], em que o gênero do espectador tem uma alta correlação com o conteúdo transmitido.

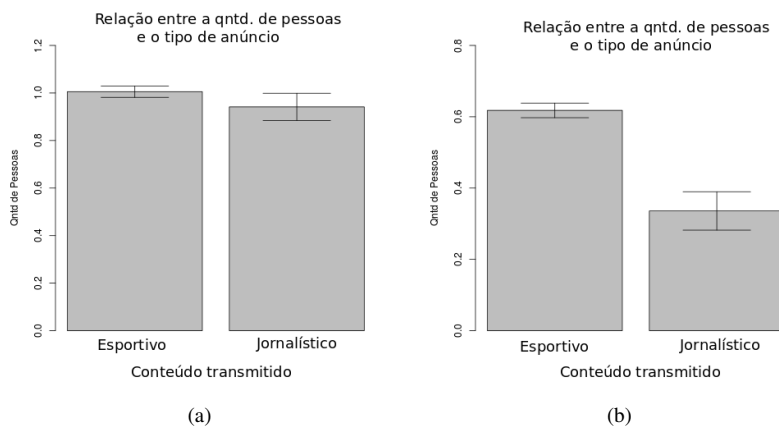


Figura 7. Relação entre o gênero e o tipo de conteúdo: (a) masculino; (b) feminino

Como exemplo de aplicação prática dos achados anteriores, ao observar diferença estatística na atenção dos espectadores considerando seu gênero, o dono de um determinado negócio, pode alterar o conteúdo transmitido automaticamente, à depender do interesse dos clientes, com o propósito de melhorar a eficácia nas transmissões de sinalização digital.

7.1 Empacotamento

Os arquivos utilizados neste experimento foram agrupados para disponibilização pública⁸, entre eles, a relação dos vídeos transmitidos no experimento e os *shell scripts* para automatizar a execução do algoritmo de classificação de gênero.

7.2 Ameaças à validade

Como o experimento foi desenvolvido em um ambiente de pequenas dimensões, com circulação limitada de pessoas, os resultados do experimento podem não refletir o comportamento geral de uma ampla população de espectadores. Para reduzir tal ameaça, poderiam ser utilizados outros ambientes mais amplos para replicação e comparação dos resultados. Além disso, a acurácia do classificador de gênero (aproximadamente 80%), pode ter impacto na confiança das inferências realizadas.

⁸Endereço para download: <https://goo.gl/iS10Hn>.

8 Conclusões e trabalhos futuros

O presente trabalho focou na formulação e teste de hipóteses no cenário de sinalização digital. Foram investigadas, preliminarmente, as seguintes relações: quantidade de imagens no conjunto de treinamento versus (tempo de treinamento e acurácia de um classificador de gênero a partir de imagens de faces). Tal classificador foi utilizado em um sistema de análise indireta da atenção dos espectadores, com o propósito de investigar a relação entre o gênero do espectador e conteúdo transmitido.

Foi observado que ao se elevar o número de imagens no conjunto de treinamento, não ocorre aumento significativo na taxa de acerto do classificador, entretanto, como esperado, o tempo de treinamento aumenta consideravelmente. Além disso, foi observado que vídeos esportivos retêm maior concentração dos espectadores, independentemente do gênero do espectador, e que homens assistem mais aos anúncios independentemente do conteúdo transmitido, enquanto que as médias de mulheres assistindo anúncios jornalísticos e esportivos são diferentes, i.e., elas estão mais presentes nos anúncios esportivos do que durante os anúncios jornalísticos.

As contribuições centrais deste trabalho residem nas análises realizadas no contexto de sinalização digital, utilizando um modelo de classificação de gênero validado estatisticamente. Resultados sem tanta análise estatística, gerando inconsistências em suas análises.

Como trabalhos futuros, pretende-se: (i) aprimorar a acurácia do classificador de gênero para maior confiança nos resultados; (ii) utilizar métodos de rastreamento facial para diminuir o custo computacional do sistema de análise; (iii) considerar outros fatores demográficos, tais como: faixa etária, tempo de visualização de cada indivíduo, etc.; e (iv) incluir exibição de vídeos contendo outros tipos de conteúdo, a fim de avaliar o impacto nas inferências.

9 Agradecimentos

Os autores agradecem ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), à responsável pelo estabelecimento comercial que cedeu espaço para realização do experimento científico e aos colegas do Laboratório de Percepção Computacional (LPC) que ajudaram com discussões sobre a pesquisa.

Contribuição dos autores:

- Ítalo de Pontes Oliveira: Escreveu primeira versão do artigo, montagem dos equipamentos, coleta dos dados, análise dos resultados.

- Eanes Torres Pereira e Herman Martins Gomes: Orientação na realização do experimento e escrita do primeiro artigo, sugestão de hipóteses, colaboraram na escrita e correções do artigo nas demais versões.

Referências

- [1] T. W. Anderson and D. A. Darling. Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes. *Ann. Math. Statist.*, 23(2):193–212, 06 1952.
- [2] Cannes Lions Archive. Strip commerce. <http://www.canneslionsarchive.com/entries/528012/strip-commerce>. acessado em 3 de Setembro de 2015.
- [3] Joan Fisher Box. Guinness, gosset, fisher, and small samples. *Statist. Sci.*, 2(1):45–52, 02 1987.
- [4] Intel. Engagin bank customers with interactive digital signage. <http://www.intel.co.za/content/dam/www/public/us/en/documents/solution-briefs/engaging-bank-customers-with-interactive-digital-signage-brief.pdf>. acessado em Setembro de 2015.
- [5] Don-Lin Yang Kuo-Cheng Yin, Hsin-Chieh Wang and Jungpin Wu. A study on the effectiveness of digital signage advertisement. In *2012 International Symposium on Computer, Consumer and Control (IS3C)*, pages 169–172, June 2012.
- [6] Alan B. Forsythe Morton B. Brown. Robust tests for the equality of variances. *Journal of the American Statistical Association*, 69(346):364–367, 1974.
- [7] Karl Pearson. Note on regression and inheritance in the case of two parents. *Proceedings of the Royal Society of London*, 58(347-352):240–242, 1895.
- [8] POPAI. Screen-media formats. <http://www.popai.com/uploads/downloads/POPAIDigitalSignage-Screen-Media-Formats-2009.pdf>. acessado em Setembro de 2015.
- [9] Robert Ravnik and Franc Solina. Audience measurement of digital signage: Quantitative study in real-world environment using computer vision. *Interacting with Computers*, 25(3):218–228, 2013.
- [10] Robert Ravnik and Franc Solina. Audience adaptive digital signage using real-time computer. *International Journal of Advanced Robotic Systems*, 2013-02-07.
- [11] Robert Ravnik and Franc Solina. Interactive and audience-adaptive information interfaces. *ELCVIA Electronic Letters on Computer Vision and Image Analysis*, 13(2), 2014.

- [12] Sheeraz Akram Sajid Ali Khan, Muhammad Nazir and Naveed Riaz. Gender classification using image processing techniques: A survey. In *2011 IEEE 14th International Multitopic Conference (INMIC)*, pages 25–30, Dec 2011.
- [13] S. S. Shapiro and M. B. Wilk. An analysis of variance test for normality (complete samples). *Biometrika*, 52(3-4):591–611, 1965.
- [14] M. Sharkas and M. A. Elenien. Eigenfaces vs. fisherfaces vs. ica for face recognition; a comparative study. In *Signal Processing, 2008. ICSP 2008. 9th International Conference on*, pages 914–919, Oct 2008.
- [15] Portal Digital Signage. Aplicações do digital signage. <http://www.portaldigitalsignage.com.br/>. acessado em 3 de Setembro de 2015.
- [16] Marina Sokolova and Guy Lapalme. A systematic analysis of performance measures for classification tasks. *Inf. Process. Manage.*, 45(4):427–437, July 2009.
- [17] C Spearman. The proof and measurement of association between two things. *International Journal of Epidemiology*, 39(5):1137–1150, 2010.
- [18] Phil Tian, Addicam V. Sanjay, Kunapareddy Chiranjeevi, and Shahzad Malik Malik. Intelligent advertising framework for digital signage. In *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD'12*, pages 1532–1535, New York, NY, USA, 2012. ACM.
- [19] Paul Viola and Michael J. Jones. Robust real-time face detection. *International Journal of Computer Vision*, 57(2):137–154, 2004.
- [20] Roy Want and Bill N. Schilit. Interactive digital signage. *Computer*, 45(5):21–24, May 2012.