



GABRIEL CAMBRAIA SOARES

**SISTEMA EMBARCADO PARA O MONITORAMENTO DA
SONOLÊNCIA E ASSISTÊNCIA AO CONDUTOR**

LAVRAS – MG

2018

GABRIEL CAMBRAIA SOARES

**SISTEMA EMBARCADO PARA O MONITORAMENTO DA SONOLÊNCIA E
ASSISTÊNCIA AO CONDUTOR**

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Engenharia de Sistemas e Automação, para a obtenção do título de Mestre.

Arthur de Miranda Neto

Orientador

Danilo Alves de Lima

Coorientador

LAVRAS – MG

2018

**Ficha catalográfica elaborada pelo Sistema de Geração de Ficha Catalográfica da Biblioteca
Universitária da UFLA, com dados informados pelo próprio autor.**

Soares, Gabriel Cambraia

Sistema embarcado para o monitoramento da sonolência e
assistência ao condutor / Gabriel Cambraia Soares. - 2018.

72 p. : il.

Orientador: Arthur de Miranda Neto.

Coorientador: Danilo Alves de Lima.

Dissertação (mestrado acadêmico) – Universidade Federal
de Lavras, 2018.

Bibliografia.

1. Identificação de sonolência. 2. PERCLOS. 3. Filtro de
Kalman. I. de Miranda Neto, Arthur II. de Lima, Danilo Alves.
III. Título.

GABRIEL CAMBRAIA SOARES

**SISTEMA EMBARCADO PARA O MONITORAMENTO DA SONOLÊNCIA E
ASSISTÊNCIA AO CONDUTOR
EMBEDDED SYSTEM FOR DROWSINESS MONITORING AND DRIVER
ASSISTANCE**

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Engenharia de Sistemas e Automação, para a obtenção do título de Mestre.

APROVADA em 26 de Março de 2018.

Leonardo Silveira Paiva UFLA
Renato Ramos da Silva UFLA
Giovani Bernardes Vitor UNIFEI

Arthur de Miranda Neto
Orientador

Danilo Alves de Lima
Co-Orientador

**LAVRAS – MG
2018**

AGRADECIMENTOS

Agradeço à Deus por me dar forças, sabedoria e por me guiar nos caminhos corretos. À Universidade Federal de Lavras, especialmente ao Departamento de Engenharia e seus excelentes professores. Gratifico o Programa de Pós-Graduação em Engenharia de Sistemas e Automação, pela oportunidade, a Capes, pela concessão da bolsa de Mestrado, a Fapemig e o CNPq. Agradeço o professor Dr. Arthur de Miranda Neto, pela orientação e, principalmente, pela confiança na minha pesquisa, e o professor Dr. Danilo Alves de Lima, pelos ensinamentos imprescindíveis, pela enorme disponibilidade e pela paciência. Sou agradecido à minha família, pelo apoio incondicional e pela confiança em minha caminhada, à minha namorada Thaís, pelo amor, carinho, atenção, paciência, apoio em todos os momentos, pela presença constante e pela torcida infindável. Gratifico, também, os professores Dr. Leonardo Silveira Paiva, Dra. Sílvia Costa Ferreira e Dr. Bruno Henrique Groenner Barbosa, pelos conhecimentos repassados e por fazerem parte da minha trajetória acadêmica.

MUITO OBRIGADO!

RESUMO

Com o avanço da tecnologia e o elevado número de acidentes envolvendo veículos terrestres causados por adormecimento, sistemas assistenciais vêm sendo desenvolvidos para monitorar características de condutores que indicam sonolência. Paralelamente, a ascensão da tecnologia relacionada aos *smartphones* e a sua propagação nas diversas camadas socioeconômicas é uma boa oportunidade para a difusão dos sistemas assistenciais, uma vez que grande parte dos módulos de identificação de sonolência são embarcados em veículos modernos. Sendo assim, este trabalho propõe o desenvolvimento de uma aplicação em *smartphone* para o monitoramento da sonolência e do adormecimento do condutor, com base na análise do estado de seus olhos segundo o indicador PERCLOS. A identificação dos olhos é feita por classificadores do tipo *Local Binary Pattern*, com a vantagem de serem rápidos e de baixa complexidade computacional, requisito muito importante em aplicações móveis. Estes foram treinados com conjuntos de dados específicos e testados por quatro indivíduos, diversificados em idade, gênero e tonalidade dos olhos, em um ambiente com iluminação natural. A ocorrência de falsos positivos foi nula e as acurácias médias dos classificadores de olhos abertos e fechados foram de 84,16% e 73,23%, respectivamente. A identificação da sonolência foi avaliada em um experimento composto por condutores vigilantes, acordados há duas horas, e sonolentos, acordados há dezoito horas. O PERCLOS de pessoas sonolentas apresentou um valor médio 209,78% superior ao de pessoas vigilantes, e foi claramente notado o sintoma de lentidão da abertura dos olhos durante o piscar de pessoas sonolentas. O sistema de detecção de adormecimento proposto opera em intervalos curtos de tempo e emite alertas nos eventos em que os olhos fechados correspondem a pelo menos 90% do valor total de olhos encontrados. Foram alcançados bons resultados na identificação do adormecimento, uma vez que nas dez simulações praticadas a detecção foi cometida, com um tempo médio de 2,034 segundos. Para o futuro, a aplicação será aprimorada com uma melhor interface com o usuário e integração de outras informações, como a velocidade do veículo, para manter uma interação eficiente com o condutor.

Palavras-chave: Identificação de sonolência. Filtro de Kalman. Classificação dos estados dos olhos. Sistemas avançados de assistência ao motorista. PERCLOS

ABSTRACT

With the advancement of technology and the high number of accidents involving land vehicles caused by drowsiness, assistance systems have been developed to monitor characteristics of drivers that indicate drowsiness. At the same time, the rise of technology related to smartphones and their spread in the different socioeconomic layers is a good opportunity for the diffusion of assistance systems, since most of the drowsiness identification modules are shipped in modern vehicles. Thus, this work proposes the development of a smartphone application for the monitoring of driver's drowsiness and sleepiness, based on the analysis of the state of their eyes according to PERCLOS indicator. The identification of the eyes is done by Local Binary Pattern classifiers, with the advantage of being fast and of low computational complexity, a very important requirement in mobile applications. These were trained with specific data sets and tested by four individuals, diversified in age, gender, and tonality of the eyes, in an environment with natural light. The occurrence of false positives was null and the average accuracy of open and closed classifiers was 84.16% and 73.23%, respectively. The identification of drowsiness was evaluated in an experiment composed by vigilant drivers, who had been awake for two hours and drowsy, who had been awake for eighteen hours. The PERCLOS of drowsy people presented an average value of 209.78% higher than that of vigilantes, and the symptom of sluggishness of the eyes during the blink of drowsy people was clearly noticed. The proposed sleepiness detection system operates at short intervals and alerts on events where closed eyes correspond to at least 90% of the total eyes found. Good results were obtained in the identification of the sleepiness, since in the ten simulations practiced the detection was committed, with an average time of 2.034 seconds. In the future, the application will be improved with a better user interface and integration of other information, such as vehicle speed, to maintain an efficient interaction with the driver.

Keywords: Drowsiness detection. Kalman filter. Eye-state classification. Advanced Driver Assistance Systems. PERCLOS

LISTA DE FIGURAS

Figura 1.1 – Número de artigos publicados na área de Detecção de Sonolência.	14
Figura 2.1 – Pulseira inteligente Microsoft BAND 2.	19
Figura 2.2 – Representação dos posicionamentos dos canais de eletrodos verticais (quadrados vermelhos) e horizontais (quadrados azuis).	20
Figura 2.3 – Ilustração dos eletrodos ligados ao EEG, ao EOG e ao EMG e seus respectivos sinais.	21
Figura 2.4 – Exemplo de sinal que representa o piscar dos olhos (esquerda) e descrição visual dos parâmetros (direita).	22
Figura 2.5 – Características de dois (A e B), de três (C) e de quatro retângulos (D). . . .	24
Figura 2.6 – Representação gráfica do valor da imagem integral.	24
Figura 2.7 – Descrição da criação de uma unidade de textura LBP.	25
Figura 2.8 – Procedimento de classificação em cascata dos classificadores.	28
Figura 2.9 – Amostras extraídas para a formação do conjunto de dados usado para o treinamento do classificador.	29
Figura 2.10 – Pontos críticos usados para o aprimoramento do classificador.	30
Figura 3.1 – Fluxograma da aplicação de identificação de sonolência.	35
Figura 3.2 – Imagem em RGB (esquerda), escalas de cinza (centro) e com o histograma normalizado (direita).	36
Figura 3.3 – Composição de um retângulo na identificação da face.	37
Figura 3.4 – Representação dos pontos rastreados pelo Filtro de Kalman.	39
Figura 3.5 – Amostras de imagens positivas.	44
Figura 3.6 – Amostras de imagens negativas.	44
Figura 3.7 – Amostras de imagens positivas do primeiro conjunto de dados.	46
Figura 3.8 – Amostras de imagens positivas do segundo conjunto de dados.	46
Figura 3.9 – Amostra original (esquerda), com o <i>gamma</i> reduzido (centro) e com o <i>gamma</i> elevado (direita).	47
Figura 4.1 – Faces dentro (superior) e fora (inferior) do intervalo de 50 à 100 centímetros.	52
Figura 4.2 – Área da face detectada (amarelo) e área dos olhos segmentada (verde). . . .	53
Figura 4.3 – Rastreamento da área dos olhos em situações de oclusão.	54
Figura 4.4 – Identificação de olhos abertos e semi-abertos.	56

Figura 4.5 – Sequência de quadros que exibe a classificação de olhos fechados (centro) e de olhos com pequenas aberturas (esquerda e direita).	58
Figura 4.6 – Valores de PERCLOS do condutor sonolento (azul) e vigilante (vermelho). A área do paralelogramo rosa representa a separação entre os dados.	62
Figura 4.7 – Sequência de quadros extraída da simulação do condutor vigilante que descreve uma amostra de piscar de olhos.	62
Figura 4.8 – Sequência de quadros extraída da simulação do condutor sonolento que descreve uma amostra de piscar de olhos.	63

LISTA DE SIGLAS

ABS	<i>Anti-lock Braking System</i>
ADAS	<i>Advanced Driver Assistance Systems</i>
ECG	Eletrocardiograma
EEG	Eletroencefalograma
EMG	Eletromiografia
EOG	Eletrooculograma
EQM	Erro quadrático médio
FA	<i>Fall Asleep</i>
FN	<i>False Negative</i>
FP	<i>False Positive</i>
fps	<i>Frame per Second</i>
HOG	<i>Histogram of Oriented Gradients</i>
HRV	<i>Heart Rate Variability</i>
Hz	<i>Hertz</i>
LBP	<i>Local Binary Pattern</i>
LED	<i>Light-emitting Diode</i>
PERCLOS	<i>Percentage of Eyelid Closure</i>
PPG	Fotopletismografia
RGB	<i>Red, Green, Blue</i>
SIFT	<i>Scale-invariant Feature Transform</i>
TN	<i>True Negative</i>
TP	<i>True Positive</i>

LISTA DE TABELAS

Tabela 4.1 – Desempenho da aplicação na configuração tradicional e na aprimorada. . .	52
Tabela 4.2 – Desempenho da aplicação com os classificadores em operação na face e na área segmentada.	53
Tabela 4.3 – EQM das variáveis de rastreamento da área dos olhos.	55
Tabela 4.4 – Dados dos indivíduos participantes do experimento.	56
Tabela 4.5 – Dados da avaliação do classificador de olhos abertos.	57
Tabela 4.6 – Dados da avaliação do classificador de olhos fechados.	59
Tabela 4.7 – Valores de PERCLOS da simulação em que o motorista está vigilante. . . .	61
Tabela 4.8 – Valores de PERCLOS da simulação em que o motorista está sonolento. . .	61
Tabela 4.9 – Valores de tempo entre o início do adormecimento e a emissão de alerta. . .	64

SUMÁRIO

1	INTRODUÇÃO	11
1.1	Objetivos	12
1.2	Justificativas e Contribuições	13
1.3	Organização do Trabalho	15
2	REVISÃO DE LITERATURA	16
2.1	Sistemas Avançados de Assistência ao Condutor	16
2.2	Identificação de Sonolência	18
2.2.1	Métodos Invasivos	18
2.2.2	Métodos Não Invasivos	21
2.3	Descritores de Textura na Identificação de Objetos	23
2.3.1	Características <i>Haar</i>	23
2.3.2	<i>Local Binary Pattern</i> – LBP	25
2.4	Algoritmo <i>Adaboost</i> e Treinamento do Classificador	26
2.5	Filtro de Kalman no Rastreamento de Características do Motorista	29
2.6	Resumo do Capítulo	32
3	METODOLOGIA	34
3.1	Aquisição e Pré-Processamento de Imagens	36
3.2	Identificação da Face	36
3.3	Segmentação da Área dos Olhos	37
3.4	Filtro de Kalman no Rastreamento da Área dos Olhos	38
3.5	Identificação dos Olhos Abertos e Semi-Abertos	43
3.6	Identificação dos Olhos Fechados	45
3.7	Deteção de Sonolência e Alerta ao Motorista	47
3.7.1	Identificação de Sonolência Segundo a Estimação do PERCLOS	47
3.7.2	Identificação de Adormecimento	48
4	RESULTADOS E DISCUSSÕES	50
4.1	Implementação	50
4.2	Identificação da Face	51
4.3	Segmentação da Área dos Olhos	53
4.4	Rastreamento da Área dos Olhos	54
4.5	Deteção de Olhos Abertos	55

4.6	Detecção de Olhos Fechados	58
4.7	Identificação de Sonolência do Condutor	60
4.8	Identificação de Adormecimento do Condutor	63
5	CONCLUSÕES E TRABALHOS FUTUROS	66
	REFERÊNCIAS	68

1 INTRODUÇÃO

A ampliação do uso de veículos pela sociedade moderna desperta uma enorme preocupação sobre o tema da segurança dos condutores. Os acidentes que envolvem veículos são responsáveis por prejuízos diversos e, sobretudo, constituem uma das maiores fontes mundiais de tragédias. As principais causas de acidentes podem ser divididas em dois grupos. O primeiro consiste de causas evitáveis, como a falta de atenção, o uso de telefone durante a condução e a falta de experiência com a direção. Essas causas podem ser erradicadas com o aprimoramento do sistema de ensino e educação de trânsito. O segundo grupo é composto por fatores incontrolláveis, como condições fisiológicas, que são passíveis de monitoramento. Inclui-se, nesse grupo, a sonolência, que embora seja de difícil prevenção, é passível de fiscalização (AKBAR et al., 2017).

O sono é um requisito necessário do corpo humano, assim como comida e água. Quando este requisito é privado, o cérebro obriga o corpo a descansar para que os ciclos de sono sejam completos. Este comportamento natural provoca a diminuição da vigilância, isto é, a capacidade de concentração, e leva à inibição do controle sobre os músculos voluntários. Um motorista com os sintomas da sonolência é incapaz de reagir aos estímulos e, levemente, pode se envolver em um acidente. Essa condição fisiológica é afetada diretamente pelas atividades em que o indivíduo está submetido e, principalmente, em tarefas monótonas, como dirigir, a sonolência é amplificada (SILVEIRA; KOZAKEVICIUS; RODRIGUES, 2016). A performance da dirigibilidade começa a decrescer após duas horas em que o indivíduo começa a dirigir. A solução apontada para essa adversidade é orientar o motorista a dormir por, aproximadamente, 15 minutos (WANG; PEI, 2014). Ainda assim, não há formas legais para evitar a dirigibilidade de condutores neste estado devido a dificuldade de avaliação e referenciação da sonolência (NESHOV; MANOLOVA, 2017).

Tendo em vista o alto grau de periculosidade dos sinais da sonolência na atividade de condução veicular e a inexistência de uma legislação preventiva responsável pelo monitoramento de motoristas, pesquisas direcionadas para o desenvolvimento de ferramentas responsáveis pela estimação do sono são recorrentes. Os procedimentos publicados são separados em grupos segundo o índice de limitação, onde são classificados por métodos invasivos e métodos não-invasivos. O primeiro grupo é composto pelos procedimentos que usam sinais elétricos do corpo humano como dados de entrada para o sistema de estimação. Esses sinais são extraídos de exames de eletroencefalogramas, eletrocardiogramas ou eletrooculogramas em tempo

real, e apresentam um grau elevado de eficácia e acurácia (TABAL; CRUZ, 2017). No entanto, a aquisição dos sinais da atividade elétrica exige a aplicação de eletrodos em regiões do corpo humano, o que é uma imposição desconfortável diante da dinamicidade da dirigibilidade (WANG; XU, 2016).

Os métodos não-invasivos dispensam a aplicação de aparelhagem diretamente no corpo do motorista e possuem maior aceitação pela sociedade. Usualmente, são sistemas compostos por câmeras para a aquisição da imagem do ambiente interno do veículo em tempo real. As imagens são processadas com o objetivo de identificação de atividades praticadas por motoristas, como o bocejo (REDDY et al., 2017) e, especialmente, o piscar dos olhos (YAN et al., 2016), (GONG; KWAK, 2017). Por meio do monitoramento dos estados dos olhos é possível calcular a duração e a frequência das piscadas, como também os tempos de abertura e os tempos de fechamento. A eficácia destes indicadores foram avaliados em Caffier, Erdmann e Ullsperger (2003). A porcentagem de fechamento da pálpebra sobre a pupila ao longo do tempo (PERCLOS), apresentada por Wierwille et al. (1994), é considerada uma das medidas mais promissoras de vigilância em tempo real para os sistemas de detecção de sonolência (DINGES; GRACE, 1998).

O avanço recente na tecnologia dos *smartphones* vem abrindo um universo de aplicações para os métodos não-invasivos. Até então, as pesquisas eram conduzidas em processadores, câmeras e alto-falantes embarcados no veículo (CHENG et al., 2009) (BERGASA et al., 2006) (GRACE et al., 1998). No entanto, atualmente, as plataformas de programação das aplicações móveis estão sendo integradas com bibliotecas de visão computacional, como a *Open Source Computer Vision Library – OpenCV*, e as primeiras aplicações assistenciais aos motoristas estão sendo desenvolvidas neste ambiente.

1.1 Objetivos

O objetivo geral deste trabalho é desenvolver uma aplicação móvel de baixo custo, responsável pelo monitoramento do condutor e pela identificação de sinais de sonolência e de adormecimento baseados em análises dos estados dos olhos. Para tanto, os seguintes objetivos específicos foram estabelecidos:

1. Criar um sistema de identificação da face, dos olhos abertos e semi-abertos e dos olhos fechados;

2. Desenvolver um método de segmentação e rastreamento da área dos olhos com o Filtro de Kalman;
3. Desenvolver classificadores do tipo *Local Binary Pattern* para os olhos abertos e fechados;
4. Elaborar um sistema de identificação de sonolência e adormecimento segundo os conceitos do indicador PERCLOS.

1.2 Justificativas e Contribuições

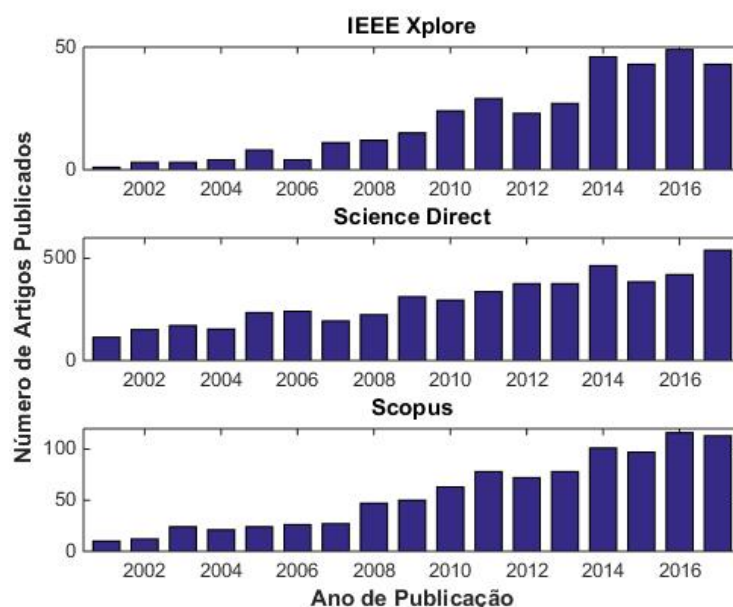
Os acidentes de transporte terrestre no Brasil matam cerca de 45 mil pessoas por ano, constituindo-se uma das principais causas de mortes no país. A perda de uma vida humana, os danos psíquicos e os estresses traumáticos aos quais as vítimas de trânsito e seus familiares estão sujeitos, é incalculável. Por outro lado, há também a formação de custos econômicos-financeiros que impactam diretamente as famílias, bem como a sociedade em geral. No ano de 2014, os acidentes ocorridos nas estradas federais ocasionaram um gasto de quase 13 bilhões de reais. Em média, cada acidente custou à sociedade brasileira aproximadamente 261 mil, sendo que um acidente envolvendo vítima fatal teve o custo médio de 664 mil reais (BRASIL, 2015).

Segundo a *National Sleep Foundation*, sediada nos Estados Unidos da América, cerca de 100.000 colisões por ano no país são causadas pelos sintomas da sonolência. Este número elevado também é justificado pela comparação de que estar acordado por 18 horas consecutivas proporciona o mesmo estado de fadiga física e mental de uma pessoa com concentração de álcool no sangue quase próxima a 0,08% (MAITA et al., 2017). No Brasil, a Polícia Rodoviária Federal é responsável pela divulgação dos dados anuais de trânsito. Em 2017, de janeiro à novembro, somente nas rodovias federais, 3.421 acidentes foram causados, especificamente, por sonolência de motoristas. Esses acidentes envolveram 6.913 pessoas, sendo que 2.568 tiveram ferimentos leves, 671 foram lesionadas gravemente e 330 pessoas morreram (BRASIL, 2017).

A preocupação com os prejuízos econômicos e sociais causados por estes tipos de acidentes é a justificativa para o desenvolvimento de ferramentas para a detecção de sonolência por parte de concessionárias, empresas e pesquisadores. Este fato pode ser observado na Figura 1.1, que mostra os resultados de uma pesquisa nas bases de dados *IEEE Xplore*, *Science Direct* e *Scopus*, realizada no dia 2 de janeiro de 2018, em que o termo *drowsiness detection* (iden-

tificação de sonolência) foi utilizado como o referencial das buscas. Em 16 anos, os artigos publicados na base *IEEE Xplore* sobre este assunto aumentaram em 4.300%. No mesmo período, as publicações nas bases *Science Direct* e *Scopus* foram ampliadas em 473,8% e 1.130%, respectivamente.

Figura 1.1 – Número de artigos publicados na área de Detecção de Sonolência.



Fonte: Do autor (2018).

A tendência é que o crescimento apresentado pela Figura 1.1 se mantenha nos próximos anos. Isso é justificado pelo surgimento de aplicações assistenciais ao motorista em plataformas móveis, como *smartphones* e *tablets*. Especificamente, os telefones da atualidade têm a mesma capacidade de processamento que os computadores de poucos anos atrás tinham, e uma penetrabilidade no mercado muito superior à dos computadores (ROMERA; BERGASA; ARROYO, 2015). No Brasil houve um crescimento de 25,4% nas vendas de celulares no primeiro trimestre de 2017, em comparação com o mesmo período do ano anterior (IDC, 2017). Paralelamente, os sistemas de processamento, armazenamento de energia e obtenção de imagem dos *smartphones* continuam sendo aprimorados constantemente. Além disso, a concorrência famigerada entre as fabricantes é fomento para a introdução de novas ferramentas, como diversos tipos de sensores, o que valoriza ainda mais o desenvolvimento de aplicações nestes dispositivos.

Neste sentido, o trabalho apresenta diferentes contribuições: i) uma abordagem inovadora referente ao desenvolvimento de um sistema identificador de sonolência para uma plataforma móvel. Esta contribuição está relacionada ao grau de mobilidade e acessibilidade que os *smartphones* proporcionam em comparação aos sistemas embarcados diretamente no veículo.

ii) um método de criação de classificadores para a identificação de olhos abertos e fechados segundo um conjunto de dados com imagens específicas. Com este método é possível desenvolver tipos de classificadores que não existem ou que não são disponibilizados para a sociedade acadêmica. Pode-se, além disso, aprimorar os classificadores existentes para que sejam compatíveis com o resultado exigido pelo projeto. iii) um sistema de segmentação e rastreamento da área dos olhos e gerenciamento de oclusão por meio do Filtro de Kalman.

1.3 Organização do Trabalho

Este documento está organizado da seguinte maneira: o Capítulo 2 é composto por conteúdos de pesquisas acadêmicas relacionadas aos temas tratados aqui. O Capítulo 3 discute a definição da metodologia usada na aplicação, onde são feitas descrições dos procedimentos que compõem o sistema de identificação de sonolência e adormecimento. No Capítulo 4 são exibidos os resultados quantitativos e qualitativos de cada procedimento proposto pela aplicação. Finalmente, no Capítulo 5, as conclusões gerais deste trabalho são apresentadas. Paralelamente, as perspectivas de continuidade são indicadas, exibindo os possíveis caminhos a serem seguidos como seguimento deste trabalho.

2 REVISÃO DE LITERATURA

2.1 Sistemas Avançados de Assistência ao Condutor

Os avanços na produção de automóveis em massa e o investimento na criação de rodovias pavimentadas refletem o alto grau de mobilidade individual obtido pelos países industrializados. Essa mobilidade está correlacionada com o bem-estar social e individual, e contribui significativamente para a qualidade de vida. Apesar de todas as vantagens que a mobilidade individual proporciona, o alto número de veículos implica no aumento do custo ecológico, econômico e social. A poluição sonora, a poluição dos gases, a perda de produtividade devido ao congestionamento e os acidentes de trânsito são exemplos clássicos (BENGLER et al., 2014).

Em países altamente industrializados, como na Alemanha, os mercados automobilísticos visam o aperfeiçoamento da mobilidade atual. Pensa-se em melhorar a eficiência dos automóveis em termos de energia, tempo e recursos. Dentre outras medidas de melhoria destaca-se a redução o número de acidentes de trânsito, que é tema do projeto *vision zero*, cuja finalidade é a erradicação de mortes e ferimentos graves causados por acidentes de trânsito. Uma abordagem diferente que contribui para esse aperfeiçoamento é o conceito de assistência ao motorista. Os sistemas assistenciais que estão sendo criados e aprimorados oferecem ferramentas para a melhoria da segurança ativa e integrada, porém, sua contribuição em larga escala depende de uma maior penetração no mercado (BENGLER et al., 2014).

A união entre a comunidade científica e a indústria automobilística tem contribuído para o desenvolvimento de sistemas de proteção para preservar a segurança dos condutores veiculares. Inicialmente, esses sistemas eram compostos por mecanismos simples, como o cinto de segurança, mas o avanço da tecnologia impulsionou a criação de dispositivos mais complexos, tais como freios ABS, programas de estabilização eletrônica e *air-bags*. Na última década, os dispositivos de proteção ao condutor evoluíram significativamente, sendo que as pesquisas foram direcionadas aos sistemas inteligentes que visam antecipar acidentes, a fim de evitá-los ou atenuar suas gravidades. Esses sistemas inteligentes são chamados de sistemas avançados de assistência ao condutor (*advanced driver assistance systems* – ADAS), visto que possuem a capacidade de auxiliar os condutores na tomada de decisões, no fornecimento de alertas em situações potencialmente perigosas e até na execução de ações autonomamente (GERONIMO et al., 2010).

Os fabricantes de automóveis estão explorando constantemente a vasta quantidade de sistemas inovadores com grande potencial de colaboração a dirigibilidade (UKE; THOOL; DHOTRE, 2012). Em particular, espera-se que as aplicações dos ADAS cresçam significativamente nos próximos anos, alcançando uma demanda de 100 milhões de unidades em 2018. Além disso, é previsto que cerca de 60% dos veículos no mundo possuirão ferramentas de conectividade à internet e ao *smartphone* até 2020, o que aumenta a possibilidade de aperfeiçoamento das aplicações existentes (VELEZ et al., 2015).

A proposta pela melhoria da segurança dos condutores vem resultando no desenvolvimento de ferramentas responsáveis pelo monitoramento das atividades do motorista em tempo real. Cita-se, como exemplo, o sistema de alerta de saída involuntária de faixa (*lane departure warning*), que tem a função de alertar o condutor em situações onde o veículo se deslocou entre as faixas sem o acionamento das setas indicadores de direção. Recentemente, Tan, Chen e Wang (2017) propuseram um modelo de identificação de faixa por meio do conceito de aprendizagem supervisionada da *Deep Fourier Neural Network*, enquanto Lee et al. (2017) usaram a transformada de Hough para a identificação da faixa e o Filtro de Kalman para o seu rastreamento.

Os sistemas assistenciais também abordam situações em que o motorista não é o principal envolvido. Tendo em vista a gravidade dos atropelamentos, as pesquisas sobre identificação de pedestres buscam alertar os motoristas para situações onde há possíveis colisões. Neste sentido, Hou et al. (2017) desenvolveram um método de identificação de pedestres por meio de Redes Neurais Convolucionais. O Filtro de Partículas foi a metodologia proposta por Gaddigoudar et al. (2017), onde o rastreamento da posição dos pedestres indica possíveis situações de colisão.

Os sistemas de monitoramento do condutor também estão inclusos nos ADAS, e são responsáveis pela medição de parâmetros que podem relatar alguma anormalidade na atenção do usuário em relação ao tráfego. Os trabalhos relacionados com a detecção de cansaço e sonolência são os mais comuns e diversificados na literatura. Lin et al. (2012) propuseram um sistema para esse tipo de detecção a partir de sinais de eletroencefalogramas, enquanto Jung, Shin e Chung (2014) utilizaram sinais de eletrocardiogramas. Takei e Furukawa (2005) estimaram o nível de cansaço de motoristas por meio da variação do ângulo de direção da roda do veículo. A visão computacional foi tema dos trabalhos de Dasgupta et al. (2013a), Yan et al. (2016), Acioğlu e Erçelebi (2016), onde os estados dos olhos foram extraídos para a estimação da sonolência segundo o indicador PERCLOS.

2.2 Identificação de Sonolência

2.2.1 Métodos Invasivos

Avanços em engenharia de materiais, circuitos eletrônicos, sensores, processamento de sinal e técnicas de classificação permitiram que sistemas computacionais interpretem quantidades biológicas, reconhecendo condições fisiológicas (SILVEIRA; KOZAKEVICIUS; RODRIGUES, 2016). Os sinais biomédicos são extremamente úteis para coletar informações da resposta do corpo durante o ciclo da sonolência. Esses sinais podem ser coletados por eletrocardiogramas (ECG) (GUPTA et al., 2017), eletrooculogramas (EOG) (ZHU et al., 2014) e por eletroencefalogramas (EEG), sendo que o último é o mais utilizado para a medição da atividade elétrica do cérebro (CORREA; OROSCO; LACIAR, 2014).

Correa, Orosco e Laciari (2014) propuseram um sistema de detecção de sonolência com base no teste de comparação de média das bandas relativas dos sinais de um canal de eletroencefalograma. O uso de um canal, em vez de dois ou quatro, diminui o grau de invasibilidade que os métodos tradicionais de EEG são caracterizados. Liang et al. (2017) apresentaram uma metodologia de estimação da sonolência que usa a fusão dos sinais elétricos provenientes do corpo humano com os sinais comportamentais que o motorista apresenta enquanto está dirigindo. Os sinais fisiológicos, provenientes de um eletroencefalograma e de um eletrooculograma, são usados para a criação de um modelo, no qual a sonolência é definida por pequenos episódios de adormecimento, dentro de um intervalo de um minuto. Os sinais comportamentais são originados pelo veículo onde o motorista está localizado, e avaliam a degradação de desempenho do condutor com base na identificação dos eventos em que o veículo cruza a faixa contínua involuntariamente.

A variabilidade do batimento cardíaco (*heart rate variability* – HRV) é o principal indicador gerado pelos sinais de eletrocardiograma. Ele é calculado por meio da análise da série entre os intervalos de batimento cardíaco. A análise da HRV é uma ferramenta confiável para a estimação da função do sistema nervoso autônomo. Esse sistema afeta, diretamente, as funções do corpo humano, e no contexto dos estudos conduzidos, sua influência sobre o ritmo diário deste. O ritmo diário é ligado as atividades que ocorrem, teoricamente, em um intervalo de 24 horas, como a sonolência e o adormecimento. O sistema nervoso autônomo é dividido em duas partes: sistema simpático e sistema parassimpático. O sistema nervoso simpático é responsável pela mobilização e excitação do corpo, uma vez que seu período de atuação ocorre durante o

dia. Por outro lado, o sistema nervoso parassimpático é responsável pelo acalmamento e pelo descanso, e atua predominantemente durante o sono. A relação entre as duas partes do sistema nervoso autônomo é visível nas mudanças diárias de frequência cardíaca e, consequentemente, nos resultados da análise da HRV. Durante períodos de fadiga ou sonolência, a dominação do sistema nervoso parassimpático aumenta, provocando um aumento de potência na banda de alta frequência (HF) e uma diminuição do poder na banda de baixa frequência (LF). As bandas são as características que podem ser analisadas para a estimativa da sonolência (PIOTROWSKI; SZYPULSKA, 2017).

Usualmente, os sinais gerados pelo eletrocardiograma são adquiridos por meio da aplicação de eletrodos na pele do indivíduo. Essa metodologia é considerada altamente invasiva e inconveniente nas situações de dirigibilidade que a vida real exige. Por esse motivo, Lee, Kim e Shin (2017) desenvolveram um sistema para adquirir os sinais de HRV por meio de fotopleismografia (PPG). Esse método baseia-se nos princípios da ótica para incidir uma luz no pulso do indivíduo, e compreender o movimento do coração, considerando a diferença da taxa de absorção da luz nos vasos sanguíneos. Para tanto, aplica-se o sistema em uma pulseira inteligente, como a mostrada na Figura 2.1, que possui um sensor PPG embutido e facilmente usável. Sendo assim, esse procedimento é muito conveniente na medição dos sinais de batimento cardíaco no cotidiano, o que o torna muito mais fácil e simples de se aplicar na vida real.

Figura 2.1 – Pulseira inteligente Microsoft BAND 2.



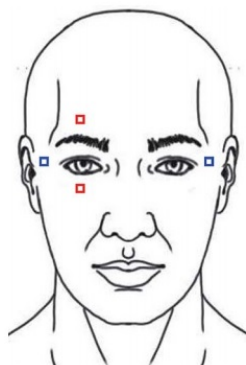
Fonte: (LEE; KIM; SHIN, 2017).

A medição da atividade elétrica dos olhos é nomeada de eletrooculograma (EOG). Especificamente, essa medição é a diferença de potencial entre a parte frontal e a parte traseira

dos olhos. Habitualmente, essa medição é realizada por meio de eletrodos, que são aplicados na pele do indivíduo. Os sinais obtidos permitem que sejam identificadas diversas atividades, como a movimentação e o piscar dos olhos. As atividades são usadas para a extração de características, como a duração do piscar, o tempo de fechamento, o tempo de abertura, a velocidade de fechamento da pálpebra, entre outras. O EOG é a técnica mais confiável para o estudo das características das piscadas dos olhos. Embora seja uma ferramenta invasiva, ele é altamente amostrado, e trabalha na faixa de 250 – 500 *Hz*. Em sistemas que operam com 250 *Hz*, por exemplo, cada amostra é adquirida em um intervalo de 0,004 segundos. Esse intervalo é muito inferior ao de câmeras que operam em 30 *fps* e fornecem *frames* a cada 0,033 segundos. A diferença no período de aquisição dos dados interfere diretamente na análise das piscadas de pessoas alertas, cuja duração pode chegar a 150 milissegundos. Cada piscada com este período seria identificada por apenas 4 *frames*, enquanto o eletrooculograma forneceria 37 amostras (PICOT; CAPLIER; CHARBONNIER, 2009).

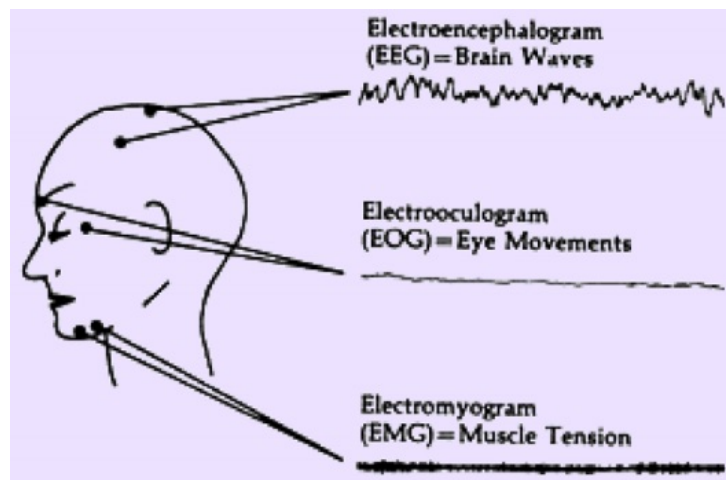
As metodologias para a estimação de sonolência por meio do EOG são diversas. Zhu et al. (2014) incorporaram os sinais do eletrooculograma em Redes Neurais Convolucionais para a criação de um modelo de treinamento não-supervisionado das características. Os sinais são extraídos de dois canais de eletrodos posicionados horizontalmente e verticalmente na face do indivíduo (FIGURA 2.2). A fusão dos sinais de eletroencefalograma, eletrooculograma e eletromiografia (EMG) foi a proposta apresentada em Kurt et al. (2009) para aprimorar a taxa de precisão da identificação de sonolência. Para tal, os sinais de interesse são gerados pelo cérebro, pelo movimento dos olhos e pelo movimento dos músculos do queixo, que são extraídos por eletrodos ligados ao EEG, ao EOG e ao EMG, respectivamente (FIGURA 2.3).

Figura 2.2 – Representação dos posicionamentos dos canais de eletrodos verticais (quadrados vermelhos) e horizontais (quadrados azuis).



Fonte: (ZHU et al., 2014).

Figura 2.3 – Ilustração dos eletrodos ligados ao EEG, ao EOG e ao EMG e seus respectivos sinais.



Fonte: (KURT et al., 2009).

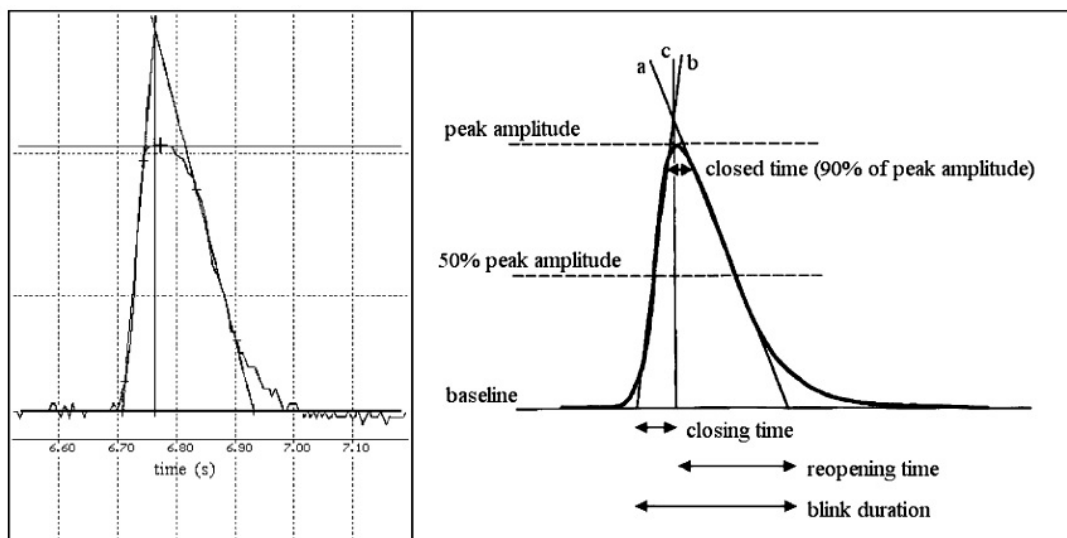
2.2.2 Métodos Não Invasivos

Os métodos citados na Subseção 2.2.1 requerem a aplicação de instrumentos para a aquisição dos sinais que são usados para a estimação da sonolência. Esses métodos apresentam índices de precisão elevados, mas são limitados ao ambiente do laboratório de pesquisa, uma vez que são invasivos e desconfortáveis (WANG; XU, 2016). Por esse motivo, estão sendo desenvolvidos métodos não-invasivos que não exigem o contato de eletrodos, como quaisquer outros aparelhos. Esses métodos são baseados na observação visual do motorista por meio de ferramentas de visão computacional (VITABILE; PAOLA; SORBELLO, 2011). Esta subseção abordará, especificamente, métodos baseados na análise do piscar dos olhos.

Caffier, Erdmann e Ullsperger (2003) descrevem que o piscar de olhos espontâneo é considerado um indicador ocular adequado para o diagnóstico da sonolência. Para eles, esse indicador oferece várias vantagens, uma vez que representa um fenômeno natural, facilmente observável e simplesmente acessível, que reflete a influência da ativação do nervo central sem manipulação voluntária. As atividades de abertura e fechamento dos olhos em condições normais são rápidas e ocorrem espontaneamente, ou seja, não são necessários estímulos externos. Segundo eles, parâmetros como a frequência e duração do piscar dos olhos são sujeitos a modificações características com o aumento da sonolência. Esses parâmetros são apresentados na Figura 2.4.

De acordo com Caffier, Erdmann e Ullsperger (2003), a duração de cada piscada é descrita pelo intervalo de tempo em que os olhos iniciam o fechamento até o momento em que eles retornam ao estado de abertura. O tempo de fechamento é o período em que os olhos saem do

Figura 2.4 – Exemplo de sinal que representa o piscar dos olhos (esquerda) e descrição visual dos parâmetros (direita).



Fonte: (CAFFIER; ERDMANN; ULLSPERGER, 2003).

estado de abertura para o estado de fechamento. De forma semelhante, o tempo de abertura é o período em que os olhos levam para voltar ao estado de abertura. Com relação ao tempo em que os olhos se encontram fechados, os resultados apontam que a duração do piscar e o tempo em os olhos se encontram fechados são os parâmetros que apresentam padrões diferentes nos estados ativo e sonolento. Pessoas sonolentas tendem a ter piscadas com durações mais longas e propendem a demorar a abrir os olhos durante as piscadas.

Um método baseado no movimento dos olhos foi desenvolvido por Dinges e Grace (1998). Nomeado de PERCLOS, cuja composição na língua inglesa é *Percentage of Eyelid Closure*, ele descreve a porcentagem de fechamento da pálpebra sobre a pupila ao longo do tempo. Os autores mostraram três métricas distintas:

- P70: proporção de tempo em que os olhos estão fechados por pelo menos 70%.
- P80: proporção de tempo em que os olhos estão fechados por pelo menos 80%.
- *EYEMEAS*: a porcentagem média quadrada da classificação do fechamento das pálpebras.

A métrica P80 do método PERCLOS foi utilizada em Yan et al. (2016) para estimar a sonolência em tempo real. Nele, o sistema de classificação dos olhos apresenta duas saídas:

- Olhos abertos: engloba os estados de olhos abertos e olhos semi-abertos que estejam até 80% fechados.

- Olhos fechados: compreende os estados de olhos fechados e olhos semi-abertos que estejam acima de 80% de fechamento.

A estimação da sonolência projetada por Yan et al. (2016) é descrita pela equação 2.1, onde TP é a proporção de tempo referente ao P80, NC é o número de imagens no estado fechado e NO é o número de imagens no estado aberto. Conforme a proporção de tempo se torna maior, o sistema reconhece isso como uma elevação na classificação de olhos fechados, o que representa que o motorista está se tornando sonolento. Uma função limiar é suficiente para a separação do estado de alerta com o estado de cansaço.

$$TP = \frac{NC}{NC + NO}. \quad (2.1)$$

2.3 Descritores de Textura na Identificação de Objetos

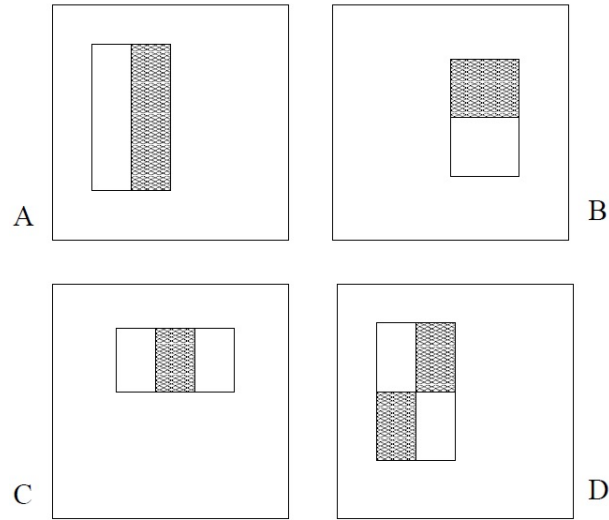
2.3.1 Características *Haar*

Viola e Jones (2004) desenvolveram um algoritmo para a construção de uma estrutura de identificação de objetos rápida e robusta. O procedimento de detecção classifica as imagens com base no valor de características, em vez de analisar diretamente cada *pixel*, uma vez que as características podem atuar na codificação do conhecimento, que é difícil de se aprender, usando uma quantidade finita de dados de treinamento. Ademais, o sistema baseado em características funciona mais rápido do que um sistema baseado em *pixels*. A Figura 2.5 exibe exemplos de características que são usadas na identificação de objetos. As características de dois retângulos são exibidas em (A) e (B), a de três retângulos é mostrada em (C) e a de quatro retângulos é apresentada em (D). Uma característica específica é definida pela diferença da soma dos *pixels* das áreas dentro dos retângulos, que pode estar em qualquer posição e escala na imagem original. Especificamente, o valor da característica *Haar-like* é a diferença entre as somas dos *pixels* do retângulo escuro pela soma dos *pixels* do retângulo claro (equação 2.2).

$$f(x)_{Haar} = Soma_{retngulo\ escuro}(pixels) - Soma_{retngulo\ claro}(pixels). \quad (2.2)$$

O cálculo das características dos retângulos pode ser rapidamente realizado por meio da representação da imagem integral, onde no local x e y contém a soma dos *pixels* acima e a esquerda desse local (FIGURA 2.6). Essa informação é expressa matematicamente pela equação 2.3, onde $ii(x,y)$ é o valor da imagem integral e $i(x,y)$ é o valor da imagem original.

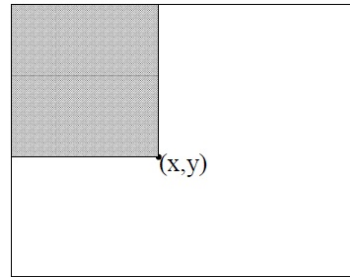
Figura 2.5 – Características de dois (A e B), de três (C) e de quatro retângulos (D).



Fonte: (VIOLA; JONES, 2004).

$$ii(x,y) = \sum_{x' \leq x, y' \leq y} i(x', y'). \quad (2.3)$$

Figura 2.6 – Representação gráfica do valor da imagem integral.



Fonte: (VIOLA; JONES, 2004).

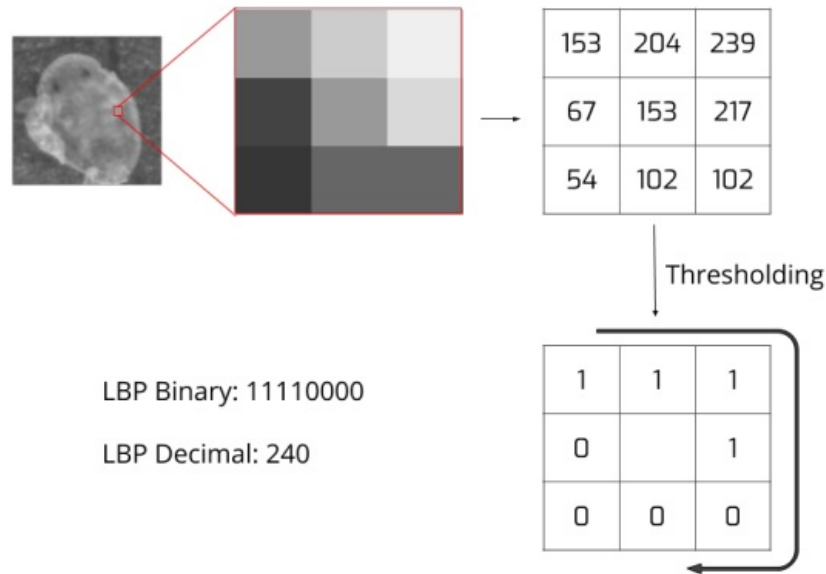
O processo de treinamento e seleção das melhores características é realizado por meio de um conjunto de dados de imagens positivas e imagens negativas, sendo que o treinamento é efetuado pelo algoritmo de aprendizado *AdaBoost* (ver Seção 2.4) (VIOLA; JONES, 2004). Esse algoritmo é projetado para selecionar a característica retangular que melhor separa os exemplos positivos dos negativos. Para cada característica, o sistema de aprendizado determina a função de classificação do limiar, $h_j(x)$, de modo que o número mínimo de exemplos sejam classificados de forma incorreta (equação 2.4). Sendo assim, a função do classificador $h_j(x)$ consiste de uma característica f_j , de um limiar θ_j e de uma paridade p_j , que indica a direção do sinal de desigualdade da equação 2.4.

$$h_j(x) = \begin{cases} 1, & \text{se } p_j f_j(x) < p_j \theta_j, \\ 0, & \text{senão.} \end{cases} \quad (2.4)$$

2.3.2 Local Binary Pattern – LBP

O padrão binário local (*local binary pattern* – LBP), proposto por Ojala, Pietikäinen e Harwood (1996), é um operador que atua em uma vizinhança com área de 3×3 *pixels*. Esse descritor, representado pela Figura 2.7, calcula a relação binária entre cada *pixel* da imagem, e seus pontos de vizinhança local em escalas de cinza. O valor do *pixel* no ponto central, v_c , é comparado com os valores dos *pixels* vizinhos, v_{ij} , na vizinhança ij (equação 2.5), na qual é formada a unidade de textura binária LBP $b_{i,j}$. Dessa forma, cada *pixel* possui uma unidade de textura com um valor específico (LINO et al., 2017).

Figura 2.7 – Descrição da criação de uma unidade de textura LBP.



Fonte: (LINO et al., 2017).

$$b_{ij} = \begin{cases} 1, & \text{se } v_{ij} \leq v_c, \\ 0, & \text{senão } > v_{ij}. \end{cases} \quad (2.5)$$

Os descritores do tipo LBP provaram ser muito eficazes para a representação de imagens que foram aplicadas a inspeção visual, detecção de movimento e análise de expressões. As propriedades mais importantes desses descritores são a sua tolerância contra mudanças de iluminação e, especialmente, sua simplicidade computacional, motivo ao qual foi definida a sua

utilização neste trabalho. Uma vez que os descritores LBP são compostos de números inteiros, tanto o treinamento quanto a detecção são substancialmente mais rápidos do que nos procedimentos que usam características *Haar* (LINO et al., 2017). Além disso, eles também identificam diversas formas de texturas primitivas, como pontos, linhas, bordas e cantos (ZHANG; HUA, 2015).

Dasgupta et al. (2013a) usaram os descritores LBP para a identificação dos olhos no sistema de estimação de sonolência. O processo de classificação é separado em fases de treinamento, apresentada na Seção 2.4, e detecção. A fase de treinamento se inicia com o redimensionamento das imagens dos olhos para a resolução 50x40 *pixels*, onde cada imagem é dividida em blocos de 5x4 *pixels*. Os valores de cada bloco de características são encontrados seus histogramas são calculados. Os histogramas de cada bloco são juntados para formar um classificador global de características. Finalmente, os vetores de características são avaliados e os 40 com melhores desempenhos são selecionados. A fase de identificação dos olhos se inicia com a obtenção da região de interesse que contém a face humana. Para cada sub-janela da região de interesse, o histograma do bloco LBP é encontrado e projetado. Posteriormente, a sub-janela com o menor erro de reconstrução é definida. Se o erro de reconstrução for menor que o limiar estipulado, a sub-janela é considerada uma detecção positiva.

2.4 Algoritmo *Adaboost* e Treinamento do Classificador

O algoritmo *Adaboost* (VIOLA; JONES, 2004) baseia-se nas características *Haar* (Subseção 2.3.1) para o treinamento de um classificador de face com alta precisão de detecção e velocidade de identificação superior a de quase todos algoritmos disponíveis. Ele é um algoritmo de aprendizagem de máquina que cria um classificador mais forte por meio da saída de algoritmos de classificação ponderada, que são conhecidos como aprendizes fracos. Um aprendiz fraco é definido de maneira que ele somente pode ser melhor do que o aleatório, ou, em outras palavras, se encontra mais próximo da orientação correta. Por outro lado, aprendizes fortes referem-se a um conjunto de algoritmos de classificação altamente correlacionados com a orientação correta. Sendo assim, o algoritmo permite a mudança de aprendizes fracos para minimizar erros em sujeitos não classificados. Desta forma, muitas das imagens que não pertencem aos rostos humanos são excluídas e mais cálculos são feitos nos quadros que têm mais chances de conter a face (CHEHREHGOSHA; EMADI, 2016).

O algoritmo de classificação seleciona as amostras de treinamento da seguinte maneira: $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$, onde x_i é a i -ésima amostra de treinamento, enquanto $y_i \in \{0, 1\}$ representa as amostras positivas e negativas, respectivamente. As amostras positivas são usadas para treinar o classificador para identificar características semelhantes, enquanto as amostras negativas são usadas para eliminar características que representam diferenças (KONG et al., 2015).

Considerando que o número de amostras positivas é s , e o número de amostras negativas é m , tem-se que o número de total de amostras $n = m + s$. O primeiro passo do algoritmo *Adaboost* é a inicialização dos pesos de cada amostra: $\omega_{1,i} = 1/n$, $i = 1, 2, \dots, n$. Em seguida, os próximos passos são repetidos para $t = 1, 2, \dots, T$, onde T é o número ótimo de classificadores fracos (KONG et al., 2015):

1. Normalize o peso em uma distribuição de probabilidade:

$$q_{t,i} = \frac{\omega_{t,i}}{\sum_{j=1}^n \omega_{t,i}}; \quad (2.6)$$

2. Para cada característica f , treine um classificador fraco h_f , e calcule as taxas de erro ponderadas ε_f dos classificadores fracos correspondentes a todas as características:

$$\varepsilon_f = \sum q_{t,i} |h_t(x_i) - y_i|; \quad (2.7)$$

3. Escolha o melhor classificador fraco $h_t(x)$ do passo (2), que apresenta a menor taxa de erro ε_t ;
4. Segundo o melhor classificador fraco, ajuste o peso:

$$\omega_{t+1,i} = \omega_{t,i} \beta_i^{1-e_i}, \quad (2.8)$$

onde $e_i = 0$ representa a classificação correta, e $e_i = 1$ representa a classificação incorreta;

5. Forme o classificador forte:

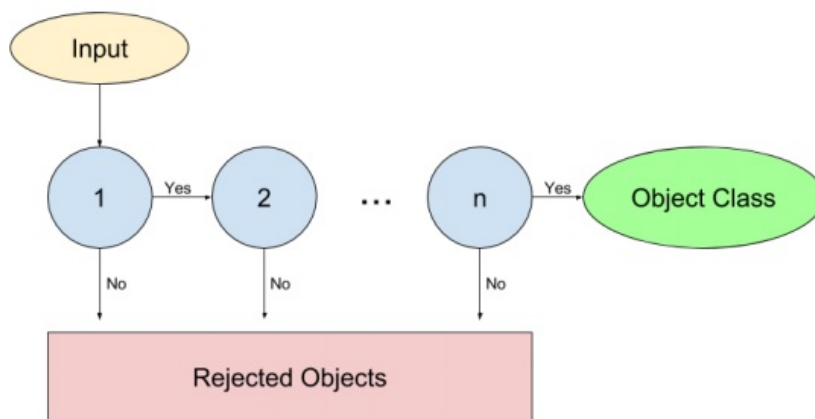
$$h(x) = \begin{cases} 1, & \sum_{t=1}^T \alpha_t h_t(x) \geq \frac{1}{2} \sum_{t=1}^T \alpha_t, \\ 0, & \text{senão,} \end{cases} \quad (2.9)$$

onde $\alpha_t = \log \frac{1-\varepsilon_t}{\varepsilon_t}$.

O algoritmo *Gentle Adaboost*, desenvolvido por Friedman et al. (2000), é um aprimoramento do *Adaboost* original. Enquanto o original calcula um classificador fraco em cada iteração e os combina de uma maneira linear para criar um classificador forte, o *Gentle Adaboost* calcula o classificador fraco pela otimização do erro de mínimos quadrados ponderado $\varepsilon_f = \sum \omega_{t,i} (y_i - f_t(x_i))^2$, onde $f_t = \alpha_t h_t(x)$. Dessa forma, o algoritmo *Gentle Adaboost* aumenta os valores dos pesos em classificações errôneas exponencialmente. Sendo assim, ele pode facilmente se concentrar nas instâncias difíceis de se classificar e tentar classificá-las corretamente (WU; NAGAHASHI, 2015).

A classificação em cascata é um processo de otimização dos algoritmos baseados no *Adaboost*. O objetivo das cascatas é descartar as regiões que não contém o objeto pesquisado (FIGURA 2.8) nos estágios iniciais. Cada primeira característica detectada em uma determinada área é delimitada com uma probabilidade de se ter um objeto nessa área, e o mesmo ocorre nos próximos estágios. Sendo assim, o treinamento sucessivo acontece apenas em regiões com alta probabilidade de ter o objeto, o que reduz drasticamente os custos computacionais de operação.

Figura 2.8 – Procedimento de classificação em cascata dos classificadores.



Fonte: (LINO et al., 2017).

O treinamento do classificador em cascata segundo o algoritmo *Adaboost* ou *Gentle Adaboost* pode ser feito por meio de características *Haar*, padrões *LBP* ou características *HOG*. Para a detecção de cansaço, em (KONG et al., 2015) foi treinado um classificador por meio de características *Haar* para a identificação de faces. No total, foram usadas 3000 imagens de faces frontais e 1500 amostras de faces com desvio de posição, que foram normalizadas e redimensionadas em 20×20 pixels. Foram treinados, também, classificadores para identificar olhos abertos e olhos fechados. Para tanto, criaram-se conjuntos de dados de imagens positivas de olhos abertos e fechados, visto que não existe um conjunto público de olhos humanos disponível para as amostras positivas e negativas (FIGURA 2.9). Foram extraídas 1190 amostras de olhos abertos e 700 amostras de olhos fechados, onde a resolução foi normalizada para 15×7 pixels.

Figura 2.9 – Amostras extraídas para a formação do conjunto de dados usado para o treinamento do classificador.



Fonte: (KONG et al., 2015).

Classificadores de faces baseados nos descritores do tipo LBP foram usados em Chehreghosha e Emadi (2016). Para tal, foram usadas somente 72 imagens para o treinamento do classificador, em vez de milhares. A análise de três grupos de pontos críticos (FIGURA 2.10), baseada na posição de elementos da face, aperfeiçoa o processo de classificação. Esse tipo de descritor apresenta a melhor performance entre os demais, visto que é possível identificar faces com expressões diversificadas, e em ambientes com situações adversas (CHEHREHGOSHA; EMADI, 2016).

2.5 Filtro de Kalman no Rastreamento de Características do Motorista

O Filtro de Kalman é um conjunto de equações matemáticas pelas quais o estado de um sistema linear é estimado de forma eficiente, por meio de uma solução recursiva do método dos mínimos quadrados. Essas equações são suficientes para a implementação de um estimador

Figura 2.10 – Pontos críticos usados para o aprimoramento do classificador.



Fonte: (CHEHREHGOSHA; EMADI, 2016).

linear do estado de um processo, com base nas informações disponíveis no momento anterior, pela correção deste estado segundo as informações disponíveis no momento atual (GARCÍA, 2009).

A qualidade da estimação é justificada pela composição do filtro, uma vez que são usadas medições atuais, estados anteriores e distribuições probabilísticas para estabelecer a próxima conjectura. A variável de estado, X_t , e a variável de medição, Z_t , formam a base da teoria do Filtro de Kalman, que é dividido em duas fases: predição e estimação (PRATIARSO et al., 2017).

A fase de predição é iniciada com a modelagem da variável de estado, que é representada pela equação 2.10 (PRATIARSO et al., 2017). Ela define o modo como o estado x varia, ao qual está relacionado no tempo com o estado anterior X_{t-1} (GARCÍA, 2009). w_{t-1} é o parâmetro de ruído do sistema, que é Gaussiano e de média zero, modelado como $p(w) = N(0, Q)$, onde a probabilidade de w segue a distribuição normal, com média 0, e covariância do ruído do processo Q (PRATIARSO et al., 2017). A matriz Q mede a variabilidade do sinal de entrada das transições ideais, definidas pela matriz de transição (DASGUPTA et al., 2013b). A matriz de transição é representada por A e contém a modelagem da transição do sistema. A matriz B é relacionada à entrada de controle u no tempo $t - 1$.

$$\hat{X}_{t|t-1} = A.\hat{X}_{t-1|t-1} + Bu_{t-1} + w_{t-1}. \quad (2.10)$$

A segunda etapa da fase de predição é a previsão da covariância do erro (equação 2.11), parâmetro utilizado no cálculo do ganho de Kalman (equação 2.12), como segue:

$$P_{t|t-1} = A.P_{t-1|t-1}.A^T + Q. \quad (2.11)$$

$$K_t = P_{t|t-1}H_t^T (H_tP_{t|t-1}H_t^T + R)^{-1}. \quad (2.12)$$

A computação do ganho de Kalman inicia a fase de estimação do estado. Nessa fase, a estimativa anterior do estado é corrigida segundo os dados das medições. A matriz H representa o estado da medição no tempo t e a matriz R define o erro na medição (PRATIARSO et al., 2017). Em seguida, o estado subsequente é estimado usando o ganho de Kalman e o resíduo proveniente da subtração entre os valores previstos e os valores medidos (equação 2.13). Por fim, a nova matriz de covariância do erro é calculada pela equação 2.14.

$$\hat{X}_{t|t} = \hat{X}_{t|t-1} + K_t(Z_t - H_t\hat{X}_{t|t-1}). \quad (2.13)$$

$$P_{t|t} = (I - K_tH_t)P_{t|t-1}. \quad (2.14)$$

A alta precisão do Filtro de Kalman é alento para a sua utilização em sistemas de rastreamento de face. No contexto dos motoristas automotivos, a posição do rosto não se altera de forma abrupta em quadros consecutivos. Sendo assim, a região da face pode ser estimada, em tempo real, sendo que esta atividade é nomeada de rastreamento. Uma vez que a posição e o tamanho do rosto são conhecidos, a partir dos resultados de detecção adquiridos por um classificador, é possível implementar um sistema que estima recursivamente essas variáveis (DASGUPTA et al., 2013b).

Uma aplicação do Filtro de Kalman no rastreamento da face pode ser vista em Dasgupta et al. (2013b), no qual o vetor de medições $z_k = [x_{1,k}, y_{1,k}, x_{2,k}, y_{2,k}]^T$ é composto pelas coordenadas do canto inferior esquerdo $(x_{1,k}, y_{1,k})$ e pelas do canto superior direito $(x_{2,k}, y_{2,k})$, que representam o limite da face. O vetor de estados foi definido como $x_k = [x_{1,k}, y_{1,k}, x_{2,k}, y_{2,k}, u_{x,k}, u_{y,k}]^T$, onde $u_{x,k}$ e $u_{y,k}$ são as velocidades do centro do retângulo na direção horizontal e vertical, respectivamente, e calculadas pelas equações 2.15 e 2.16.

$$u_{x,k} = \frac{x_{1,k} + x_{2,k}}{2} - \frac{x_{1,k-1} + x_{2,k-1}}{2}. \quad (2.15)$$

$$u_{y,k} = \frac{y_{1,k} + y_{2,k}}{2} - \frac{y_{1,k-1} + y_{2,k-1}}{2}. \quad (2.16)$$

No rastreamento dos olhos, por sua vez, opta-se por estimar a posição das coordenadas x e y da pupila dos olhos (ZHANG; ZHANG, 2010) e (GU; JI; ZHU, 2002). Em Zhang e Zhang (2010), a identificação de sonolência, por meio do rastreamento dos olhos, é processada segundo a estimação do indicador PERCLOS, nos quais os estados dos olhos são classificados em uma janela de tempo. Em Gu, Ji e Zhu (2002), um sistema de estimação de sonolência foi elaborado com base em 22 características, onde em uma delas é utilizado o rastreamento da pupila para descrever a movimentação da face. Em ambos, a matriz de estados é representada por $X_t = (c_t, r_t, u_t, v_t)$, onde c_t, r_t são as posições da pupila e u_t, v_t são as velocidades da pupila em x e y respectivamente. A correção da estimação é feita com as medições z_t , na qual $z_t = (c_t, r_t)$. A semelhança na composição do Filtro de Kalman também é verificada na taxa de precisão do rastreamento da pupila. Para Gu, Ji e Zhu (2002), o procedimento de rastreamento obteve 97.6% de acerto, e em Zhang e Zhang (2010) a precisão excedeu 99%.

2.6 Resumo do Capítulo

Este capítulo exibiu os principais pontos que devem ser compreendidos para o entendimento das ferramentas usadas neste trabalho. A identificação de sonolência em condutores, principal objetivo desta pesquisa, é uma tarefa que pode ser estimada por meio de diversos sinais. Os sinais biológicos são os mais confiáveis, e apresentam índices de precisão elevados, mas são excessivamente invasivos. Por outro lado, a qualidade de estimação da sonolência segundo os sinais visuais dependem do equipamento de aquisição e processamento da imagem e da iluminação do ambiente. Em meio as limitações dos dois recursos, a métrica usada em Yan et al. (2016), baseada na classificação do estado dos olhos para a estimação do PERCLOS, foi usada neste trabalho para a identificação da sonolência.

O sistema de classificação do estado dos olhos deve ser preciso e, principalmente, leve, uma vez que a aplicação é destinada a dispositivos móveis. Sendo assim, os classificadores criados são baseados nos descritores LBP, uma vez que estes trabalham com operações de números inteiros e demandam menos recursos computacionais (LINO et al., 2017). Ademais, o

sistema de classificação opera em cascata para focalizar na busca das janelas positivas, em vez de desperdiçar processamento com as janelas negativas (LINO et al., 2017). O algoritmo de aprendizado encarregado da classificação é o *Gentle Adaboost*, pois a forma como os pesos são ajustados em classificações errôneas permite que ele se concentre nas instâncias de difíceis classificações, que poderiam ser descartadas pelo algoritmo *Adaboost* tradicional (WU; NAGAHASHI, 2015).

A primordialidade da aquisição das imagens que contenham os olhos do motorista é a justificativa para o uso do Filtro de Kalman. O rastreamento de características é uma das saídas para gerenciar episódios de oclusão. Nesse sentido, Yan et al. (2016) propuseram um sistema de rastreamento da face, onde as velocidades nas coordenadas x e y se baseiam na movimentação dos pontos iniciais e finais. Na aplicação desenvolvida por este trabalho, o rastreamento é feito com velocidade constante, uma vez que os deslocamentos do motorista dentro do veículo durante a atividade de dirigir são pequenos.

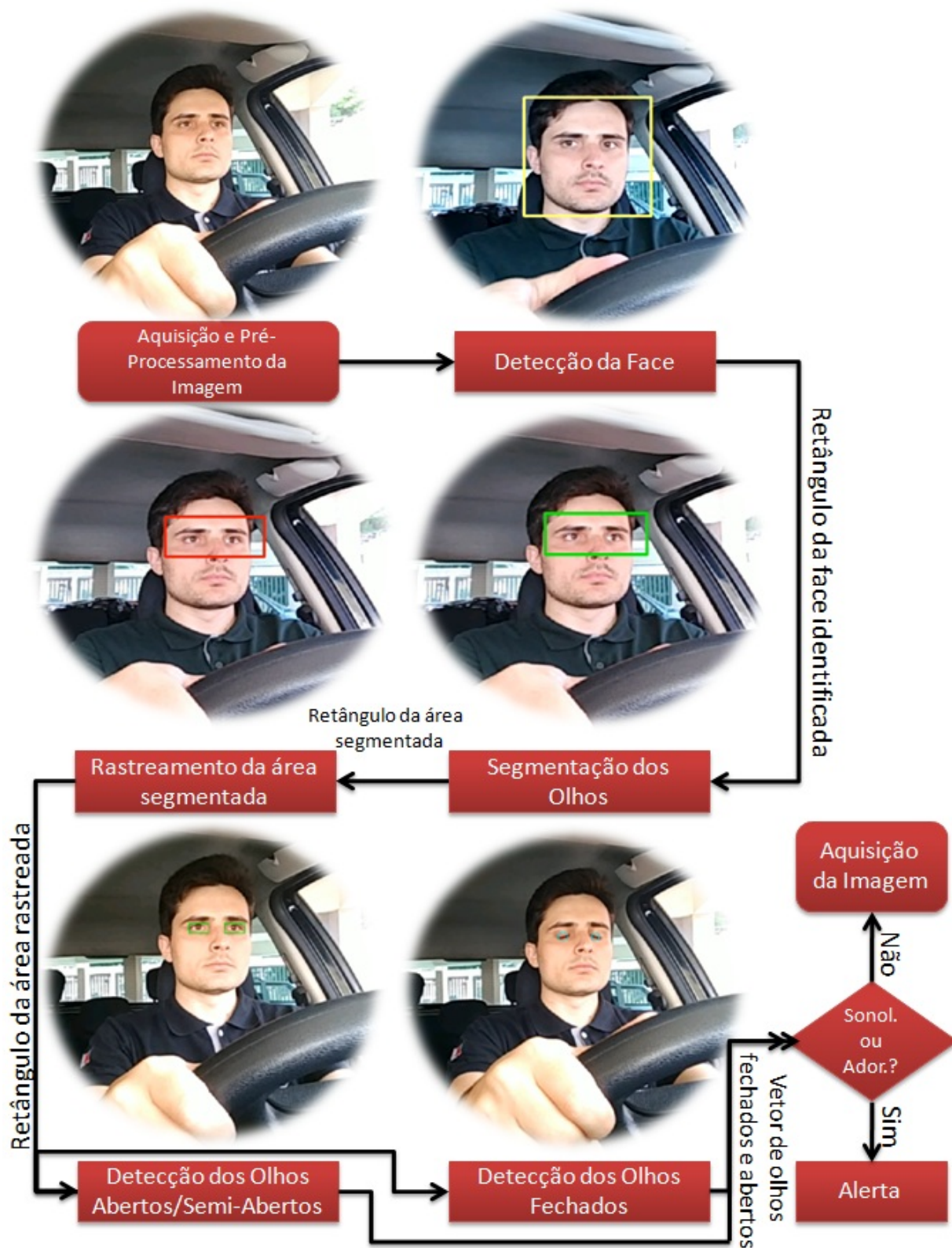
3 METODOLOGIA

Na Seção 2.2 foram apresentados os métodos de identificação de sonolência mais comuns da literatura. Os métodos mais confiáveis usam os sinais elétricos do cérebro, dos olhos e/ou do coração, mas exigem a aplicação de eletrodos, ou outros dispositivos receptores de sinal, sendo que esses procedimentos são considerados invasivos e desconfortáveis. Essa adversidade não existe nos métodos que usam as características visuais do motorista, descritos na Subseção 2.2.2. Tendo em vista a liberdade e o conforto do motorista durante a atividade de dirigir, a aplicação desenvolvida neste trabalho se insere no campo da visão computacional para identificar características visuais que possam indicar sinais de sonolência. Ela é composta pelos processos descritos na Figura 3.1.

A aquisição e o pré-processamento da imagem representam o recebimento e o tratamento dos quadros para criar um ambiente favorável para a detecção de características. A identificação da face é o primeiro procedimento de busca que o algoritmo executa. A principal função da extração da face humana é a aplicação dos identificadores dos olhos em uma região específica, onde sabe-se que os olhos estejam inseridos. Com isso, ao pesquisar pelos olhos somente na região da face, descarta-se o restante da imagem, o que evita um desperdício computacional. O aperfeiçoamento deste processo consiste na segmentação da região dos olhos, onde são descartados, também, regiões superiores e inferiores da face. Esse método foi apresentado em Dasgupta et al. (2013a), mas as equações foram alteradas para adaptar à resolução da aplicação.

O procedimento de rastreamento da região dos olhos é baseado na estimação do Filtro de Kalman, sendo que suas principais contribuições são o gerenciamento de oclusão e a conversão da identificação para um processo contínuo, em vez de discreto. Os olhos abertos e fechados são pesquisados na região dos olhos rastreada pelo Filtro de Kalman. Por fim, é feita a estimação da sonolência e do adormecimento segundo o estimador PERCLOS, que é composto pela razão do número de olhos classificados como fechados, pela soma do número de olhos classificados como abertos com a quantidade de olhos agrupados como fechados. Ele indica o aumento no período de fechamento dos olhos durante o piscar de pessoas sonolentas. A identificação de adormecimento ocorre nas instâncias de fechamento sequencial dos olhos, e é encarregada de identificar o adormecimento repentino e assintomático do condutor. Estas etapas serão melhores detalhadas nas seções a seguir.

Figura 3.1 – Fluxograma da aplicação de identificação de sonolência.



Fonte: Do autor (2018).

3.1 Aquisição e Pré-Processamento de Imagens

A aquisição da imagem é feita pela câmera frontal do *smartphone*, em uma resolução de 640x480 *pixels*, com o formato do tipo paisagem. Um objeto da classe *JavaCameraView* é instanciado e transfere cada *frame* da câmera para uma classe nativa programada na linguagem C++. Esse recurso permite o uso da biblioteca *OpenCV* no ambiente onde ela apresenta melhor performance.

A imagem original tem o modelo de cor vermelho, verde e azul (*red, blue, green* – RGB), o qual é o padrão exibido pela tela do celular. O procedimento de conversão do modelo para escalas de cinza transfere a imagem em *grayscale* para uma nova matriz, na qual são realizados os métodos de identificação. A conversão da imagem em escalas de cinza é justificada pelo uso da equalização do histograma, que é um procedimento encarregado de normalizar o brilho e aumentar o contraste da imagem. Esse tratamento, exibido na Figura 3.2, é usado para criar um ambiente onde os classificadores consigam operar com maior distinção entre os diferentes tons de cinza.

Figura 3.2 – Imagem em RGB (esquerda), escalas de cinza (centro) e com o histograma normalizado (direita).

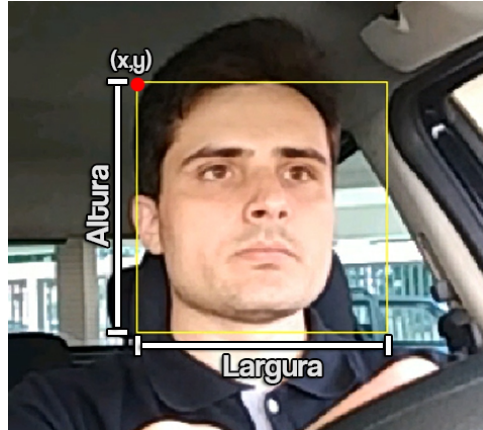


Fonte: Do autor (2018).

3.2 Identificação da Face

A identificação da face é um processo discreto, que é efetuado em cada quadro do vídeo, e resulta em uma área retangular circundante ao rosto. As faces detectadas são armazenadas em um vetor de retângulos, onde cada índice do vetor representa uma face distinta. Os retângulos são compostos por um ponto de origem (x,y), e as medidas de altura e largura (FIGURA 3.3).

Figura 3.3 – Composição de um retângulo na identificação da face.



Fonte: Do autor (2018).

A análise das características da face é efetuada por classificadores treinados com descritores de textura. A decisão pelo uso do classificador com o descritor do tipo *LBP* tem como causa a sua simplicidade computacional (ZHANG; HUA, 2015), citada na Subseção 2.3.2, o que reflete em uma menor diminuição do número de quadros por segundo (*fps*), devido ao pesado processo de detecção da face. O classificador utilizado foi desenvolvido pela equipe do *OpenCV* e está disponibilizado no portal de repositório da biblioteca¹. Nomeado de *lbpcascade_frontalface*, ele identifica faces frontais e rostos com pequenos graus de angulação.

3.3 Segmentação da Área dos Olhos

A segmentação da área dos olhos é um processo de aperfeiçoamento projetado para reduzir a área de pesquisa durante o processo de identificação dos olhos abertos e fechados, tendo em vista a redução de processamento e o ganho de performance (quadros por segundo). Esse procedimento recebe a área retangular da face como entrada e a redimensiona para atingir a área retangular próxima dos olhos, segundo as equações 3.1, 3.2, 3.3 e 3.4.

$$face_segmentada_y = face_y + \frac{face_altura}{5.2}, \quad (3.1)$$

$$face_segmentada_x = face_x + \frac{face_largura}{8}, \quad (3.2)$$

$$face_segmentada_altura = \frac{face_altura}{3}, \quad (3.3)$$

¹ Disponível em: <<https://goo.gl/u2qgFf>>.

$$face_segmentada_{largura} = \frac{face_{largura}}{1.32}. \quad (3.4)$$

A seleção dos valores apresentados nas equações 3.1, 3.2, 3.3 e 3.4 foi baseada na inspeção visual da área dos olhos gerada na resolução da aplicação, considerando uma pequena região adicional para assegurar que os olhos estejam sempre dentro da área de interesse.

3.4 Filtro de Kalman no Rastreamento da Área dos Olhos

O procedimento de detecção da face, que resulta na aquisição da área dos olhos, é executado em cada quadro do vídeo. Nesse método, a área dos olhos somente é calculada quando a face é identificada. Se, por ventura, o motorista efetuar uma atividade que venha a cobrir parte do seu rosto, a identificação da face pode ser prejudicada e não será possível estimar os estados dos olhos. Sendo assim, foi proposto um sistema de rastreamento da área dos olhos com o Filtro de Kalman, que efetua a sua estimação segundo a modelagem baseada na equação do movimento retilíneo uniforme, representada pela equação 3.5.

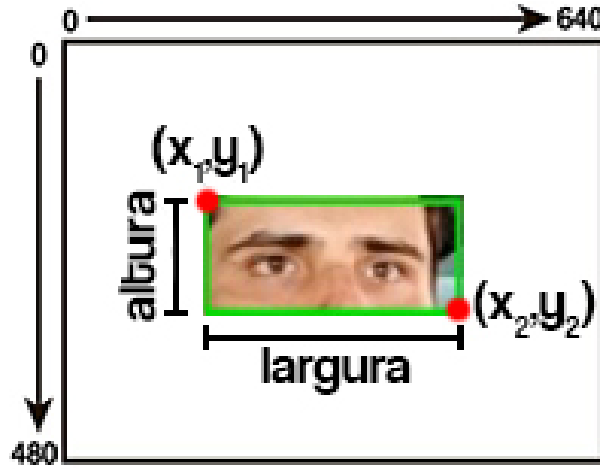
$$x_t = x_{t-1} + \frac{dx_{t-1}}{dt} \cdot \Delta t. \quad (3.5)$$

A nova posição de um ponto unidimensional x_t provém do deslocamento do ponto original x_{t-1} com velocidade $\frac{dx_{t-1}}{dt}$ em um intervalo de tempo Δt . Pode-se, também, calcular a nova velocidade $\frac{dx_t}{dt}$ por meio da derivação da equação 3.5. Esse procedimento, representado pela equação 3.6, confirma a imutabilidade da velocidade no movimento retilíneo uniforme.

$$\frac{dx_t}{dt} = \frac{dx_{t-1}}{dt}. \quad (3.6)$$

A modelagem do sistema de rastreamento, proposta por este trabalho, é feita por meio da estimação de 4 parâmetros, os quais formam dois pontos, como mostrado na Figura 3.4. O primeiro parâmetro, x_1 , é a posição inicial da área dos olhos no eixo x , que representa a localização da borda superior esquerda no eixo das abcissas. O segundo indicador, x_2 representa a posição inicial da área dos olhos no eixo y . O ponto inicial em que o rastreamento é efetuado é composto pelos dois primeiros indicadores. Os parâmetros 3 e 4 são as posições iniciais, x e y , somadas da largura e da altura do retângulo que envolve a área dos olhos, e formam o ponto final do rastreamento (x_2, y_2) . A modelagem dos parâmetros, segundo o movimento retilíneo uniforme, é apresentada, matricialmente, pela equação 3.7.

Figura 3.4 – Representação dos pontos rastreados pelo Filtro de Kalman.



Fonte: Do autor (2018).

$$\begin{bmatrix} \hat{x}_{1,t|t-1} \\ \hat{y}_{1,t|t-1} \\ \hat{x}_{2,t|t-1} \\ \hat{y}_{2,t|t-1} \\ \frac{d\hat{x}_{t|t-1}}{dt} \\ \frac{d\hat{y}_{t|t-1}}{dt} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & \Delta t & 0 \\ 0 & 1 & 0 & 0 & 0 & \Delta t \\ 0 & 0 & 1 & 0 & \Delta t & 0 \\ 0 & 0 & 0 & 1 & 0 & \Delta t \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \hat{x}_{1,t-1|t-1} \\ \hat{y}_{1,t-1|t-1} \\ \hat{x}_{2,t-1|t-1} \\ \hat{y}_{2,t-1|t-1} \\ \frac{d\hat{x}_{t-1|t-1}}{dt} \\ \frac{d\hat{y}_{t-1|t-1}}{dt} \end{bmatrix}. \quad (3.7)$$

A equação 3.7 é a representação da matriz de estado segundo o modelo adotado pela matriz de transição. Essa é a primeira etapa do processo de predição do Filtro de Kalman. A partir de agora, os parâmetros são descritos com 2 subscritos. O primeiro indica o tempo ($t - 2$, $t - 1$, t) em que o parâmetro está sendo abordado, e o segundo subscrito indica a fase do Filtro de Kalman ($t - 1$ é a fase de predição; t é a fase de estimação). Eventualmente, as equações matriciais são resumidas em equações simples, onde cada matriz é representada por uma letra. Sendo assim, a equação 3.8 apresenta a matriz de estado no formato simbólico, onde $\hat{X}$ é a matriz de estado, e A é a matriz de transição. Devido à simplicidade do sistema de estimação de posição, proposto neste trabalho, as matrizes de controle e de ruído foram desconsideradas.

$$\hat{X}_{t|t-1} = A \cdot \hat{X}_{t-1|t-1}. \quad (3.8)$$

A segunda etapa consiste na predição da covariância, que é a medição estimada da precisão dos parâmetros em um tempo determinado. A estimação é calculada por meio da equação 3.9, onde A representa o modelo do sistema, P_t é a predição da covariância no tempo t , e Q é

a matriz de ruído de processo. Esta última mede a variabilidade dos sinais de entrada longe das transições ideais definidas para a matriz de transição. Maiores valores nesta matriz significam que os sinais de entrada têm maior variação, e o filtro precisa ser mais adaptável. Valores menores resultam em uma saída mais suave, mas o filtro perde a adaptabilidade às grandes mudanças.

$$P_{t|t-1} = A.P_{t-1|t-1}.A^T + Q. \quad (3.9)$$

A inicialização da matriz de predição da covariância é requerida para preencher o primeiro elemento de $P_{t-1|t-1}$. Os valores podem ser preenchidos com os desvios padrões de cada parâmetro, no formato de uma matriz diagonal, ou por meio de testes ajustáveis, quando os desvios não são notórios. Neste trabalho, foram realizados testes para verificar a velocidade de inicialização, segundo a variação de $P_{t-1|t-1}$, uma vez que $P_{t|t-1}$ se ajusta com o tempo. Foi notado que em valores próximos a zero, o sistema de rastreamento inicializa vagarosamente. O motivo dessa ocorrência é a influência que a matriz de predição da covariância tem sobre o Ganho de Kalman, apresentado pela equação 3.12. Quando o sistema é inicializado com o ganho próximo a zero, o rastreamento é feito, majoritariamente, pelo modelo da matriz de transição. Sendo assim, como os ponderamentos das medições são baixos para influenciar no ajuste da estimação inicial e o estado inicial foi definido como zero, o modelo não consegue atuar isolado. Foi definida, então, a composição da matriz $P_{t-1|t-1}$ que permite a atuação do modelo e a correção segundo as medições (equação 3.10). A matriz diagonal representa a não correlação entre os elementos.

$$P_{t-1|t-1} = \begin{bmatrix} 10 & 0 & 0 & 0 & 0 & 0 \\ 0 & 10 & 0 & 0 & 0 & 0 \\ 0 & 0 & 10 & 0 & 0 & 0 \\ 0 & 0 & 0 & 10 & 0 & 0 \\ 0 & 0 & 0 & 0 & 10 & 0 \\ 0 & 0 & 0 & 0 & 0 & 10 \end{bmatrix}. \quad (3.10)$$

A composição da matriz de ruído de processo Q é importante para um bom desempenho do filtro. Deve-se fazer uma ponderação entre velocidade de rastreamento e estabilidade para atingir o valor ótimo. A velocidade de rastreamento indica o quão rápido o sistema consegue se adaptar para rastrear um alvo com uma velocidade definida. Teoricamente, é desejado que o ras-

treamento atinja alvos com valores diversos de velocidade. Porém, na prática, é observado que a estabilidade é afetada quando o sistema apresenta alto grau de adaptabilidade de velocidade. Quando a matriz de ruído de processo Q é preenchida com valores superiores a 1, é possível rastrear a área dos olhos em situações com alto grau de mobilidade, isto é, situações em que o motorista se movimenta dentro do veículo constantemente com uma velocidade considerável. Por outro lado, em situações usuais de oclusão da face, o rastreamento se perde vagamente, mesmo quando o deslocamento horizontal ou vertical da face é pequeno. Sendo assim, foi feita uma análise dos valores da matriz Q para compensar o desempenho das propriedades de velocidade e estabilidade, e o resultado é dado por:

$$Q = \begin{bmatrix} 0.1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.1 \end{bmatrix}. \quad (3.11)$$

O estágio de estimação é iniciado com o cálculo do ganho de Kalman (equação 3.12), K_t , que é usado para ponderar entre os estados previstos e os estados medidos (equação 3.16). Valores de ganho inferiores indicam que as medições dos parâmetros possuem pouca representabilidade na estimação do filtro, isto é, o sistema atribui maior importância aos estados previstos. Distintivamente, elevações no ganho de Kalman indicam que o sistema transfere maior relevância aos estados medidos.

$$K_t = P_{t|t-1} H_t^T (H_t P_{t|t-1} H_t^T + R)^{-1}. \quad (3.12)$$

R é a matriz de covariância do ruído de medição e representa as incertezas das localizações dos parâmetros medidos em relação as localizações dos parâmetros estimados. Essencialmente, R define a quantidade de informação dos parâmetros medidos que é usada no rastreamento. Valores elevados de R consideram que as medições não são precisas, portanto o modelo é supervalorizado. Tendo em vista a variabilidade de diferentes câmeras usadas pelos *smartphones* e a dificuldade na definição de uma incerteza de medição, a matriz R , representada pela equação 3.13, foi estabelecida com valores referentes a 1% da resolução da câmera.

$$R = \begin{bmatrix} 6,4 & 0 & 0 & 0 \\ 0 & 4,8 & 0 & 0 \\ 0 & 0 & 6,4 & 0 \\ 0 & 0 & 0 & 4,8 \end{bmatrix}. \quad (3.13)$$

A matriz H é o modelo de observação que mapeia o espaço de estados no espaço observado. Essa matriz é composta unitariamente pelos estados em que há medições disponíveis. Neste trabalho, a matriz de transição foi modelada para atender 6 parâmetros (equação 3.7), onde 4 parâmetros são mensuráveis (x_1, y_1, x_2 e y_2). Sendo assim, a matriz de observação H tem a dimensão 4x6, e é representada pela equação 3.14.

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}. \quad (3.14)$$

A matriz Y , nominada de resíduo de medição, e representada pela equação 3.15, é o resultado da subtração entre os valores medidos, Z , e os valores dos estados obtidos na fase de predição, $\hat{X}_{t|t-1}$. Ela indica a precisão dos valores estimados pelo Filtro de Kalman perante o valor medido pela ferramenta de identificação. No sistema proposto, os valores medidos compõem os pontos iniciais, (x_1, y_1) , e os pontos finais, (x_2, y_2) , do retângulo que delimita a área dos olhos. Esses valores são obtidos pela saída do procedimento de segmentação da área dos olhos, que foram descritos na Seção 3.3. Os valores estimados na fase de predição são adquiridos segundo o modelo apresentado na equação 3.7.

$$\tilde{Y}_t = Z_t - H_t \hat{X}_{t|t-1}. \quad (3.15)$$

Durante episódios de oclusão parcial ou total, a matriz dos valores medidos Z_t não possui nenhum valor, visto que o procedimento de identificação da área dos olhos não consegue detectar a face para efetuar a segmentação. Essa atividade é verificada pelo tamanho do vetor de faces, que retorna zero quando nenhuma face é identificada. Sendo assim, no decorrer da oclusão, o resíduo de medição deve ser considerado como nulo, para que o sistema aja consoante com o modelo. Essa atividade permite o rastreamento da área dos olhos sem o requerimento dos dados de medição. Desse modo, é possível analisar o estado dos olhos mesmo quando parte da

face do motorista está encoberta, ou quando o sistema de identificação da face apresenta falhas na detecção.

A segunda etapa do estágio de estimação consiste na atualização da matriz de estados estimada, exibida pela equação 3.16. A nova matriz de estados é o resultado da soma da matriz de estados extraída na fase de predição, com a multiplicação entre o ganho de Kalman e os resíduos de medição. A saída desse procedimento representa os pontos estimados, que são usados para o rastreamento da área dos olhos.

$$\hat{X}_{t|t} = \hat{X}_{t|t-1} + K_t \tilde{Y}_t. \quad (3.16)$$

O último procedimento adotado na implementação do filtro de Kalman para o rastreamento da área dos olhos é a estimação da covariância (equação 3.17). A nova covariância estimada é usada no próximo ciclo para a atualização da predição da covariância.

$$P_{t|t} = (I - K_t H_t) P_{t|t-1}. \quad (3.17)$$

3.5 Identificação dos Olhos Abertos e Semi-Abertos

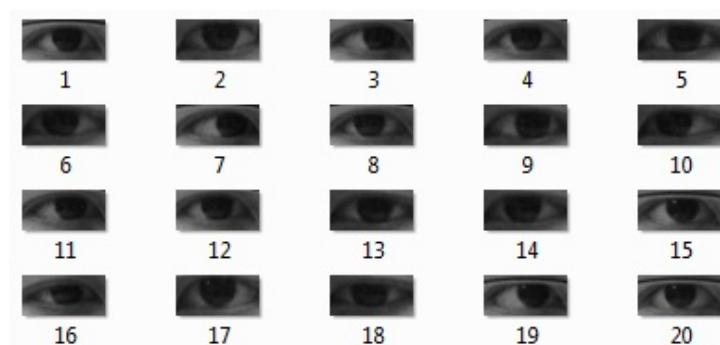
Na Seção 3.2, foi salientado que a biblioteca *OpenCV* disponibiliza classificadores em cascata, treinados com descritores de textura, para a comunidade científica. Especificamente, os classificadores treinados com descritores *LBP* acessíveis pela biblioteca identificam a face de humanos (frontal ou perfil), a face frontal de gatos e talheres de cozinha. A necessidade pela identificação de olhos abertos ou semi-abertos (até 20% de abertura) levou à criação de um classificador próprio.

A criação de um classificador em cascata envolve a elaboração e o treinamento de um conjunto de dados com imagens positivas e imagens negativas, conforme apresentado na Seção 2.4. As imagens positivas devem conter itens os quais se deseja identificar. As imagens negativas devem descrever o que não deve ser classificado como positivo. Sendo assim, o conjunto de dados desenvolvido contém 5330 imagens de olhos abertos, o que correspondem as amostras positivas. Os olhos foram recortados manualmente de imagens provenientes dos conjuntos de dados de Smith et al. (2013) e de Wood et al. (2015).

Os olhos recortados são de imagens de faces frontais ou com pequenos graus de angulação, sendo que a resolução de todas as amostras é de 100x50 *pixels*. Esse valor é justificado

pelo maior comprimento horizontal que os olhos possuem. Todas as imagens estão no formato *.jpg* e foram numeradas sequencialmente. A Figura 3.5 apresenta algumas destas amostras. Elas possuem brilhos diversificados e passaram pelo procedimento de espelhamento horizontal para aumentar a generalização do conjunto de dados.

Figura 3.5 – Amostras de imagens positivas.



Fonte: Do autor (2018).

As imagens negativas devem conter elementos presentes no cenário de busca que não devem ser classificados como positivos. Inicialmente, a identificação dos olhos havia sido desenvolvida para a região da face humana. Sendo assim, foram recortadas 6000 imagens de elementos como boca, nariz, queixo, sobrancelha, bochecha, testa e, principalmente, olhos fechados (FIGURA 3.6). Porém, após a criação do procedimento de segmentação dos olhos, apresentado na Seção 3.3, vários desses elementos se tornaram dispensáveis. Apesar disso, decidiu-se manter o classificador original, visto que sua eficácia em ambas regiões é elevada, como será quantificado nos resultados. As amostras negativas foram recortadas das imagens do conjunto de dados criado por Gao et al. (2008).

Figura 3.6 – Amostras de imagens negativas.



Fonte: Do autor (2018).

O treinamento foi realizado pelo algoritmo *Gentle Adaboost*, descrito na Seção 2.4, onde 90% das amostras positivas foram usadas para treinamento, e 10% foram usadas para teste. Inicialmente, o classificador havia sido treinado para identificar olhos com resolução mínima de 10 por 5 *pixels*. Porém, durante o treinamento, verificou-se que o mesmo apresentava dificuldade para alcançar a taxa de sucesso mínima definida para o treinamento, que é 99,2%. Em função disso, os testes não apresentaram o desempenho requerido para a aplicação. Sendo assim, a resolução mínima do classificador foi duplicada para 20 por 10 *pixels*, onde foi possível treiná-lo com uma taxa de sucesso mínima de 99,2%, e com um valor de alarme falso máximo de $10^{-7}\%$. O treinamento foi realizado em 8 estágios e o arquivo do classificador gerado possui 54,4 KB.

A área dos olhos, segmentada na Seção 3.3, é rastreada e enviada para o procedimento de identificação dos olhos abertos por meio de um objeto do tipo retângulo (ver Seção 3.3). Essa área é a região de interesse ao qual os olhos são pesquisados, onde descarta-se o restante da imagem. Portanto, diminui-se o custo computacional e a probabilidade de classificações falsas que interfeririam na estimação da sonolência.

A identificação dos olhos abertos é reportada pelo tamanho do vetor de retângulos que é retornado da função de detecção. Se o tamanho desse vetor for maior que zero, pode-se inferir que, pelo menos, uma janela foi classificada como positiva. Se a classificação estiver correta, significa que ao menos um dos olhos abertos foi identificado.

3.6 Identificação dos Olhos Fechados

O procedimento de identificação dos olhos fechados é semelhante à identificação dos olhos abertos. A criação de um classificador em cascata de olhos fechados com boa performance requer um conjunto de dados com imagens positivas que apresentam boa qualidade. O primeiro conjunto de dados elaborado contou com 10000 imagens de olhos fechados e olhos parcialmente fechados (FIGURA 3.7), redimensionadas em 60x30 *pixels*, recortadas das faces oriundas de Gao et al. (2008). As imagens originais contém faces de pessoas a uma distância considerável da câmera e, em função disso, a qualidade das imagens dos olhos não era adequada. As amostras negativas foram extraídas de Gao et al. (2008), Smith et al. (2013) e Wood et al. (2015), e representam imagens negativas de elementos da face, como nariz, boca, queixo, sobrancelha e olhos abertos.

Figura 3.7 – Amostras de imagens positivas do primeiro conjunto de dados.



Fonte: Do autor (2018).

O treinamento do classificador de olhos fechados foi feito pelo algoritmo Gentle Adaboost (ver Seção 2.4) com 90% das amostras positivas usadas para treinamento e 10% para teste. O classificador foi treinado com o descritor *LBP* para identificar janelas de 20x10 *pixels*, de forma similar ao classificador dos olhos abertos. Porém, durante os testes, os resultados foram extremamente insatisfatórios e decidiu-se criar um novo conjunto de dados com imagens de melhor qualidade para analisar a diferença de desempenho.

As novas amostras positivas foram extraídas de imagens no portal *Google*. No total, 918 figuras de olhos fechados, com qualidade superior, foram adquiridas (FIGURA 3.9). As novas amostras passaram pela operação de ajuste do *gamma*, que consiste na alteração do brilho original da figura. O *gamma* foi ajustado em 0.7 e em 1.1 (FIGURA 3.9), visto que as imagens originais possuem níveis de brilho elevado. Essa operação tem como objetivo aumentar a generalização do sistema de classificação dos olhos fechados, uma vez que a iluminação natural da imagem varia segundo a posição da câmera e o horário do dia.

Figura 3.8 – Amostras de imagens positivas do segundo conjunto de dados.



Fonte: Do autor (2018).

O novo conjunto de dados possui as imagens positivas originais e as ajustadas, o que totaliza 2754 amostras, e as 5330 imagens negativas que foram usadas no primeiro conjunto

Figura 3.9 – Amostra original (esquerda), com o *gamma* reduzido (centro) e com o *gamma* elevado (direita).



Fonte: Do autor (2018).

de dados. Novamente, todas as imagens se encontram no formato *.jpg*. No segundo conjunto, entretanto, as amostras possuem tamanho de 80×40 *pixels*. Novamente, o treinamento do novo classificador foi feito pelo algoritmo *Gentle Adaboost* com descritores do tipo *LBP* para identificar janelas de 20×10 *pixels*, em um número total de 8 estágios. 90% das amostras positivas foram usadas para o treinamento e 10% foram usadas para o teste.

A identificação dos olhos fechados é relatada de modo igual à identificação dos olhos abertos. O vetor de retângulos que retorna da função de detecção possui tamanho maior que zero se alguma janela for classificada como positiva. Pode-se inferir que, se a classificação estiver correta, significa que ao menos um dos olhos fechados foi identificado.

3.7 Detecção de Sonolência e Alerta ao Motorista

A identificação de sonolência é realizada por meio de dois procedimentos distintos. O primeiro é baseado na estimação contínua do índice PERCLOS, em tempo-real, durante uma janela definida de tempo. Esse método identifica o sintoma de sonolência que faz com que os olhos permaneçam fechados por um período longo de tempo durante o piscar. O segundo procedimento analisa os estados dos olhos para verificar o adormecimento repentino do motorista, em um curto intervalo de tempo.

3.7.1 Identificação de Sonolência Segundo a Estimação do PERCLOS

O método de estimação do PERCLOS usado na aplicação é baseado no procedimento apresentado por Yan et al. (2016). Os olhos são monitorados, em tempo-real, e são classificados em abertos e fechados. A equação 3.18 apresenta a formulação para o cálculo do PERCLOS. Os olhos classificados como fechados são representados por *NC*, e os olhos classificados como abertos são denotados por *NO*.

$$PERCLOS = \frac{NC}{NC + NO}. \quad (3.18)$$

O incremento de *NC* acontece quando existem *frames* em que os olhos classificados como fechados são identificados. A metodologia proposta necessita que pelo menos um dos olhos fechados seja detectado, uma vez que é analisado se o tamanho do vetor de saída da função *detectMultiScale* é maior que zero. Esse recurso é um alento em ocorrências de falsos negativos em um dos olhos, visto que o outro identificado é suficiente para a incrementação de *NC*. De forma semelhante, quando os dois olhos são classificados como fechados em um mesmo *frame*, o incremento em *NC* é de apenas uma unidade. Sendo assim, é possível identificar a quantidade de quadros em que o motorista apresenta-se com os olhos fechados dentro de um intervalo de tempo. O incremento de *NO* realiza-se identicamente ao de *NC*, mas com os olhos classificados como abertos.

A identificação de sonolência é uma função do tipo limiar (equação 3.19) e possui como saída o parâmetro de alerta responsável por sinalizar o motorista de que ele apresenta sinais de sonolência.

$$WARNING = \begin{cases} 1, & \text{se } PERCLOS > \text{limiar da sonolência,} \\ 0, & \text{senão.} \end{cases} \quad (3.19)$$

3.7.2 Identificação de Adormecimento

A detecção de adormecimento é um procedimento de análise dos olhos, em um curto intervalo de tempo, que identifica sinais de que o motorista está dormindo. Esse método é complementar à estimação do PERCLOS, e foi criado para atender as situações onde o motorista adormece rapidamente sem apresentar sinais de sonolência.

O adormecimento de motoristas é identificado pelo parâmetro *FALL ASLEEP* – FA, no qual é baseado nos eventos de classificação de olhos abertos, NO, e fechados, NC, conforme a equação 3.20. Em situações onde o motorista transita entre o estado de sonolência e adormecimento, o fechamento ininterrupto dos olhos é um dos primeiros sintomas. Sendo assim, a equação 3.21 foi desenvolvida para identificar esse sintoma. A variável *WARNING* recebe o valor 2 quando o limiar de adormecimento é ultrapassado pelo parâmetro FA. Sendo assim, são emitidos dois tipos diferentes de alarme: suave para os sinais de sonolência, e intenso para os sinais de adormecimento.

$$FA = \frac{NC}{NC + NO}, \quad (3.20)$$

$$WARNING = \begin{cases} 2, & \text{se } FA > \text{limiar do adormecimento,} \\ 0, & \text{senão.} \end{cases} \quad (3.21)$$

4 RESULTADOS E DISCUSSÕES

4.1 Implementação

Os procedimentos apresentados no Capítulo 3 foram implementados no *smartphone* Motorola Z Play XT1635-02. Esse telefone possui um processador octa-core 2.0 GHz Cortex-A53 e uma unidade de processamento gráfico Adreno 506. Sua memória RAM é de 3 GB e a câmera frontal dispõe de 5 MP com uma abertura de $f/2.2$. O sistema operacional a partir do qual os testes deste capítulo foram efetuados é o Android 7.1.1 (Nougat).

A versão da biblioteca de visão computacional *OpenCV* adotada foi a 3.1.0, lançada em 18/12/2015. Na época em que o projeto foi iniciado, essa era atualização mais recente. A versão atual é a 3.4.0, que foi lançada em 23/12/2017.

O procedimento de identificação das características proposto por este trabalho usa o algoritmo de aprendizado de máquina *Gentle Adaboost* (ver Seção 2.4) em cascata para classificar cada sub-janela da imagem. O processo de classificação exige dois parâmetros: fator de redução de escala (*scaleFactor*) e número de vizinhos mínimos (*minNeighbors*).

A redução de escala é a medida de redução da imagem (em x e em y) onde são efetuadas as buscas. Inicialmente, o algoritmo procura indícios da característica em sub-janelas da imagem original e armazena as áreas localizadas. Em seguida, a mesma busca é feita nas imagem reduzidas até que se atinja um tamanho mínimo, onde as características localizadas são redimensionadas para o tamanho da imagem original. Esse procedimento aumenta a possibilidade de identificação, uma vez que são realizadas buscas em ambientes de diferentes dimensões. Entretanto, quanto menor a redução de escala, maior é o número de operações executadas pelo algoritmo e eleva-se o custo computacional.

O parâmetro *minNeighbors* indica a quantidade de janelas adjacentes classificadas como positivas que são requeridas para uma classificação positiva. Durante as buscas na imagem original e nas imagens reduzidas, cada janela classificada como positiva tem sua localização armazenada. No final do procedimento de identificação, as posições dessas janelas são comparadas. Em localizações onde a quantidade de sub-janelas positivas adjacentes é maior do que o parâmetro *minNeighbors*, a classificação é dada como positiva. Esse parâmetro é essencial para corrigir os falsos positivos, que correspondem as classificações errôneas de características.

Os parâmetros *minSize* e *maxSize* não são requeridos no processo de classificação, mas são essenciais para o aprimoramento da performance da aplicação. Eles são encarregados de

limitar o tamanho mínimo e o tamanho máximo das janelas de busca, com o intuito de reduzir o número de operações executadas durante as buscas das características. Como consequência desse procedimento, objetos menores que *minSize* ou objetos maiores que *maxSize* não são pesquisados e, obviamente, não são identificados.

4.2 Identificação da Face

Na Seção 4.1, foi apresentado um método de aperfeiçoamento da performance da aplicação, que consiste na limitação do tamanho das janelas de busca. A definição de um intervalo específico tem a finalidade de reduzir a quantidade de operações que o algoritmo de pesquisa efetua e, conseqüentemente, elevar o número de quadros por segundo da aplicação. O modelo de identificação tradicional, divulgado e usado como exemplo na página da biblioteca OpenCV¹, adota um limiar de busca inferior, de 30x30 *pixels*. As janelas de pesquisa iniciam-se nesse tamanho e vão se expandindo até o tamanho total da resolução da imagem.

Neste trabalho, o intervalo de distâncias entre a câmera e o motorista foi definido entre 50 a 100 centímetros. O espaço de 50 centímetros é justificado pela grande variabilidade de posições em que o banco do motorista pode ser deslocado. Ao definir um espaço de distâncias onde a aplicação deve operar, faces menores que 50 ou maiores que 100 centímetros não precisam ser pesquisadas. As faces de motoristas que se encontram a cem centímetros de distância da câmera, ocupam uma área aproximada de 140x140 *pixels* na imagem, enquanto faces à 50 centímetros da câmera, envolvem 100x100 *pixels*. Sendo assim, os parâmetros *minSize* e *maxSize* foram preenchidos com (100,100) e (140,140), respectivamente. Dessa forma, as janelas de busca iniciam no tamanho mínimo e finalizam no tamanho máximo. A Figura 4.1 ilustra a execução da aplicação dentro e fora do intervalo descrito. As faces, representadas pelo retângulo amarelo, somente são pesquisadas dentro do espaço definido. Foi definido que o número de *minNeighbors* deve ser igual ou superior a 3, uma vez que esse valor é suficiente para que não sejam geradas excessivas classificações falsamente positivas. Por fim, a escala foi reduzida em 20%.

O sistema de identificação de faces apresentou bons resultados para faces frontais e com angulações de até 30 graus. Sendo assim, é recomendável que a aplicação seja usada dentro desse intervalo, uma vez que a detecção da face é primordial para que a área dos olhos seja calculada e a estimativa seja atualizada.

¹ Disponível em: <<https://goo.gl/kJE25K>>.

Figura 4.1 – Faces dentro (superior) e fora (inferior) do intervalo de 50 à 100 centímetros.



Fonte: Do autor (2018).

A avaliação do ganho de performance é baseada no valor de quadros por segundo (fps - do inglês *frame per second*) que a aplicação processa. A análise foi feita em 20 segundos de vídeos gravados em situações de iluminação diferentes, com a resolução de 640x480 *pixels*, onde os valores de média e desvio padrão são apresentados pela Tabela 4.1. Os números foram arredondados a partir da terceira casa decimal. A primeira linha de dados representa o método tradicional, no qual é aplicado somente a limitação inferior, de tamanho 30x30 *pixels*. O aperfeiçoamento empregado por este trabalho é representado pela segunda linha de dados. A limitação do tamanho de busca potencializou a performance da aplicação em 93,01%, segundo os valores das médias do número de quadros por segundo. Essa alta disparidade é justificada pela quantidade elevada de operações que o método tradicional efetua, em comparação com o método aprimorado.

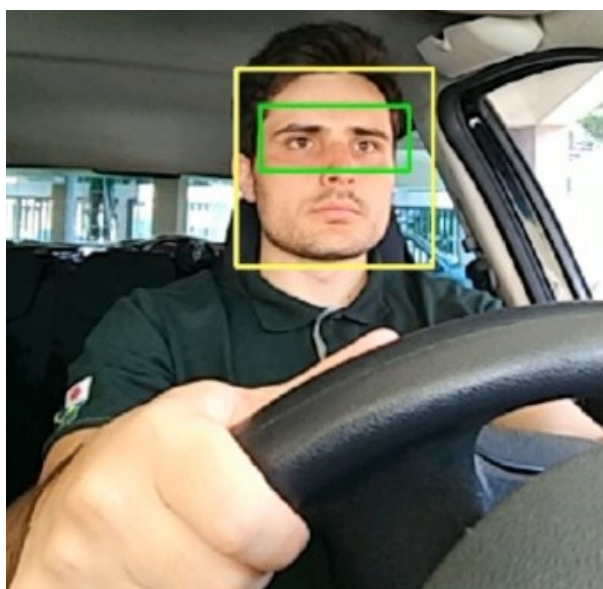
Tabela 4.1 – Desempenho da aplicação na configuração tradicional e na aprimorada.

Configuração de parâmetros	Fps
$minSize = (30,30)$	$15,427 \pm 0,412$
$minSize = (100,100)$ e $maxSize = (140,140)$	$29,789 \pm 0,527$

4.3 Segmentação da Área dos Olhos

A segmentação da área dos olhos, apresentada na Seção 3.3, é um procedimento destinado a reduzir a área da face para uma região mínima que contenha os olhos. A área segmentada, detalhada na Figura 4.2 e descrita pelas equações 3.1, 3.2, 3.3 e 3.4, representa 25.25% da área da face original. A justificativa para a redução da região de busca dos olhos é a diminuição da demanda de recursos computacionais que a aplicação requer para identificar os olhos abertos e fechados.

Figura 4.2 – Área da face detectada (amarelo) e área dos olhos segmentada (verde).



Fonte: Do autor (2018).

A Tabela 4.2 apresenta o desempenho da aplicação operando com a identificação dos olhos na área da face e na região segmentada. O objetivo da comparação é visualizar o ganho de performance com a operação dos classificadores na região próxima aos olhos. A análise foi praticada em uma gravação de 34 segundos da aplicação, da qual os valores de fps foram extraídos e as médias e os desvios padrões foram computados. A melhoria proporcionada pela segmentação da região dos olhos é de 2,620 fps, segundo a diferença das médias, o que representa 12,6% de evolução.

Tabela 4.2 – Desempenho da aplicação com os classificadores em operação na face e na área segmentada.

Busca dos olhos	Fps
Área da Face	20,792 $\pm$ 0,433
Área Segmentada	23,412 $\pm$ 0,345

4.4 Rastreamento da Área dos Olhos

Na Seção 3.4, foi citado que o rastreamento da área dos olhos é necessário para gerenciar situações de oclusão em que a face não é identificada. O procedimento de rastreamento com o filtro de Kalman estima a posição da região dos olhos e permite que os classificadores operem na avaliação da sonolência mesmo quando a face está encoberta. A Figura 4.3 expõe quatro situações comuns em que a face se encontra semi-oclusa. O retângulo vermelho representa a região estimada pelo filtro de Kalman, e os retângulos verdes retratam a classificação positiva de olhos abertos. Em todos os eventos, a face não foi identificada, mas o rastreamento permitiu que os olhos fossem avaliados e classificados.

Figura 4.3 – Rastreamento da área dos olhos em situações de oclusão.



Fonte: Do autor (2018).

A avaliação quantitativa do filtro foi feita com base no erro quadrático médio (EQM) entre as variáveis de medição (X_1, Y_1, X_2, Y_2) e os valores estimados ($\hat{X}_1, \hat{Y}_1, \hat{X}_2, \hat{Y}_2$). O objetivo dessa medição é avaliar a média dos desvios ao quadrado entre a medição e a estimação, em

que valores próximos a zero correspondem a melhores estimadores. A equação 4.1 descreve o cálculo do EQM das variáveis que formam os pontos em que a área dos olhos é rastreada.

$$EQM = \frac{1}{n} \sum_{i=1}^n (X - \hat{X})^2, \quad (4.1)$$

Uma situação comum de dirigibilidade foi simulada, na qual o motorista mantém a face direcionada para a pista. As diferenças entre os valores medidos e os valores estimados são obtidas pelo vetor de resíduos (ver equação 3.15). A avaliação consiste no cálculo do erro quadrático médio de cada variável em 30 quadros sequenciais. Os resultados, descritos na Tabela 4.3, foram arredondados na terceira casa decimal.

Tabela 4.3 – EQM das variáveis de rastreamento da área dos olhos.

Variável	EQM (<i>pixels</i> ²)
x_1	0,971
y_1	0,732
x_2	1,739
y_2	1,198

Os erros quadráticos médios de x_1 e x_2 representam, respectivamente, 0,151% e 0,272% da resolução total da câmera no eixo x . Semelhantemente, os números de EQM das variáveis em y representam 0,153% e 0,249%. Os baixos valores de EQM demonstram que o modelo definido para a matriz de transição é preciso, e o preenchimento das matrizes de ruído de processo, Q , e de covariância do ruído de medição, R , foi suficiente para estimar a região dos olhos do condutor próxima a da medição obtida pelo processo de identificação.

4.5 Detecção de Olhos Abertos

O processo de identificação de olhos abertos é encarregado de informar eventos em que os olhos abertos ou semi-abertos se encontram presentes na imagem. Os classificadores operam na região segmentada dos olhos, e as amostras classificadas como positivas são ilustradas por retângulos verdes (ver Figura 4.4). Para reduzir o número de operações efetuadas pelo classificador, o intervalo de buscas foi definido em 20x10 à 30x15 *pixels*. Esse valor é suficiente para identificar olhos abertos em distâncias de 50 à 100 centímetros.

A redução de escala que apresentou a melhor razão de performance/esforço foi a de 10%. A elevação desse valor diminui o tamanho da imagem a ser pesquisada, e isso diminui a quantidade de processamento requerido pela aplicação. Entretanto, em imagens com tamanhos

Figura 4.4 – Identificação de olhos abertos e semi-abertos.



Fonte: Do autor (2018).

inferiores, a performance da detecção foi reduzida significativamente, uma vez que a quantidade de *pixels* dos alvos a serem analisados é menor. Em contrapartida, a diminuição na redução de escala elevou o gasto com processamento, o que reduziu o número de fps.

O número mínimo de identificações vizinhas (*minNeighbors*) requeridas para a classificação positiva de uma janela foi definido como 3. Esse é o menor valor que apresentou nenhuma amostra classificada como falso-positiva em todos os testes quantitativos. Sendo assim, torna-se desnecessário o uso de um valor superior em *minNeighbors*.

A avaliação quantitativa do classificador de olhos abertos foi feita em um experimento com 4 pessoas, as quais possuem idades diferentes e são compostas por três mulheres e um homem. As descrições dos indivíduos, nomeados por letras, são descritas na Tabela 4.4.

Tabela 4.4 – Dados dos indivíduos participantes do experimento.

Indivíduo	Idade	Gênero	Tonalidade dos Olhos
A	24	Feminino	Claro
B	48	Feminino	Claro
C	26	Feminino	Escuro
D	21	Masculino	Escuro

O experimento é baseado na análise dos parâmetros verdadeiro positivo (*true positive* – TP), falso positivo (*false positive* – FP), falso negativo (*false negative* – FN) e verdadeiro negativo (*true negative* – TN). Os indicadores verdadeiro positivo e negativo indicam as amostras que foram corretamente classificadas como positivas e negativas, respectivamente. O falso positivo representa uma amostra que não é um olho aberto e foi classificada como positiva. O falso negativo é a ocorrência de classificação negativa de uma amostra de olho aberto. Os in-

dicadores são usados para a computação da precisão (*precision*), revocação (*recall*), medida-F (*F-measure*) e acurácia (*accuracy*), os quais são retratados pelas equações 4.2, 4.3, 4.4 e 4.5.

$$Precision = \frac{TP}{TP + FP}, \quad (4.2)$$

$$Recall = \frac{TP}{TP + FN}, \quad (4.3)$$

$$F - measure = (1 + \beta^2) \frac{Precision \cdot Recall}{Precision + Recall}, \quad (4.4)$$

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}, \quad (4.5)$$

onde $\beta = 1$ define a média harmônica (*F1-measure*) entre a precisão e a revocação. Os indivíduos foram orientados a simularem uma situação comum de dirigibilidade, onde a face e os olhos estão voltados para a pista. O experimento foi registrado entre 13:10 e 13:30, em um dia ensolarado, uma vez que é desejável descobrir se os efeitos da iluminação natural afetam a musculatura dos olhos abertos. Durante o experimento, o veículo se encontrava parado, visto que seria perigoso realizar os testes com o condutor no controle, pois foi pedido para que os indivíduos permanecessem com os olhos abertos o máximo de tempo que conseguirem. No total, para cada indivíduo, foram usados 90 quadros sequenciais em que os olhos estão abertos ou semi-abertos. Os resultados da análise quantitativa são retratados na Tabela 4.5.

Tabela 4.5 – Dados da avaliação do classificador de olhos abertos.

Indivíduo	Precisão	Revocação	<i>F-Measure</i>	Acurácia
A	100,00	67,22	80,41	67,22
B	100,00	83,88	91,23	83,88
C	100,00	88,33	93,80	88,33
D	100,00	97,22	98,59	97,22

A precisão perfeita em todos os indivíduos é justificada pela ausência de classificações com falsos positivos. Essa é uma consequência do uso de imagens de elementos do rosto no treinamento negativo do classificador, como nariz, testa, e sobrancelha. A revocação de indivíduos com olhos claros se mostrou inferior a de pessoas com olhos escuros. Notou-se que os indivíduos com olhos claros possuem sensibilidade aos raios solares e, por esse motivo, deformam as pálpebras. Uma vez que o classificador não havia sido preenchido com amostras positivas

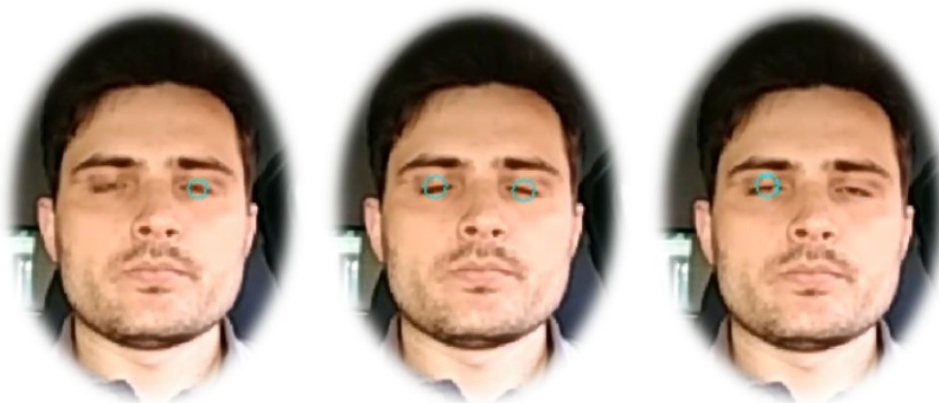
que ilustram essa situação, a classificação dos olhos claros apresentou muitos falsos negativos, e a revocação foi impactada. A acurácia de todos os indivíduos apresentou valor semelhante à revocação, visto que o experimento não possui verdadeiros negativos e falsos positivos.

A quase totalidade das ocorrências de falsos negativos (94,2%) ocorreram somente em um dos olhos. Sendo assim, como o sistema de estimação de sonolência necessita de pelo menos uma amostra positiva classificada corretamente para incrementar o contador de olhos abertos, pode-se inferir que os efeitos dos falsos negativos na avaliação da sonolência são pequenos. Uma amostra em vídeo do sistema de identificação de olhos abertos exhibe os resultados qualitativos em diversas posições e angulações (SOARES; NETO; LIMA, 2018c).

4.6 Detecção de Olhos Fechados

Olhos fechados, ou olhos com pequenas aberturas, são identificados pelo algoritmo *Gente Adaboost*, em conjunção com o classificador de olhos fechados. A operação do procedimento de identificação se limita à região dos olhos, semelhantemente à identificação de olhos abertos, na qual os círculos azuis representam classificações positivas (ver Figura 4.5).

Figura 4.5 – Sequência de quadros que exhibe a classificação de olhos fechados (centro) e de olhos com pequenas aberturas (esquerda e direita).



Fonte: Do autor (2018).

A detecção de olhos fechados é um processo difícil, uma vez que existem características semelhantes às das sobrancelhas que devem ser agrupadas distintivamente. Identificar olhos fechados com pequenas aberturas é mais penoso ainda, uma vez que imagens de olhos totalmente abertos foram usadas no treinamento negativo do classificador e presença da pupila (ou parte dela) pode ser associada a um evento negativo. Sendo assim, o balanceamento entre o desempenho e o custo computacional levou em consideração a eficiência da classificação. Portanto, o

parâmetro de redução de escala foi definido como 4%, em vez de 10% (usado na identificação de olhos abertos). Desse modo, são efetuadas mais operações de buscas (uma vez que a imagem é maior), e a possibilidade de classificação correta é superior.

Os parâmetros *minSize* e *maxSize*, encarregados de demarcar o tamanho das janelas de busca, foram fixados em 20x10 e 30x15 *pixels*, respectivamente. Inicialmente, os classificadores treinados classificavam elementos como boca, sobrancelhas e queixo como positivos. Por esse motivo, amostras específicas desses elementos foram adicionadas ao conjunto de dados de imagens negativas, e o novo classificador gerado conseguiu diminuir as ocorrências de falsos positivos. Esse tratamento específico permitiu que o parâmetro *minNeighbors* seja determinado como 1, isto é, durante a varredura da imagem com classificadores operando em janelas, basta que duas janelas adjacentes sejam positivas para que o classificador classifique essa região como positiva.

A análise quantitativa do classificador de olhos fechados foi feita com o mesmo grupo de indivíduos do classificador de olhos abertos, descrita na Tabela 4.4. Cada indivíduo gravou uma simulação em que a única orientação foi a de pretender que era um ato comum de dirigir. O experimento ocorreu entre 13:10 e 13:30 da tarde, em um dia ensolarado, com o veículo parado, para garantir a segurança dos indivíduos. Os parâmetros verdadeiro positivo, falso positivo, falso negativo e verdadeiro negativo foram extraídos de 55 quadros para o cálculo da precisão, revocação, *F-measure* e acurácia (ver Seção 4.5). Os resultados são exibidos na Tabela 4.6.

Tabela 4.6 – Dados da avaliação do classificador de olhos fechados.

Indivíduo	Precisão	Revocação	<i>F-Measure</i>	Acurácia
A	100,00	72,72	84,20	72,72
B	100,00	70,90	82,87	70,90
C	100,00	78,18	87,75	78,18
D	100,00	83,63	91,08	83,63

Semelhantemente ao classificador de olhos abertos, este não apresentou classificações falsamente positivas. Entretanto, é notável a ocorrência de falsos negativos, principalmente em olhos com pouca abertura que deveriam ser classificados como fechados, o que justificado pela pequena quantidade de imagens adicionadas no conjunto de dados positivo com essa característica. Embora exista essa adversidade, a maioria (86,1%) das ocorrências de falsos negativos aconteceram somente em um dos olhos. Portanto, como o sistema de identificação de sono-lência e adormecimento necessita de pelo menos uma classificação positiva, os resultados dos falsos negativos são atenuados. Uma amostra em vídeo do sistema de identificação de olhos

fechados exibe os resultados qualitativos em diversas posições e angulações (SOARES; NETO; LIMA, 2018b).

4.7 Identificação de Sonolência do Condutor

A identificação de sonolência baseia-se na classificação do estado dos olhos segundo o parâmetro PERCLOS, descrito na Subseção 2.2.2, em que o sintoma de lentidão da abertura dos olhos nas piscadas é detectado. Para tal, o PERCLOS é composto pela contagem das ocorrências de olhos abertos e fechados em uma longa janela de tempo (ver equação 3.18).

A estimação do PERCLOS acontece, aproximadamente, de minuto em minuto. A incerteza se justifica pelo fato de que o intervalo é medido segundo a aquisição de novas imagens, e não de forma temporal. A aplicação possui uma taxa de fps que tange 23,4 (ver Tabela 4.2). Sendo assim, em sessenta segundos são gerados cerca de 1400 quadros. Portanto, o indicador PERCLOS é calculado a cada 1400 quadros. Como a taxa de quadros por segundo é pouco variável, pode-se considerar o intervalo de tempo de um minuto com um pequeno desvio superior ou inferior.

Motoristas sonolentos tendem a permanecer com os olhos fechados por mais tempo durante as piscadas. Portanto, neste evento, espera-se que mais olhos fechados sejam identificados, e que a magnitude do PERCLOS seja elevada. A avaliação deste método baseia-se na simulação da atividade de dirigir em que o condutor se encontra em duas situações: vigilante e sonolento, no qual o termo vigilante é o antônimo de sonolento.

No experimento elaborado, os dados do parâmetro PERCLOS foram coletados no dia em que o condutor acordou 08:30 da manhã, após ter tido 07 horas de sono. A avaliação do condutor vigilante foi feita no intervalo de 10:30 às 11:00 horas da manhã, e a simulação do motorista cansado ocorreu entre 02:00 e 02:30 da manhã, quase 18 horas após este ter acordado. Uma amostra em vídeo do experimento exibe a comparação qualitativa entre o indivíduo vigilante e sonolento (SOARES; NETO; LIMA, 2018a). O principal objetivo desse experimento é analisar os valores do PERCLOS, exibidos nas Tabelas 4.7 e 4.8, e os efeitos que o excesso de tempo sem dormir causam nas piscadas dos olhos.

Os resultados apontam que, em todas simulações, os valores de PERCLOS do condutor vigilante foram inferiores aos do indivíduo sonolento. As amostras mais próximas, representadas pela simulação 2 e 3 do motorista vigilante e sonolento, respectivamente, estão separadas

Tabela 4.7 – Valores de PERCLOS da simulação em que o motorista está vigilante.

Amostra	PERCLOS (vigilante)
1	0,0299
2	0,0312
3	0,0230
4	0,0271
5	0,0211
6	0,0198
7	0,0199
8	0,0242
9	0,0190
10	0,0195
Média	0,0235

Tabela 4.8 – Valores de PERCLOS da simulação em que o motorista está sonolento.

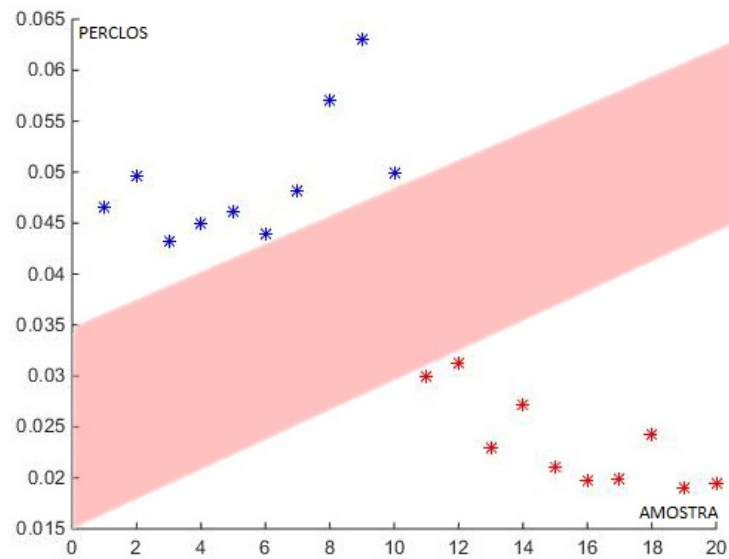
Amostra	PERCLOS (sonolento)
1	0,0466
2	0,0496
3	0,0432
4	0,0450
5	0,0461
6	0,0439
7	0,0482
8	0,0571
9	0,0630
10	0,0499
Média	0,0493

por 0,012. A Figura 4.6 descreve os dados da Tabela 4.7 no formato de um gráfico, onde é possível visualizar a separação dos valores de cada simulação nitidamente.

A discrepância numérica do parâmetro PERCLOS entre o condutor vigilante e sonolento é justificada pela quantidade de quadros em que os olhos fechados são identificados durante uma piscada. Na simulação do condutor vigilante, foi notado que os olhos fechados eram identificados em 2 ou 3 quadros no decorrer de cada piscada (ver Figura 4.7). Entretanto, a simulação do condutor sonolento apontou que a aplicação identifica 5 ou 6 quadros de olhos fechados (ver Figura 4.8). Esse resultado justifica o motivo pelo qual a média do PERCLOS de um condutor sonolento equivale a 209,78% do PERCLOS médio de um vigilante, segundo os dados das Tabelas 4.7 e 4.8.

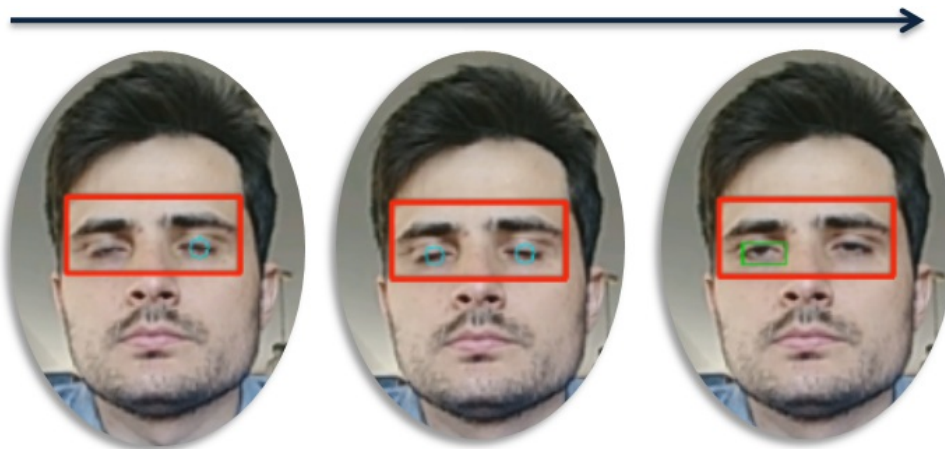
A equação 4.6 descreve a função de emissão de alerta para os eventos em que o condutor apresenta sinais de sonolência. O limiar foi definido como a média da simulação do condutor

Figura 4.6 – Valores de PERCLOS do condutor sonolento (azul) e vigilante (vermelho). A área do paralelogramo rosa representa a separação entre os dados.



Fonte: Do autor (2018).

Figura 4.7 – Sequência de quadros extraída da simulação do condutor vigilante que descreve uma amostra de piscar de olhos.

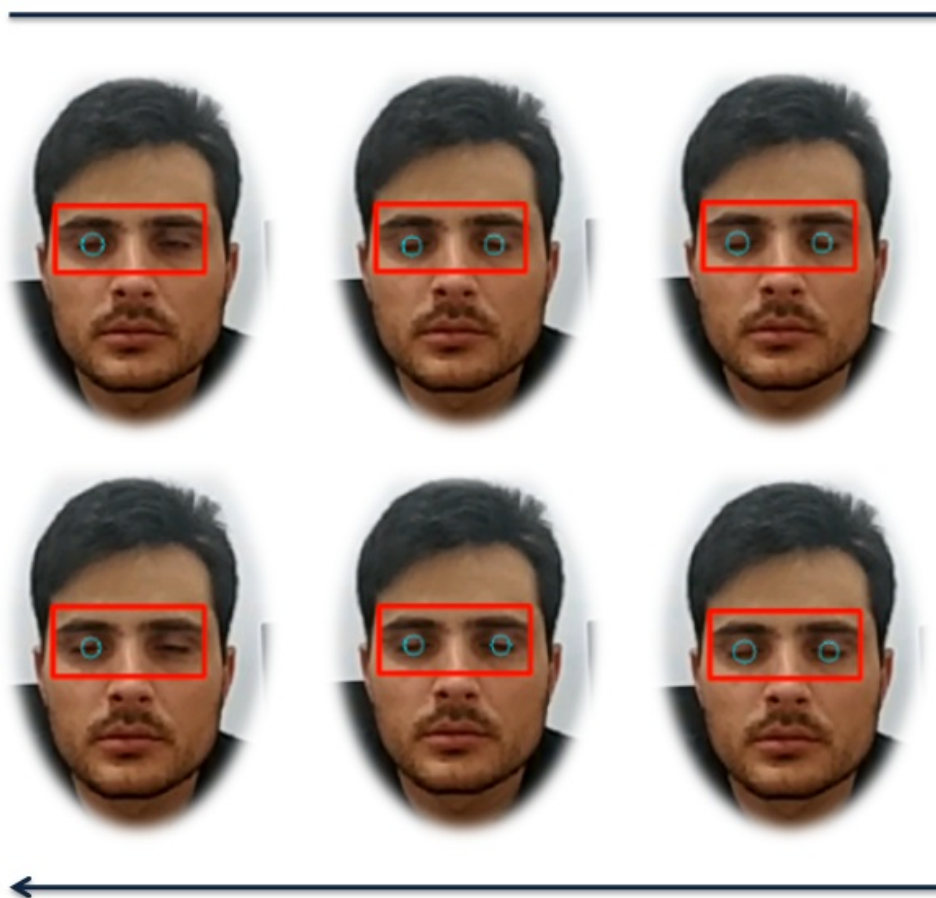


Fonte: Do autor (2018).

sonolento. Nos eventos em que o limiar é ultrapassado, o alerta de sonolência é acionado para orientar o condutor a estacionar o veículo e, se possível, repousar por um tempo.

$$WARNING = \begin{cases} 1, & \text{se } PERCLOS > 0,0493, \\ 0, & \text{senão.} \end{cases} \quad (4.6)$$

Figura 4.8 – Sequência de quadros extraída da simulação do condutor sonolento que descreve uma amostra de piscar de olhos.



Fonte: Do autor (2018).

4.8 Identificação de Adormecimento do Condutor

A identificação de adormecimento é semelhante à detecção de sonolência, no qual os olhos são monitorados e classificados. A principal diferença entre os métodos é o intervalo de estimação e o limiar em que o alerta é acionado. Na Subseção 3.7.2, foi citado que um dos primeiros sintomas de motoristas entrando no estado de adormecimento é o fechamento ininterrupto dos olhos. Nesse evento, a razão do parâmetro FA (ver equação 3.20) tende ao unitário, uma vez que os olhos fechados ocupam a totalidade das classificações. Entretanto, nos cenários onde o motorista começa a adormecer nos instantes medianos e finais da janela de observação, o parâmetro FA considera as classificações dos olhos abertos que ocorreram no início, e o seu valor decresce. Sendo assim, o limiar de adormecimento foi definido como 0,9, segundo na equação 4.7. Ademais, no sentido de atuar velozmente, o intervalo temporal de estimação do adormecimento é de 30 imagens, ou aproximadamente 1,4 segundo.

$$WARNING = \begin{cases} 2, & \text{se } FA > 0,9, \\ 0, & \text{senão.} \end{cases} \quad (4.7)$$

A avaliação quantitativa do identificador de adormecimento se baseia no tempo de espera entre o primeiro quadro em que o condutor fechou os olhos, até o momento em que o alerta foi acionado. Para tal, foi realizado uma simulação composta por 10 amostras de adormecimento, onde em cada uma o motorista manteve os olhos fechados ininterruptamente até que o alerta fosse acionado (SOARES; NETO; LIMA, 2018d). A simulação foi gravada por um programa de registro de tela do *smartphone*. Foi notado que esse *software* ocasionava quedas eventuais de até 3 fps na aplicação. Os eventos de ocorrências dos alarmes e seus períodos foram observados e apresentados na Tabela 4.9.

Tabela 4.9 – Valores de tempo entre o início do adormecimento e a emissão de alerta.

Amostra	Tempo (seg)
1	2,775
2	1,998
3	2,030
4	1,863
5	1,491
6	1,906
7	2,154
8	2,154
9	1,947
10	1,519
Média	2,034

O sistema de identificação de adormecimento foi capaz de alertar o motorista em todas as simulações de adormecimento. O tempo de emissão de alerta variou significativamente, principalmente na primeira simulação, o que é explanado pela análise do intervalo de quadros. Quando o motorista adormece a partir do quarto *frame* da janela, sem ter piscado anteriormente, o valor de FA não alcança o limiar de 0,9. Mas, na próxima janela, se o adormecimento continua, o alerta é acionado no final do trigésimo quadro. Sendo assim, somente depois do 56º quadro, desde o adormecimento do motorista, que o alerta é emitido. Durante a primeira simulação, a aplicação operou flutuando entre 20 e 21 fps. Portanto, o adormecimento a partir do quarto quadro demoraria entre 2,66 e 2,8 segundos para ser identificado, o que esclarece a divergência da primeira simulação. A solução para este problema, deixada para trabalhos futuros, é a implantação de uma janela móvel, no formato de fila, em que os dados entrariam no final

e saíam no início. Dessa forma, a razão FA seria calculada continuamente após o trigésimo quadro, e a posição da janela na qual o motorista adormeceu não influenciaria no tempo de alerta.

Uma amostra em vídeo em que o condutor adormece propositalmente foi gravada (SOARES; NETO; LIMA, 2018e). Nela, a característica da inclinação frontal da cabeça, que é comum em eventos de adormecimento, foi simulada. É notável que os olhos fechados são identificados até nesta situação, na qual a posição da face não está direcionada à câmera.

5 CONCLUSÕES E TRABALHOS FUTUROS

Este trabalho teve por finalidade o desenvolvimento de uma aplicação de monitoramento de condutores encarregada de identificar eventos de sonolência e adormecimento. O uso de *smartphones* para este fim foi uma ideia inovadora, que proporciona uma maior penetrabilidade na sociedade e permite que condutores de diversas camadas econômicas possam guiar seus veículos com mais segurança. O método proposto usa imagens capturadas pela câmera do dispositivo móvel para identificar a face, segmentá-la na região dos olhos, rastrear essa região para gerenciar eventos de oclusão e classificar o estado dos olhos em aberto ou fechado.

A identificação da face, dos olhos abertos e fechados é feita por classificadores treinados com descritores do tipo LBP. Embora esses classificadores sejam caracterizados pela leveza, sendo indicados em aplicações móveis, a limitação da área de busca se mostrou extremamente importante para diminuir o número de operações executadas pelo algoritmo e, consequentemente, elevar o fps da aplicação. Sem a limitação não seria possível atingir uma taxa de fps aceitável para classificar os olhos em quadros suficientes.

A criação dos classificadores requisitou análises sistemáticas sobre os dados fornecidos durante a fase de treinamento. Nos primeiros testes, o número de falsos positivos era grande, principalmente em elementos da face. Gradualmente, foram adicionadas informações específicas no conjunto de dados referente às amostras negativas, como imagens de bocas, queixos, sobancelhas, bochechas e narizes, até que um resultado satisfatório fosse obtido. Os falsos negativos foram tratados por meio do aumento da variabilidade de intensidade de brilho na imagem. Especificamente no classificador de olhos fechados, foi necessário modificar todo o conjunto de dados de amostras positivas, uma vez que a qualidade das imagens não era conveniente para que o treinamento apresentasse bons resultados. As modificações feitas após as análises foram essenciais para que os classificadores alcançassem um bom desempenho.

O sistema de avaliação da sonolência foi analisado quantitativamente por meio de um experimento com um indivíduo nos estados de vigilante e sonolento. O sintoma da sonolência responsável por diminuir a velocidade de abertura dos olhos durante as piscadas foi identificado, e foi possível elaborar uma função limiar que separasse os níveis de PERCLOS dos dois estados. Sendo assim, a emissão de alerta somente é realizada em eventos onde a sonolência do condutor é identificada.

O método proposto para a identificação do adormecimento de condutores efetua a estimação do parâmetro FA a cada 30 quadros. Essa janela fixa se mostrou ineficiente, uma vez que

o tempo de alerta emitido ao motorista apresentou variação significativa. Propõe-se, então, para trabalhos futuros, a implantação de uma janela móvel, no formato de fila, onde a amostra de entrada é inserida no último lugar, e o cálculo de FA é efetuado em todos os quadros de forma contínua.

A aplicação desenvolvida mostrou-se robusta em ambientes diurnos, mas a limitação na operação em baixas iluminações é uma enorme adversidade. Uma vez que grande parte das câmeras dos *smartphones* não possuem a tecnologia infravermelho, as características dos condutores dificilmente são reconhecidas. Como solução, pode-se elaborar um circuito com LEDs infravermelhos e direcioná-lo para a face do motorista, enquanto as câmeras de telefones móveis não implementem essa tecnologia.

Os classificadores foram treinados para atender situações de faces e olhos frontais e com leves angulações. Por esse motivo, o suporte anexado ao para-brisa do veículo é situado frontalmente ao volante. Para aumentar o ambiente de operação da aplicação, pode-se treinar classificadores encarregados de identificar faces e olhos com angulações, e executá-los paralelamente aos classificadores frontais. Desse modo, a restrição do local de instalação do suporte seria mínima e o condutor teria mais liberdade para posicionar o seu *smartphone*.

Os alertas emitidos pela aplicação devem ser os mais precisos possíveis. No contexto atual, eles são emitidos independentemente da velocidade do veículo. Entretanto, em situações de congestionamento, onde o veículo está parado, a emissão de alerta é dispensável. Nesse sentido, pode-se desenvolver um sistema de estimação da velocidade do veículo por meio dos sensores inerciais do *smartphone*, para que a emissão de alerta siga um padrão definido. Pode-se, também, aprimorar o alerta enviado para motoristas adormecidos. Em velocidades elevadas, deve-se enviar o alerta rapidamente, mas em velocidades reduzidas pode-se considerar um intervalo mais longo, uma vez que o condutor pode estar descansando conscientemente com os olhos fechados.

REFERÊNCIAS

- ACIOGLU, A.; ERÇELEBI, E. Real time eye detection algorithm for perclos calculation. In: **2016 24th Signal Processing and Communication Application Conference (SIU)**. [S.l.: s.n.], 2016. p. 1641–1644.
- AKBAR, I. A. et al. Three drowsiness categories assessment by electroencephalogram in driving simulator environment. In: IEEE. **Engineering in Medicine and Biology Society (EMBC), 2017 39th Annual International Conference of the IEEE**. [S.l.], 2017. p. 2904–2907.
- BENGLER, K. et al. Three decades of driver assistance systems: Review and future perspectives. **IEEE Intelligent Transportation Systems Magazine**, IEEE, v. 6, n. 4, p. 6–22, 2014.
- BERGASA, L. M. et al. Real-time system for monitoring driver vigilance. **IEEE Transactions on Intelligent Transportation Systems**, IEEE, v. 7, n. 1, p. 63–77, 2006.
- BRASIL. Estimativa dos custos dos acidentes de trânsito no brasil com base na atualização simplificada das pesquisas anteriores do ipea. In: INSTITUTO DE PESQUISA ECONÔMICA APLICADA. [S.l.], 2015.
- BRASIL. Dados de trânsito: acidentes. In: POLÍCIA RODOVIÁRIA FEDERAL. [S.l.], 2017.
- CAFFIER, P. P.; ERDMANN, U.; ULLSPERGER, P. Experimental evaluation of eye-blink parameters as a drowsiness measure. **European journal of applied physiology**, Springer, v. 89, n. 3-4, p. 319–325, 2003.
- CHEHREHGOSHA, A.; EMADI, M. Face detection using fusion of lbp and adaboost. **Journal of Soft Computing and Applications**, v. 2016, n. 1, p. 1–10, 2016.
- CHENG, X. et al. A real-time driving fatigue monitoring dsp device based on computing complexity of binarized image. In: IEEE. **Computer Science and Engineering, 2009. WCSE'09. Second International Workshop on**. [S.l.], 2009. v. 1, p. 84–88.
- CORREA, A. G.; OROSCO, L.; LACIAR, E. Automatic detection of drowsiness in eeg records based on multimodal analysis. **Medical engineering & physics**, Elsevier, v. 36, n. 2, p. 244–249, 2014.
- DASGUPTA, A. et al. An on-board vision based system for drowsiness detection in automotive drivers. **International Journal of Advances in Engineering Sciences and Applied Mathematics**, Springer, v. 5, n. 2-3, p. 94–103, 2013.
- DASGUPTA, A. et al. A vision-based system for monitoring the loss of attention in automotive drivers. **IEEE Transactions on Intelligent Transportation Systems**, IEEE, v. 14, n. 4, p. 1825–1838, 2013.
- DINGES, D. F.; GRACE, R. Perclos: A valid psychophysiological measure of alertness as assessed by psychomotor vigilance. **US Department of Transportation, Federal Highway Administration, Publication Number FHWA-MCRT-98-006**, 1998.
- FRIEDMAN, J. et al. Additive logistic regression: a statistical view of boosting. **The annals of statistics**, Institute of Mathematical Statistics, v. 28, n. 2, p. 337–407, 2000.

GADDIGOUDAR, P. K. et al. Pedestrian detection and tracking using particle filtering. In: **2017 International Conference on Computing, Communication and Automation (ICCCA)**. [S.l.: s.n.], 2017. p. 110–115.

GAO, W. et al. The cas-peal large-scale chinese face database and baseline evaluations. **IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans**, IEEE, v. 38, n. 1, p. 149–161, 2008.

GARCÍA, D. A. N. **Contribución a la autolocalización de robots móviles basada en la fusión de información multisensorial**. Tese (Doutorado), 2009.

GERONIMO, D. et al. Survey of pedestrian detection for advanced driver assistance systems. **IEEE transactions on pattern analysis and machine intelligence**, IEEE, v. 32, n. 7, p. 1239–1258, 2010.

GONG, D. H.; KWAK, K. C. Face detection and status analysis algorithms in day and night environments. In: IEEE. **Advanced Informatics, Concepts, Theory, and Applications (ICAICTA), 2017 International Conference on**. [S.l.], 2017. p. 1–4.

GRACE, R. et al. A drowsy driver detection system for heavy vehicles. In: IEEE. **Digital Avionics Systems Conference, 1998. Proceedings., 17th DASC. The AIAA/IEEE/SAE**. [S.l.], 1998. v. 2, p. I36–1.

GU, H.; JI, Q.; ZHU, Z. Active facial tracking for fatigue detection. In: IEEE. **Applications of Computer Vision, 2002.(WACV 2002). Proceedings. Sixth IEEE Workshop on**. [S.l.], 2002. p. 137–142.

GUPTA, N. et al. Mobile ecg-based drowsiness detection. In: **2017 14th IEEE Annual Consumer Communications Networking Conference (CCNC)**. [S.l.: s.n.], 2017. p. 29–32.

HOU, Y. L. et al. Multispectral pedestrian detection based on deep convolutional neural networks. In: **2017 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC)**. [S.l.: s.n.], 2017. p. 1–4.

IDC. **Estudo da IDC Brasil aponta crescimento de 25,4% nas vendas de celulares no primeiro trimestre de 2017**. 2017. Disponível em: <<https://goo.gl/Pfo4RD>>. Acesso em: 10 jan. 2018.

JUNG, S. J.; SHIN, H. S.; CHUNG, W. Y. Driver fatigue and drowsiness monitoring system with embedded electrocardiogram sensor on steering wheel. **IET Intelligent Transport Systems**, IET, v. 8, n. 1, p. 43–50, 2014.

KONG, W. et al. A system of driving fatigue detection based on machine vision and its application on smart device. **Journal of Sensors**, Hindawi Publishing Corporation, v. 2015, 2015.

KURT, M. B. et al. The ann-based computing of drowsy level. **Expert Systems with Applications**, Elsevier, v. 36, n. 2, p. 2534–2542, 2009.

LEE, D. K. et al. Real-time lane detection and tracking system using simple filter and kalman filter. In: IEEE. **Ubiquitous and Future Networks (ICUFN), 2017 Ninth International Conference on**. [S.l.], 2017. p. 275–277.

- LEE, J.; KIM, J.; SHIN, M. Correlation analysis between electrocardiography (ecg) and photoplethysmogram (ppg) data for driver's drowsiness detection using noise replacement method. **Procedia Computer Science**, Elsevier, v. 116, p. 421–426, 2017.
- LIANG, Y. et al. Prediction of drowsiness events in night shift workers during morning driving. **Accident Analysis & Prevention**, Elsevier, 2017.
- LIN, F. C. et al. Generalized eeg-based drowsiness prediction system by using a self-organizing neural fuzzy system. **IEEE Transactions on Circuits and Systems I: Regular Papers**, IEEE, v. 59, n. 9, p. 2044–2055, 2012.
- LINO, A. F. S. et al. Performance of haar and lbp features in cascade classifiers to whiteflies detection and counting. In: **2017 CHILEAN Conference on Electrical, Electronics Engineering, Information and Communication Technologies (CHILECON)**. [S.l.: s.n.], 2017. p. 1–6.
- MAITA, F. et al. Integrated steering wheel system based on nanostructured elastomeric sensors for real time detection of driver drowsiness status. In: IEEE. **2017 IEEE SENSORS**. [S.l.], 2017. p. 1–3.
- NESHOV, N.; MANOLOVA, A. Drowsiness monitoring in real-time based on supervised descent method. In: IEEE. **Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS), 2017 9th IEEE International Conference on**. [S.l.], 2017. v. 2, p. 660–663.
- OJALA, T.; PIETIKÄINEN, M.; HARWOOD, D. A comparative study of texture measures with classification based on featured distributions. **Pattern recognition**, Elsevier, v. 29, n. 1, p. 51–59, 1996.
- PICOT, A.; CAPLIER, A.; CHARBONNIER, S. Comparison between eeg and high frame rate camera for drowsiness detection. In: IEEE. **Applications of Computer Vision (WACV), 2009 Workshop on**. [S.l.], 2009. p. 1–6.
- PIOTROWSKI, Z.; SZYPULSKA, M. Classification of falling asleep states using hrv analysis. **Biocybernetics and Biomedical Engineering**, Elsevier, 2017.
- PRATIARSO, A. et al. Implementation of kalman filter method for tracking position of alzheimer's patients. In: **2017 International Conference on Smart Cities, Automation Intelligent Computing Systems (ICON-SONICS)**. [S.l.: s.n.], 2017. p. 135–140.
- REDDY, B. et al. Real-time driver drowsiness detection for embedded system using model compression of deep neural networks. In: IEEE. **Computer Vision and Pattern Recognition Workshops (CVPRW), 2017 IEEE Conference on**. [S.l.], 2017. p. 438–445.
- ROMERA, E.; BERGASA, L. M.; ARROYO, R. A real-time multi-scale vehicle detection and tracking approach for smartphones. In: IEEE. **Intelligent Transportation Systems (ITSC), 2015 IEEE 18th International Conference on**. [S.l.], 2015. p. 1298–1303.
- SILVEIRA, T. L. da; KOZAKEVICIUS, A. J.; RODRIGUES, C. R. Automated drowsiness detection through wavelet packet analysis of a single eeg channel. **Expert Systems with Applications**, v. 55, n. Supplement C, p. 559 – 565, 2016. ISSN 0957-4174. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0957417416300744>>.

SMITH, B. A. et al. Gaze locking: passive eye contact detection for human-object interaction. In: ACM. **Proceedings of the 26th annual ACM symposium on User interface software and technology**. [S.l.], 2013. p. 271–280.

SOARES, G. C.; NETO, A. de M.; LIMA, D. A. de. **Amostra do vídeo usado para o cálculo do PERCLOS**. 2018. Disponível em: <https://youtu.be/McyuT6x_Cxo>. Acesso em: 12 fev. 2018.

SOARES, G. C.; NETO, A. de M.; LIMA, D. A. de. **Closed-Eye LBP Classifier**. 2018. Disponível em: <https://youtu.be/mFfW_5RCpVk>. Acesso em: 12 fev. 2018.

SOARES, G. C.; NETO, A. de M.; LIMA, D. A. de. **Open-Eye LBP Classifier**. 2018. Disponível em: <<https://youtu.be/aMmSOEXwiUU>>. Acesso em: 12 fev. 2018.

SOARES, G. C.; NETO, A. de M.; LIMA, D. A. de. **Sleepiness Detection Application: first simulation**. 2018. Disponível em: <<https://youtu.be/Sg3YCTbGtGY>>. Acesso em: 12 fev. 2018.

SOARES, G. C.; NETO, A. de M.; LIMA, D. A. de. **Sleepiness Detection Application: second simulation**. 2018. Disponível em: <<https://youtu.be/08fLYYiGyyo>>. Acesso em: 12 fev. 2018.

TABAL, K. M.; CRUZ, J. D. Development of low-cost embedded-based electrooculogram blink pulse classifier for drowsiness detection system. In: IEEE. **Signal Processing & its Applications (CSPA), 2017 IEEE 13th International Colloquium on**. [S.l.], 2017. p. 29–34.

TAKEI, Y.; FURUKAWA, Y. Estimate of driver's fatigue through steering motion. In: IEEE. **2005 IEEE international conference on systems, man and cybernetics**. [S.l.], 2005. v. 2, p. 1765–1770.

TAN, D.; CHEN, W.; WANG, H. On the use of monte-carlo simulation and deep fourier neural network in lane departure warning. **IEEE Intelligent Transportation Systems Magazine**, IEEE, v. 9, n. 4, p. 76–90, 2017.

UKE, N. J.; THOOL, D. R.; DHOTRE, P. S. Drowsiness detection – a visual system driver support. **International Journal of Electronics Communication and Computer Engineering**, v. 3, p. 29–33, 2012.

VELEZ, G. et al. A reconfigurable embedded vision system for advanced driver assistance. **Journal of Real-Time Image Processing**, Springer, v. 10, n. 4, p. 725–739, 2015.

VIOLA, P.; JONES, M. J. Robust real-time face detection. **International journal of computer vision**, Springer, v. 57, n. 2, p. 137–154, 2004.

VITABILE, S.; PAOLA, A. D.; SORBELLO, F. A real-time non-intrusive fpga-based drowsiness detection system. **Journal of Ambient Intelligence and Humanized Computing**, Springer, v. 2, n. 4, p. 251, 2011.

WANG, L.; PEI, Y. The impact of continuous driving time and rest time on commercial drivers' driving performance and recovery. **Journal of safety research**, Elsevier, v. 50, p. 11–15, 2014.

WANG, X.; XU, C. Driver drowsiness detection based on non-intrusive metrics considering individual specifics. **Accident Analysis & Prevention**, Elsevier, v. 95, p. 350–357, 2016.

WIERWILLE, W. W. et al. **Research on vehicle-based driver status/performance monitoring; development, validation, and refinement of algorithms for detection of driver drowsiness. final report.** [S.l.], 1994.

WOOD, E. et al. Rendering of eyes for eye-shape registration and gaze estimation. In: **Proc. of the IEEE International Conference on Computer Vision (ICCV 2015)**. [S.l.: s.n.], 2015.

WU, S.; NAGAHASHI, H. Penalized adaboost: improving the generalization error of gentle adaboost through a margin distribution. **IEICE TRANSACTIONS on Information and Systems**, The Institute of Electronics, Information and Communication Engineers, v. 98, n. 11, p. 1906–1915, 2015.

YAN, J. J. et al. Real-time driver drowsiness detection system based on perclos and grayscale image processing. In: IEEE. **Computer, Consumer and Control (IS3C), 2016 International Symposium on**. [S.l.], 2016. p. 243–246.

ZHANG, Y.; HUA, C. Driver fatigue recognition based on facial expression analysis using local binary patterns. **Optik-International Journal for Light and Electron Optics**, Elsevier, v. 126, n. 23, p. 4501–4505, 2015.

ZHANG, Z.; ZHANG, J. A new real-time eye tracking based on nonlinear unscented kalman filter for monitoring driver fatigue. **Journal of Control Theory and Applications**, Springer, v. 8, n. 2, p. 181–188, 2010.

ZHU, X. et al. Eog-based drowsiness detection using convolutional neural networks. In: **2014 International Joint Conference on Neural Networks (IJCNN)**. [S.l.: s.n.], 2014. p. 128–134. ISSN 2161-4393.