

Detecção Automática dos Momentos Relevantes em Vídeos de Vigilância Veicular

Laionara de Araújo Gonçalves
Especialização em Computação Distribuída e Ubíqua
Grupo de Sistemas Distribuídos,
Otimização, Redes e Tempo-Real – IFBA
Salvador, Bahia
Email: laionara@ifba.edu.br

Manoel Carvalho Neto
Especialização em Computação Distribuída e Ubíqua
Grupo de Sistemas Distribuídos,
Otimização, Redes e Tempo-Real – IFBA
Salvador, Bahia
Email: manoel@ifba.edu.br

Abstract—O objetivo deste projeto é apresentar um sistema capaz de detectar os momentos relevantes de vídeos de vigilância veicular. Para este trabalho, os momentos relevantes são definidos como os primeiros instantes antes e após a ocorrência de eventos previamente definidos no sistema. Além disso, são destacadas, das imagens consideradas relevantes, placas veiculares e face humana, caso estejam contidas nas mesmas. Esse projeto faz-se necessário devido ao tempo gastos para analisar vídeos que possuem muitas horas de duração. Acredita-se que com a utilização desse sistema seja possível reduzir o trabalho de um especialista na análise de vídeos. Além disso, por se tratar de vídeos de vigilância, espera-se que os dados retirados dos principais momentos auxiliem no esclarecimento de ocorrências em crimes e acidentes.

Palavras-Chave: detectar, veículo, vídeo, melhores, momentos, vigilância, face, placas

I. INTRODUÇÃO

É cada vez mais frequente, em diversos cenários das grandes cidades, a utilização de sistemas de vídeo de vigilância com a finalidade de inibir ações não desejadas, como: crimes, delitos e contravenções. Segundo [22], proteção e segurança tornaram-se essenciais em muitas áreas, especialmente nas áreas públicas, e há uma necessidade de possibilitar aos operadores o monitoramento remoto dessas áreas. São elas: sistemas de transporte (transporte ferroviário, aeroportos, redes rodoviárias, e transporte marítimo), os bancos, shopping centers, parques de estacionamento, e edifícios públicos, os ambientes industriais, e os estabelecimentos governamentais (bases militares, prisões, centros de radar, e hospitais). Essa necessidade, somada aos avanços em desenvolvimento de hardware de computação e sensores eficientes são fatores que contribuem para a criação de complexos sistemas de vigilância.

A utilização desses dispositivos eletrônicos para vídeo de vigilância vem crescendo nos últimos anos. Um dos fatores que proporciona esse aumento é o avanço de tecnologias para captura de imagens que fazem com que equipamentos sofisticados sejam, cada vez mais, acessíveis para pessoas e organizações. Com o barateamento desses dispositivos e com a crescente necessidade de garantir a segurança, administrar incidentes e obter dados para fazer frente a reclamações de responsabilidade civil, sistemas de vídeo de vigilância estão se tornando pervasivos na sociedade.

Para este projeto interessa, principalmente, a utilização de sistemas de vídeo monitoramento em veículos de trans-

porte terrestre, por exemplo: carros, motocicletas, ônibus, caminhões, entre outros. No entanto, a grande maioria dos consumidores de sistemas de vídeo de vigilância em veículos são empresas de transporte público, como ônibus e metrô, e além destes, viaturas da polícia, ambulâncias e carros-fortes.

Além de manter a segurança, os sistemas de vídeo de vigilância oferecem a possibilidade de analisar os eventos após a sua ocorrência. Entretanto, analisar vídeos de monitoramento que possuem muitas horas de duração demanda, normalmente, um esforço (tempo) muito grande. Este projeto apresentará um sistema que permite detectar os principais momentos em vídeos de vigilância veicular. O objetivo deste sistema é contribuir para redução dos esforços demandados para analisar vídeos de vigilância e, além disso, extrair dados como placas veiculares e face humana que podem ser úteis para um especialista neste domínio.

O monitoramento do veículo se deu pelo uso de um protótipo, desenvolvido neste projeto, com câmera, sensor ultrassônico de distância e GPS acoplados. O sistema indica o período de tempo o qual foi identificado um evento que indique um momento relevante. Em seguida, o conteúdo do momento relevante é processado para que possa ser possível detectar as placas de identificação e reconhecer de pessoas (faces). São considerados momentos relevantes em um vídeo de vigilância veicular, a ocorrência, em qualquer quantidade, de qualquer um dos eventos listados seguir:

- A identificação de algum obstáculo a frente do veículo feita através do uso de sensores de distância. É importante ressaltar que, para este evento, considera-se como condição de ocorrência a velocidade na qual um veículo se aproxima de um obstáculo.
- Alterações bruscas de velocidade do veículo. Por exemplo, no caso de uma colisão séria existe sempre uma alteração brusca de velocidade.
- Detecção e identificação da face do motorista. Por exemplo, Uma identificação de um motorista diferente do pré-cadastrado dispararia um evento e consequentemente uma marcação no vídeo.

Para os sistemas de vídeo de vigilância faz-se necessário o estudo de teorias e tecnologias para possibilitar a interpretação de imagens. A área de pesquisa em que se realizam esses estudos é a Visão Computacional, onde um dos problemas

básicos é a detecção e reconhecimento de objetos em imagens. A detecção consiste em determinar se um dado tipo de objeto está presente em uma imagem, enquanto que o reconhecimento consiste em relacionar algum dos objetos detectados a um dado exemplar daquele tipo de objeto [24]. Para este trabalho são utilizadas as técnicas para a detecção de placas e para o reconhecimento de pessoas em imagens, além de sensores de distância e velocidade.

A iniciativa deste trabalho foi inspirada na importância dos sistemas de vídeo vigilância para sociedade, especialmente para prevenção/elucidação de crimes, acidentes e fiscalização de veículos no trânsito. Espera-se com os estudos aqui desenvolvidos, contribuir para a redução de tempo e custos em análises de vídeos de vigilância veicular além de fornecer dados que possam auxiliar no esclarecimento de ocorrências em crimes e acidentes.

A hipótese do uso deste sistema como ferramenta de apoio depende da boa qualidade das imagens dos vídeos. Existem também fatores como luminosidade, oscilações na pista dentre outros que podem tornar difícil a extração dos dados. Além disso, ainda há limitações impostas pelo método utilizado para o processamento das imagens. Cada método possui um nível de eficiência específico e portanto a eficiência do sistema depende do método escolhido. Outro fator que deve ser levado em consideração é a velocidade de processamento do computador. O sistema deve ser ajustado para trabalhar com uma velocidade de processamento equivalente a velocidade de gravação da câmera para possibilitar a detecção em tempo real [4].

Conhecer e entender quais são os estudos mais importantes nesta área é um ponto fundamental para encontrar e reduzir os problemas mencionados anteriormente. Neste contexto, este texto apresenta um mapeamento do estado da arte da área de detecção de eventos em vídeos a partir de uma revisão sistemática da literatura (*Systematic Literature Review - SLR*). Esta revisão sistemática é apresentada na Seção II. A seção seguinte refere-se a um estudo de caso desenvolvido com o intuito de identificar o perfil dos possíveis usuários da solução apresentada neste projeto. O próximo tópico aborda sobre as tecnologias utilizadas e a arquitetura escolhida. A Seção IV refere-se a etapa de implementação do protótipo desenvolvido, indicando o algoritmo criado juntamente com a biblioteca OpenCV e os testes aplicados ao mesmo. Enquanto que na Seção V são expostas as limitações desse projeto e as condições ideais para a aplicação do mesmo. Por fim, tem-se os trabalhos que podem ser desenvolvidos no futuro e as conclusões.

II. REVISÃO SISTEMÁTICA

Para o desenvolvimento deste projeto foi necessário a realização de uma revisão sistemática da literatura. Essa última se deu posteriormente a realização de um Mapeamento Sistemático da Literatura, cuja finalidade é encontrar e classificar estudos referentes a vídeos de vigilância veicular e processamento de imagem.

No Mapeamento Sistemático da Literatura, questões simples e abrangentes, sobre o tema de pesquisa, são respondidas para identificar os estudos disponíveis. Apesar do método de pesquisa e extração de dados ser o mesmo realizado em revisão sistemática, as questões de pesquisa são mais amplas

e em maior número. Além disso, os dados são extraídos com foco na classificação, sem despender muito tempo, e a análise dos dados é feita, principalmente, por sumarização e uso de gráficos.

Já a Revisão Sistemática da Literatura visa identificar, avaliar e interpretar todos os estudos disponíveis e relevantes para uma questão de pesquisa específica. A busca pelos estudos é feita por termos altamente focados ao tema, assim como as questões de pesquisa. Consequentemente, é necessário mais tempo para uma análise cuidadosa dos estudos encontrados, no entanto, após a revisão, os resultados obtidos são mais conclusivos para o projeto.

Espera-se com este Mapeamento Sistemático da Literatura obter as respostas para as seguintes questões de pesquisa:

- 1) Quais métodos de detecção de placas de veículos em imagens são mais utilizados atualmente?
- 2) Quais métodos de detecção e reconhecimento de pessoas em imagens são mais utilizados atualmente?
- 3) Quais os principais desafios em detecção de objetos em imagens?
- 4) Quais tecnologias podem ser utilizadas no desenvolvimento deste projeto?

A partir dessas questões foi definida a estratégia de pesquisa de estudos. Inicialmente foram definidos três grupos de pesquisa denominados A, B e C. Em cada grupo buscou-se artigos produzidos entre os anos de 2000 e 2014 e que apresentem, em qualquer local do texto, pelo menos uma das seguintes expressões:

- Grupo A: “*detection of highlights in video*” OR “*video surveillance vehicle*” OR “*video of vehicle monitoring*” OR “*video vehicle tracking*”.
- Grupo B: “*detecção de face humana*” OR “*detection of human face in video*” OR “*reconhecimento de face*” OR “*face recognition in video*”.
- Grupo C: “*detection of vehicle plates*” OR “*vehicle plate location*”.

As fontes de pesquisa dos estudos selecionados foram:

- ACM Portal (<http://dl.acm.org/>),
- IEEE Xplore (<http://ieeexplore.ieee.org/>),
- Google Scholar (<http://scholar.google.com.br>).

Estas bibliotecas digitais foram escolhidas pois elas:

- 1) possuem motores de busca que permitem o uso de expressões lógicas ou algum mecanismo equivalente,
- 2) incluem publicações em ciência da computação ou em tópicos relacionados que possuem uma relação com as questões da pesquisa, e
- 3) são acessíveis a partir da rede de computadores da instituição onde estão os autores destes trabalhos.

Estas bibliotecas digitais são usadas comumente em mapeamentos sistemáticos na área de ciência da computação. É importante ressaltar que, nem todas elas têm as mesmas funcionalidades. Assim, foi necessário fazer algumas adaptações nas buscas em cada uma delas. Especialmente o *Google*

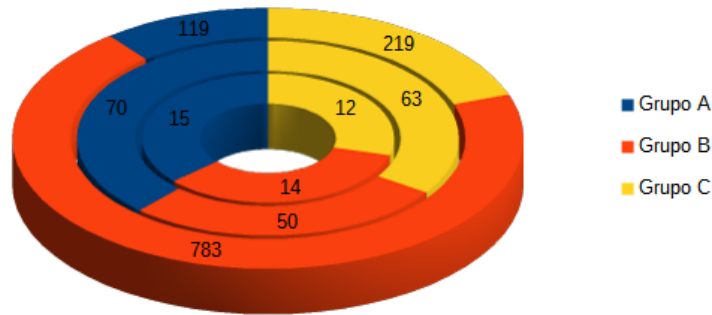


Figura 1. Seleção de Estudos do mapeamento sistemático

Scholar, apesar de não ser uma plataforma de publicação, foi escolhido pois, apresenta resultados de outras fontes não mencionadas aqui como por exemplo ERBASE, Workshop de Visão Computacional entre outras.

A Figura 1 apresenta a distribuição dos artigos selecionados na revisão sistemática. Os números iniciais dos trabalhos obtidos pela busca foram: Grupo A - 119 , Grupo B - 783 e Grupo C - 219. Depois disso, estes resultados foram filtrados a partir da análise dos títulos e/ou resumos de cada um dos trabalhos, além de considerar o número de citações e publicações. Este processo de filtragem reduziu a quantidade de material para a leitura completa e, no final, os grupos ficaram com a seguinte configuração: Grupo A - 15 , Grupo B - 12 e Grupo C - 14. Esta seleção é representada pela Figura 2.

É importante ressaltar que os critérios de seleção, inclusão e exclusão dos artigos estão relacionados às questões de pesquisa definidas anteriormente. Outra observação importante é que o filtro aplicado sobre os resultados iniciais da pesquisa foi baseado apenas nos critérios de exclusão.

A. Resultados

Inicialmente, foram extraídas dos artigos obtidos nas buscas, as seguintes informações: nome do autor, título, ano, publicações, quantidade de citações, métodos e tecnologias utilizadas. Esses dados foram sintetizados e serão apresentados a seguir:

- **Número de artigos por ano:** A Figura 3 apresenta a distribuição dos números de artigos de cada grupo ordenado pelo ano de publicação. Constatou-se que os artigos do Grupo A foram encontrados em quase todos os anos de publicação considerados na pesquisa (de 2000 até 2014), exceto no ano 2000, e em maior quantidade no ano de 2012. Já os artigos do Grupo B foram encontrados em maior quantidade no ano de 2009. Os artigos pertencentes ao Grupo C foram publicados, em sua maioria, no ano de 2010.
- **Métodos de Detecção de objetos em Imagens**
A Tabela I apresenta os alguns métodos para detecção de objetos em imagens e seus respectivos níveis de precisão, de acordo com o estudo realizado em [13].
- **Tema dos estudos analisados**

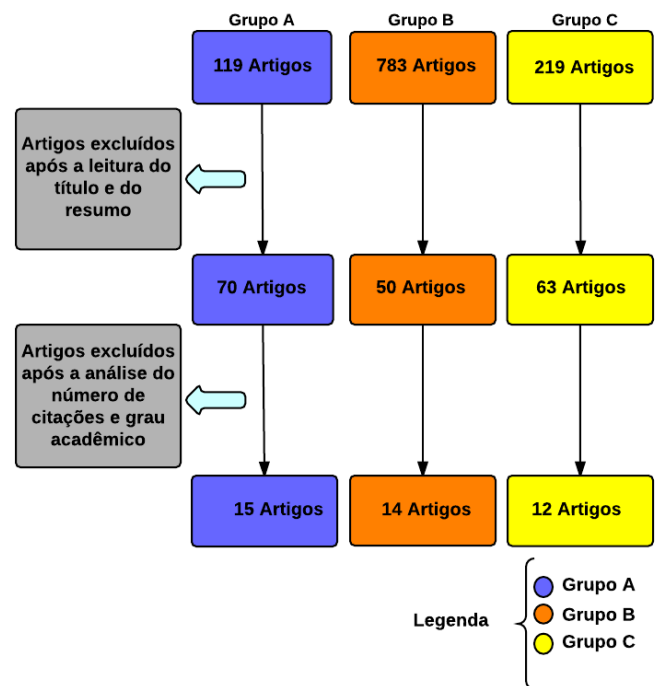


Figura 2. Etapas de seleção dos artigos

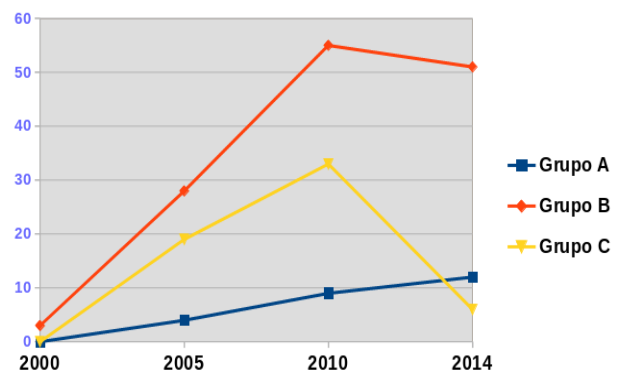


Figura 3. Análise da quantidade de artigos selecionados em relação aos anos definidos na pesquisa

TABELA I. MÉTODOS DE DETECÇÃO DE OBJETOS EM IMAGENS

Método	Categoria	Precisão
AdaBoost	Classificador	95,6%
SVM	Classificador	91,3%
Transf. Wavelet	Proc. de Tons de Cinza	92,4%
Operador de Sobel	Proc. de Imagens Binárias	97%
Filtro Gabor	Proc. de Tons de Cinza	98%

A Tabela II apresenta as características encontradas nos estudos vistos nessa revisão sistemática. Foi possível concluir que a maioria dos trabalhos sobre vídeo vigilância estão relacionados ao estudo ou proposta de algoritmos, poucos estudos visam propor novas aplicações para vídeo vigilância. E ainda dentre os sistemas apresentados, um número pequeno possibilita a interação com usuário.

Com relação aos artigos que abordam a detecção de momentos relevantes em vídeo, conclui-se que a maioria deles estão relacionados a vídeos de programas de esporte. Nesses trabalhos estudou-se diversas formas de detecção do momento importante do vídeo, desde a análise do áudio [30], [28], [27], cores [26], até imagens do gramado e textos nas imagens [28].

B. Referencial Teórico

Nesta seção são apresentados os aspectos mais importantes dos principais artigos lidos nesta revisão sistemática. A partir destas leituras foi possível obter a fundamentação necessária para o desenvolvimento deste trabalho.

Os artigos coletados nas buscas realizadas com as expressões do "Grupo A", supracitado nesta revisão, estão relacionados a detecção de destaques em vídeos de esportes [26], [27], [28], [29], [11], [30] e a sistemas de vídeo vigilância [22], [23], [21].

Em [29] é descrito um método estatístico para detecção de destaques em vídeos de jogos de beisebol. Nessa abordagem, o vídeo é segmentado em tiros de cenas que são classificados em diferentes tipos. Essa abordagem é baseada em dois critérios de observação, o primeiro de que os destaque em um vídeo geralmente pertencem ao mesmo tipo de tiro de cena, e o segundo, de que esses tiros de cena possuem contextos especiais em transição no tempo.

Enquanto que em [30], é apresentado um algoritmo para identificar momentos importantes em jogos de beisebol a partir do áudio do vídeo. Para a identificação do destaque foi definido que: segmentos emocionantes estão altamente correlacionados com o discurso animado dos locutores, e que os segmentos mais emocionantes nos vídeos de basebol ocorrem logo após um arremesso. O algoritmo deve diferenciar a voz do narrador

dos outros sons no vídeo, e identificar os momentos de animação na voz desse narrador. Por fim, esses momentos são fundidos e apresentados como destaque.

Outro contexto em que pode ser útil a obtenção de melhores momentos é o de vídeos de jogos de tênis, que é o alvo do estudo realizado em [11]. No entanto o autor esclarece que essa metodologia pode ser útil também em outros tipos de vídeos. Este estudo propõe o uso de um conjunto combinado de estratégias inteligentes para encontrar automaticamente os destaques nas transmissões de tênis.

O objetivo do trabalho desenvolvido em [28] é apresentar um método multimídia para a extração de melhores momentos em vídeos de programas de esporte. O autor considera um diferencial em seu trabalho a possibilidade de obter os destaques do vídeo através de mais de um componente do mesmo, são eles: componentes visuais, de áudio e textos. A extração utilizando áudio é realizada a partir da excitação no áudio do vídeo. Já o texto não é suficiente para determinar um momento importante, logo é utilizado em conjunto com o áudio. Enquanto que os componentes visuais determinam um destaque a partir da observação de que, geralmente, nessas situações é realizado um *close-up* nos jogadores e espectadores, logo pouca ou nenhuma parte do gramado é visível na imagem. Por esse motivo, o algoritmo utilizado nesse trabalho detecta cenas em que há pouca ou nenhuma grama.

Em [26] é apresentada uma nova abordagem para detecção de melhores momentos em vídeos de esportes. O autor critica as abordagens para detecção de melhores momentos a partir de recursos áudio-visuais. Segundo ele, a obtenção desses recursos dos vídeos é difícil e costuma restringir a aplicação do detector a um determinado esporte. Nesse trabalho é proposta a utilização apenas de características visuais, tais como histograma de cores (CH), histograma de gradientes orientados (HOG) ou ambos, pois são consideradas de fácil extração. Além disso, o Autor explica que, dessa forma, a metodologia pode ser aplicada a vídeos de esportes diferentes.

[27] também propõe um método para detecção/geração de melhores momentos em vídeos de programas esportivos sem a utilização de recursos áudio-visuais. Esse método detecta segmentos no vídeo com movimento de reprodução lenta, para localizar momentos semanticamente importantes. Além disso, é permitido por esse método criar resumos, desses momentos, com diferentes durações. O método composto por três componentes: um detector de cenas em câmera lenta, que localiza e determina os seus limites, um filtro de comercial, que exclui as cenas referentes a comerciais, e um gerador de resumo, que destaca e sintetiza as cenas em câmera lenta.

Em [23] é apresentado um método para detecção de movimento, em tempo real, em sistemas de vídeo vigilância de ambientes internos. A detecção do movimento é realizada através da comparação entre valores atribuídos a cada pixel das imagens, capturada anteriormente, com valores atuais e remoção de fundo, deixando apenas elementos que se moveram os novos elementos. Além disso, esse sistema permite o envio de mensagens, através de uma rede GSM (*Global System for Mobile Communication*) ligado ao sistema, indicando a presença de um possível intruso no local.

Métodos para detecção de mudança de fundo de imagens em vídeos de vigilância também são abordados em [22]. Além

TABELA II. TEMA DOS ESTUDOS ANALISADOS

Característica	Referência
Sistemas de vídeo vigilância	[22] , [23] , [21]
Sistemas de vídeo vigilância em veículos	[12], [25]
Deteção de Momentos relevantes em vídeos	[29] , [30] , [28] , [11] , [26] , [27]
Deteção de Movimento em video	[22], [23]
Sistema de Multicameras	[22]
Estudo ou proposta de Algoritmo	[30]
Deteção de pessoas por Redes Neurais	[20]
Deteção de pessoas por pontos fiduciais	[3] , [19] , [6]
Deteção de pessoas po Análise de Componentes Principais	[10]
Detector de placas de veículos por Viola-Jones	[8], [13]

disso, nesse último foram realizados estudos sobre técnicas de processamento de imagem e deteção de mudança em imagens para câmeras fixas e moveis e métodos de registro para os sistemas de multicâmera. O objetivo foi métodos que possibilitem a interpretação coerente de uma sequência de cenas em vídeos de monitoramento por múltiplas câmeras.

Já em [25] é proposto um método para a compreensão do movimento de veículos em vídeos monitoramento de trânsito. O objetivo foi obter a análise do comportamento dos veículos sem nenhuma informação prévia de baixo nível, apenas com o conhecimento prévio de cena de tráfego. Para este fim, foram utilizados a tecnologia *Fuzzy C Mean* para agrupar as partes semelhantes do movimento capturado de trajetórias e a medida de distância de *Hausdorff* para classificar a trajetória do veículo.

Em [21] foi realizada uma revisão sistemática de artigos acadêmicos publicados entre 2000 e 2011 referentes a sistemas de video vigilância. Foram utilizadas como fontes, seis bases de dados online em que se buscou estudos através da palavra-chave vídeo vigilância. A primeira da seleção, consistiu em buscas pelos estudos nas fontes definidas, resultando em 4037 artigos. Desse total foram excluídos aqueles artigos que não apresentam a palavra-chave, no título ou no resumo. Na segunda etapa, foram aplicados critérios de exclusão que resultou em 143 artigos que foram analisados e classificados. O autor concluiu que às pesquisa em sistemas de videovigilância estão continuamente crescendo, porém a maioria dos estudos analisam e propõem algoritmos para video vigilância e a minoria abordam a interação do usuário com o sistema.

Os estudos selecionados, após as buscas com as expressões pertencentes ao “Grupo B” desta revisão sistemática, são:

[3], [19] e [6] que abordam aspectos do reconhecimento de pessoas a partir da deteção de pontos fiduciais em faces. Além desses, foram analisados o [10], que apresenta uma abordagem para o reconhecimento de pessoas utilizando os métodos de Transformada Discreta de Cosseno, [7] que utiliza esse mesmo método juntamente com o método Realce do Contraste Local para o reconhecimento seja invariante a iluminação, entre outros.

Em [19] foi desenvolvido um arcabouço de treinamento e validação de detectores que permite a deteção em tempo real de faces humanas. Foram avaliados cinco métodos de classificação linear: detectores DPI, filtros discriminativos, e os filtros conhecidos como *Unconstrained Average Correlation Energy* (UMACE), *Average of Synthetic Exact Filters* (ASEF), e *Minimum Output Sum of Squared Error* (MOSSE). Outra contribuição desse trabalho foi a criação de novas abordagens para o detector IPD com o objetivo de melhorar o desempenho do mesmo ou permitir o treinamento online. São elas: Regressor DPI, *Multiple Instance Learning IPD*, *Boosting / Bagging IPD* e *Online IPD*.

Em [3], é proposto um sistema para o reconhecimento de características faciais capaz de detectar um conjunto de treze pontos fiduciais. O modelo proposto neste trabalho é dividido em Pré-processamento, Modelo Probabilístico para Pontos Fiduciais, Detector por Produto Interno - DPI e Pós-processamento. A etapa de Pré-processamento é aplicada com o objetivo de reduzir as interferências (ruídos) nas imagens como, por exemplo, os efeitos da iluminação. Uma sequência de passos são aplicados nas imagens de face antes do treinamento dos classificadores. Estes passos são: correção de gama, filtragem por diferença gaussianas e a equalização do contraste. A segunda etapa adotou o modelo gaussiano com o objetivo

restringir o espaço de busca por pontos fiduciais da face. Em seguida foi aplicado o DPI para filtragem de correlação para reconhecimento da face. Por fim, o Pós-processamento é realizado pois é necessária a determinação de um limiar que, dada a saída do DPI, separe os pontos fiduciais dos outros pontos. Essa tarefa é atribuída aos classificadores: o método *Fisher*, SVM (*Support Vectors Machine*) e AdaBoost (*Adaptive Boosting*).

O método *Fisher* é baseado em histogramas referentes a distribuição de níveis de cinza e caracteriza-se por encontrar, dentre um conjunto de dados, os limiares entre as classes de níveis de cinza, podendo ser utilizado também para redução de dimensionalidade ou como pós-processamento de classificadores mais complexos [5]. A técnica AdaBoost (*Adaptive Boosting*) é uma das variações mais utilizadas do *Boosting*. Este último segue a ideia de que é possível construir um classificador forte a partir de um conjunto de classificadores fracos. O objetivo do AdaBoost é encontrar um conjunto de característica, dentre o conjunto de treinamento, para compor o classificador forte associando cada amostra de treinamento a um peso que corresponde à probabilidade desta amostra ser selecionada. Permitindo, dessa forma, uma melhor classificação do conjunto de entrada. Já a técnica SVM (*Support Vectors Machine*) consiste na análise de dois conjuntos de dados separados por um hiperplano cujo objetivo é agrupar a maior quantidade de dados possível em um só conjunto, ao mesmo tempo que maximiza a distância entre eles.

A proposta do trabalho desenvolvido por [6] é contribuir para a aquisição de informações relevantes para representação e avaliação das características das regiões da face humana por método de ponderação dessas regiões utilizando a técnica de transformada *Wavelets*. As características analisadas pertencem as regiões da face humana (olhos, nariz e boca), o objetivo é determinar se alguma região possui maior relevância para o reconhecimento facial, ou seja, destaque sobre as demais regiões. O modelo aplicado define duas formas distintas de parametrização das imagens da face: a primeira, utiliza apenas uma imagem para representar a região dos olhos, enquanto que a segunda utiliza duas imagens para representar a região dos olhos (uma para cada olho). As imagens são submetidas à conversão para escala de cinza para trabalhar no modo bidimensional e sua dimensionalidade é reduzida para composição dos espaços de características. Em seguida um método seleciona o valor decisivo que represente a característica da imagem no vetor de características.

As imagens utilizadas nesse trabalho foram concedidas, principalmente, pela base de dados do laboratório LAV2OI (Laboratório de Visão, Visualização e Otimização iterativa do Instituto de Informática da Universidade Federal de Goiás). O processo de extração de caracteres relevantes nessas imagens foi realizado por [6] utilizando duas técnicas do método de transformada: a transformada *wavelet de Daubechies* e de Haar.

Aplicando uma transformada wavelet bidimensional se obtém a decomposição da imagem (matriz), com seus coeficientes de aproximação (cA), coeficientes horizontais (cH), coeficientes verticais (cV) e coeficientes diagonais (cD). Para cada nível de transformada se utiliza o coeficiente de aproximação para derivada de uma nova imagem, de forma que uma imagem 64x64 pixels submetida a uma transformada (nível

1) terá resolução 32x32 pixels, se submetida a uma nova transformada (nível 2) sua dimensão será 16x16 pixels. E assim sucessivamente. As etapas para extração de características incluem as etapas de Normalização e energia, Composição dos vetores de características, Vetores de características com três e quatro imagens, Ponderação dos vetores de características e Análise e Representações.

Após obter os coeficientes da imagem (matrizes) por meio da decomposição da transformada wavelets se faz necessário a normalização de todos os coeficientes obtidos exceto os coeficientes de aproximação que são usados para obter uma nova resolução da imagem. Nesta normalização é feita a adequação dos valores no intervalo de 0 a 1, em seguida calcula-se a energia dos coeficientes que foram normalizados. Para se calcular o percentual dos coeficientes faz-se necessário obter o histograma da imagem. Assim cada elemento da matriz de um determinado coeficiente é dividido pela resolução da matriz de coeficientes. Uma vez obtido o percentual de cada elemento da matriz do coeficiente, é feito o cálculo para se obter a energia, que nada mais é do que fazer o somatório elevando ao quadrado para cada percentual encontrado.

O vetor de característica é construído após a etapa de normalização e contém todas as informações que são julgadas importantes na especificação de uma imagem ou ponto característico da face. Cada vetor possui informações pertinentes a cada região distinta. Os vetores são constituídos de três posições para cada nível de transformada, sendo que a primeira posição contém informações horizontais (cH), segunda posição informações verticais (cV) e a terceira informações diagonais (cD). Ao final o vetor de características possuirá dimensão de acordo com o número de transformadas utilizadas na imagem.

A partir do vetor de característica de uma face, podem ser construídos os vetores de características ponderados, que se destina as faces em que serão aplicadas o filtro gaussiano (ruído). Como é possível verificar na imagem a seguir. E por fim, é elaborada uma forma de representação e comparação dos vetores, para isso se faz necessário a utilização de uma técnica para obter as distâncias entre pontos, a técnica utilizada foi a do cálculo da distância Euclidiana. Após gerar todas as distâncias para todas as faces do banco de imagens utilizado, é gerada uma matriz com todas as distâncias, onde cada linha da matriz representa uma face testada e as colunas às distâncias para cada teste de ponderação.

O objetivo do estudo realizado por [10] foi analisar os efeitos da aplicação de um filtro de processamento de imagens, denominado Transformada Census, em uma base de dados com imagens da face em diferentes condições de iluminação. Para isso, se fez necessário associar este filtro ao método de Análise de Componentes Principais no reconhecimento de faces (PCA). Esse último permite identificar padrões em dados, e expressar esses dados de tal maneira a realçar suas similaridades e diferenças. Além disso, PCA tem a vantagem de que uma vez encontrados tais padrões, é possível comprimir os dados reduzindo-se o número de dimensões sem que haja muita perda de informação [10].

De acordo com os estudos realizados em [10], o funcionamento do PCA se dá seguinte maneira. Cada imagem de face do conjunto de treinamento é representada em um vetor, em seguida é calcula-se o vetor médio das faces e o subtrai

Olhos	Nariz	Boca
0.5	0.25	0.25
0.25	0.5	0.25
0.25	0.25	0.5
0.33	0.33	0.33
0.75	0.125	0.125
0.125	0.75	0.125
0.125	0.125	0.75

Figura 4. Matriz de ponderação para regiões da face [6]

por cada vetor do conjunto de treinamento. Posteriormente, calcula-se a matriz de covariância dos vetores resultantes em que os elementos da diagonal principal representam a variância da mesma característica, e os demais elementos representam a covariância entre duas diferentes características.

Ainda para a análise de componentes principais, é necessário obter o valor autovetores a partir da matriz de covariância, pois para o reconhecimento de padrões, é desejável que a covariância entre características diferentes seja zero ou próxima de zero. O último passo consiste em calcular uma combinação linear dos autovetores com o conjunto de dados originais a fim de se obter o espaço de faces. Dado o espaço de face, cada face de treinamento é projetada neste espaço e seus coeficientes são armazenados para as comparações futuras.

O filtro Transformada Census, alvo deste estudo de [10], é definido como um método de processamento de imagem baseado no mapeamento de bits de uma janela. A janela é definida basicamente como um conjunto de comparações entre intensidades de pixels em uma vizinhança local, representando quais deles possuem valores de brilho menores que um dado pixel central. Experimentos foram realizados utilizando a técnica PCA com imagens da base de dados AR antes e depois da aplicação da Transformada Census. Os resultados desses experimentos mostraram que a aplicação da Transformada Census melhorou o resultado do reconhecimento das faces.

Em [20] é apresentada uma metodologia de visão computacional que utiliza redes neurais MLP (*Perceptron Multicamadas*) para segmentar a cor da pele e a textura da face, de outros objetos presentes em uma imagem de fundo complexo. A imagem resultante é dividida em regiões, e para cada região são extraídas características que são aplicadas em outra rede neural MLP para identificar se naquela região contem face ou não

O objetivo do trabalho desenvolvido por [8] é criar detector de placas de veículos que alcance bons desempenhos de acurácia e velocidade. Para isso, foi utilizada uma adaptação do método Viola-Jones (muito utilizado em detecção de face) que consiste numa cascata de classificadores que avaliam caracte-

terísticas baseadas nas *wavelets de Haar*, todos treinados pelo algoritmo de aprendizagem supervisionada *AdaBoost* sobre um conjunto de imagens contendo exemplos positivos e negativos. A estrutura de cascata permite descartar rapidamente as regiões de fundo da imagem, focando no processamento em áreas com maior chance de conter o objeto [8].

A construção do detector ocorreu em três etapas, são elas: o treinamento de classificador, adaptação do detector ao domínio de placas e aprimoramentos no mecanismo de detecção. Para a construção do detector de placas implementou-se o detector de Viola-Jones disponível na biblioteca OpenCV. Sendo assim, para realizar a etapa de treinamento, as placas de veículos foram separadas em dois tipos: foram consideradas placas do “tipo 1” aquelas cujo fundo é mais claro que os caracteres, enquanto o “tipo 2” é formado pelas placas cujo fundo é mais escuro que os caracteres. Foram treinados portanto dois classificadores em cascata, cada um responsável por um tipo de placa. Para o treino de cada classificador, foram usadas 10.000 imagens positivas, às quais foram aplicadas rotações aleatórias.

Para a adaptação do detector ao domínio de placas foi escolhida uma abordagem simples de varrer a imagem duas vezes. Foi utilizado um classificador auxiliar simples, cuja meta é diferenciar um tipo de placa do outro. O número de estágios foi limitado a apenas 1, visando a eficiência. O resultado desse treinamento foi um classificador composto de apenas uma característica.

O objetivo do aprimoramento no mecanismo de detecção é aprimorar o nível de acurácia da imagem diminuindo o mínimo possível a velocidade de processamento. Uma consequência de se fazer uma varredura mais minuciosa são ocorrências de múltiplas detecções ligeiramente deslocadas para uma mesma instância (positiva ou não). O algoritmo do detector de Viola-Jones trata essa questão agrupando os retângulos vizinhos e retornando uma única detecção para cada grupo, cujas coordenadas são determinadas pela média das coordenadas dos retângulos no grupo.

De acordo com os testes realizados por [8], O detector desenvolvido é flexível, podendo ser sintonizado para obter bons desempenhos de acurácia e velocidade em diversas aplicações. Atende requisitos de tempo real para a aplicação em foco, com taxa de detecção (99,4%) e taxas de falso positivo e falso negativo da ordem de 0,5%.

O objetivo dos estudos realizados em [13] é apresentar e avaliar técnicas de reconhecimento de placas de veículos em imagens e vídeos. Segundo o autor, os algoritmos para reconhecimento de placas são, geralmente, compostos por três etapas, são elas: localização da placa, segmentação dos caracteres da placa, e o reconhecimento de cada caractere. As duas primeiras etapas incorporam técnicas de processamento de imagem em imagens estáticas ou vídeos, cuja avaliação se baseia na taxa de reconhecimento verdadeiro e a taxa de reconhecimento de erro. Para este projeto, é interessante os estudos referentes a etapa de localização de placa em imagens. Para essa etapa foram avaliados métodos das seguintes categorias:

- 1) Processamento de imagem binária: estão incluídas nesta categoria, técnicas baseadas em combinações de estatísticas de ponta e morfologia matemática.

Os resultados obtidos por métodos dessa categoria são considerados muito bons, apesar de que, quando aplicados sozinhos, os mesmos não são eficientes para imagens complexas devido alta sensibilidade a bordas indesejadas. Nestes métodos, a grandeza de gradiente e a variância de uma imagem local, são calculados com base no princípio de que a alteração de brilho na região de placa é mais notável e mais frequente do que em outros lugares. São exemplos de métodos pertencentes a essa categoria:

- a) Top Hat, capaz de detectar objetos pequenos de brilho significativamente diferente (esse algoritmo, apesar de oferece uma taxa de detecção de 80%, é altamente dependente da distância entre a câmara e o veículo).
 - b) Algoritmo de extração de placa híbrido com base em estatísticas de ponta e morfologia, proposto em [14], que oferece cerca de 99,6% de precisão.
 - c) A análise de componentes Connected (CCA) é uma técnica que "varre" uma imagem binarizada e rotula seus pixels em componentes baseados em conectividade (4 ou 8). Uma vez que todos grupos de pixels tenham sido determinados, a cada pixel é definido um valor de acordo com o componente a que foi atribuído. São úteis para medições espaciais, como áreas.
 - d) Operador de Sobel: esse operador aplica duas máscaras de filtragem de tamanho 3x3, uma máscara para detecção vertical e outra para detecção horizontal, que são comparadas com a imagem original e resultam em outras duas imagens. Para determinar se um ponto é uma borda, é estabelecido um valor de limiar e analisado o valor do gradiente em cada ponto. Se o valor do ponto for maior que o do limiar imposto considera-se que é um ponto de borda [15].
- 2) Processamento de Tons de Cinza: nessa categoria a localização de placas é feita através da identificação de arestas por pixels escuros. Neste estudo foram apresentadas algumas abordagens para essa categoria, tais como:
- a) Processamento de Imagem Global: o autor apresenta os algoritmos utilizados nos trabalhos, [16] e [17] como exemplos de utilização dessa abordagem. Em [16] o algoritmo seleciona a zona que apresenta o máximo de contraste local que (possivelmente) corresponde ao retângulo que contém o LP. Em [17] o contraste ente os caracteres da placa e o fundo branco também é utilizado para localizar a região da placa.
 - b) Análise de Imagem Parcial: nesta abordagem a imagem do veículo é digitalizada com a distância N-linha, contando as arestas existentes. Caso o número de arestas seja maior do que um valor limiar, a presença de uma placa pode ser assumida.
 - c) Medidas Estatísticas: o processamento da im-

agem é feito em "blocos de cinza" nesta abordagem. Os blocos com bordas altas ou com bordas variadas são consideradas possíveis regiões de placa. Essa abordagem foi utilizada no trabalho desenvolvido em [18] e ofereceu cerca de 92,5% de precisão.

- d) Representações Hierárquicas: consiste no mapeamento de regiões de imagem de acordo com o contraste apresentado pelas mesmas. As regiões com maior contraste são agrupadas em blocos menores, como é o caso da placa veicular, e dessa forma tornam o processamento mais rápido.
 - e) Segmentação de Região: nessa abordagem utiliza-se estatísticas de imagem, como o desvio-padrão e o valor médio como uma "heurística" para uma possível localização da placa.
 - f) Transformações da Imagem: para essa abordagem, o filtro de Gabor é um exemplo de técnica muito utilizada para localizar placas de veículos. Essa técnica é utilizada em análise de textura em um número ilimitado de direções e escalas e de acordo com este estudo pode oferecer até 98% de precisão, no entanto não deve ser aplicado em imagens complexas devido ao seu processamento lento. A técnica de transformada de Hough (HT) e a transformada wavelet (WT) também pertencem a esta abordagem.
- 3) Processamento de Cor: devido à instabilidade da cor da placa de veículos com relação a mudanças de iluminação, as técnicas para localização de placas baseada em cor não oferecem um alto grau de precisão. Outro problema é que como as cores das placas variam de acordo com o país, os algoritmos baseados em cor só podem ser aplicados para um país específico.
- 4) Classificadores: para esta categoria o autor cita as técnicas AdaBoost e SVM já apresentadas acima pelos estudos de [3].

III. ESTUDO DE CASO

Com o intuito de adquirir mais informações sobre a utilização de sistemas de vídeo vigilância, particularmente para o monitoramento de veículos, foi elaborado um formulário de pesquisa. O formulário foi submetido a três perfis de pessoas, são elas:

- Grupo A: Pessoas que não utilizam sistemas de vídeo vigilância
- Grupo B: Pessoas que utilizam sistemas de vídeo vigilância como ferramenta de trabalho, mas não exclusivamente para o monitoramento de veículos, tais como: funcionários de segurança de edifícios empresariais, residenciais e centros de lojas.
- Grupo C: Pessoas que utilizam sistemas de vídeo vigilância como ferramenta de trabalho e exclusivamente para o monitoramento de veículos. Nesse grupo foram

incluídos os policiais militares que analisam vídeos de sistemas de vigilância de viaturas.

O objetivo principal do formulário é identificar o método de análise de vídeos de sistemas de vigilância, especialmente em sistemas de vigilância veicular. Através do formulário buscou-se conhecer as ferramentas utilizadas e o grau de dificuldade inerente a essa atividade. Essas questões foram discutidas com pessoas que utilizam sistemas de vídeo monitoramento como ferramenta de trabalho, e que foram incluídas nos grupos A e B dessa pesquisa.

Outro objetivo foi ampliar o entendimento de como esses sistemas são inseridos no contexto urbano e qual o interesse de pessoas físicas (pessoas não vinculadas a uma instituição) em obter tais sistemas, principalmente em seus veículos particulares. Para essa finalidade, o formulário de pesquisa também foi apresentado às pessoas que não utilizam sistemas de vigilância, mas que frequentam ambientes onde eles estão ou poderiam estar inseridos como por exemplo os ônibus coletivos e os veículos particulares.

O formulário foi apresentado a um total de vinte pessoas, sendo que dez dessas se enquadram no "Grupo A" dessa pesquisa, sete ao "Grupo B", e três ao "Grupo C". O formulário possui quinze questões abordando diferentes aspectos sobre sistemas de vídeo vigilância e a atividade de análise dos vídeos resultante desse monitoramento. São elas:

- 1) Você acredita que os dados capturados pelo sistema de vídeo vigilância podem ser utilizados para auxiliar no desvendamento de crimes e contravenções?
- 2) Em uma escala de zero a dez, qual o nível de importância você acredita que esses dados têm?
- 3) Você utiliza sistemas de vídeo monitoramento?
- 4) Se você possui veículo próprio, por qual motivo não utiliza sistemas de vídeo monitoramento nele?
- 5) O sistema de monitoramento, caso você utilize, permite obter informações específicas? Por exemplo: obter imagens de um determinado período de tempo, dia da semana, etc
- 6) Com qual frequência você analisa vídeos de monitoramento?
- 7) Quantas pessoas analisam os vídeos junto com você?
- 8) Como é realizada a análise dos vídeos de vigilância?
- 9) Algum dado é coletado após a análise dos vídeos?
- 10) Em uma escala de zero a dez, qual nível de atenção você acha que é necessário para analisar os vídeos de vigilância?
- 11) Em uma escala de zero a dez, qual o grau de dificuldade você encontra para identificar uma placa de veículo no vídeo?
- 12) Em uma escala de zero a dez, qual o grau de dificuldade você encontra para identificar e reconhecer uma pessoa no vídeo?
- 13) Você considera útil o reconhecimento do condutor para o monitoramento de veículos?
- 14) Você considera que alterações bruscas de velocidade de um veículo podem indicar um momento importante em vídeo de vigilância veicular?
- 15) Você considera que a presença de um obstáculo próximo ao veículo em movimento pode indicar um momento importante em vídeos de vigilância veicular?

Todas as pessoas questionadas acreditam na importância dos dados provenientes de vídeos de vigilância, alcançando uma média nove em uma escala de zero a dez. No entanto, não foram encontradas pessoas que utilizam esses sistemas em veículos particulares. As dificuldades que desestimulam a aquisição desses sistemas em um veículo particular foram apontadas, como: alto custo do serviço de instalação dos dispositivos de vigilância no veículo e falta de conhecimento com relação a essa tecnologia. Metade dos entrevistados, pertencentes ao grupo A dessa pesquisa, justificaram não utilizar esse tipo de monitoramento em veículos particulares por falta de conhecimento dessa tecnologia, a outra parte consideram alto o custo desses dispositivos.

As informações recolhidas das pessoas que utilizam sistemas de vídeo vigilância, incluídas nos "Grupos B e C", revelaram que esses sistemas, em sua maioria, permitem a pesquisa de trechos do vídeo de acordo com a data e horário. No entanto, em um dos edifícios empresariais visitados, O Amaratuba, as imagens capturadas só são armazenadas após a identificação de movimento através de sensores, além disso, a observação do vídeo com pouca ou nenhuma iluminação é facilitada pela utilização de sensores de calor para imagens noturnas. Em outro edifício, o Trade Center de Salvador, o sistema de vigilância permite o reconhecimento de funcionários através da face obtida no vídeo. Na Figura 5 são apresentadas as técnicas, encontradas nessa pesquisa, que facilitam a análise dos vídeos.

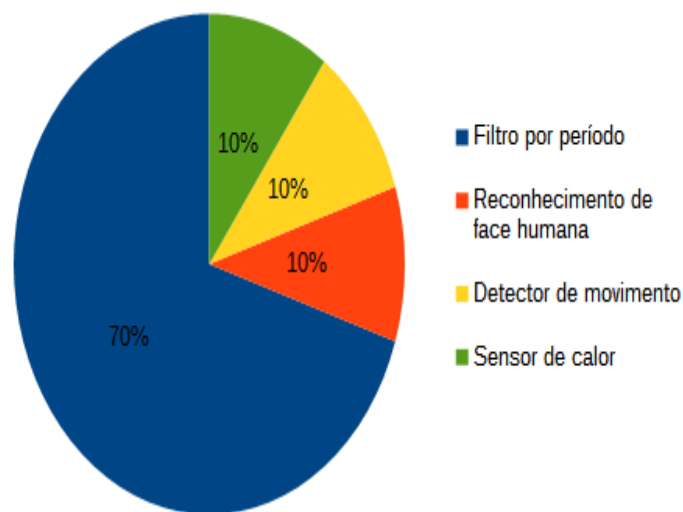


Figura 5. Técnicas para análise de vídeos encontradas em estudo de caso

Notou-se também que, nas empresas ou estabelecimentos, a atividade de observação dos monitores do sistema de vigilância é realizada por apenas um profissional por turno. E segundo os mesmos, essas atividades demanda um nível mediano de atenção, sendo que os vídeos são revistos pela direção da empresa em casos de crimes ou incidentes. Além disso, foi relatado que o reconhecimento de pessoas e placas de veículos no vídeo depende principalmente da qualidade da imagem, sendo que na maioria dos casos é considerado mais difícil reconhecer uma placa de veículo do que uma pessoa.

E ainda, por ser o foco desse projeto, foram entrevistados policiais militares (instalados na secretaria de segurança

pública sediada no Parque Tecnológico de Salvador) que analisam os vídeos resultantes do sistema de vigilância em viaturas. São acopladas internamente e externamente (ao redor) desses veículos cameras para captura de imagens. As imagens são armazenadas em um HD para posterior análise de policiais internos, os policiais que utilizam a viatura não têm acesso ao conteúdo gravado pelas câmeras. O objetivo dessa vigilância é utilizar as imagens capturadas para auxiliar no desvendamento de crimes ou condutas errôneas dos policiais.

A esses policiais, incluídos ao "Grupo C" desse estudo de caso, foi questionado se seria útil a indicação no vídeo de vigilância dos eventos definidos nesse projeto como momentos relevantes do vídeo. E a resposta, tanto para a identificação do período do vídeo em que o veículo em movimento estiver próximo de um obstáculo, quanto para quando o veículo tem sua velocidade alterada bruscamente foi positiva. A resposta também foi positiva para a localização de placas de veículos e face humana nas imagens, sendo constatada que a atividade de identificação da placa é mais difícil para também para os policiais.

IV. PROJETO

Sabe-se da importância dos vídeos de vigilância para auxiliar no esclarecimento de ocorrências de crimes e acidentes e até mesmo para inibir a execução de delitos. Especialmente no trânsito, o monitoramento por vídeo é bastante utilizado para fiscalizar infrações de trânsito, tais como: condutores que ultrapassam o limite de velocidade na via, avanço de sinal em sinalizadas, veículos acima do peso e/ou altura permitida trafegando na via, etc. Esses vídeos fornecem dados que podem ser úteis para desvendar esses casos, por exemplo: imagens de acidentes e atropelamentos, placas de veículos irregulares ou roubados, imagens de pessoas envolvidas em crimes no trânsito, entre outros. No entanto, a análise desses vídeos demanda tempo, fato que pode vir a favorecer a impunidade em episódios de crimes.

O objetivo desse projeto é reduzir o tempo em análise de vídeos de vigilância veicular, destacando os momentos mais relevantes do vídeo. Para essa finalidade foi desenvolvido nesse projeto um protótipo para detectar eventos que podem indicar um momento relevante no vídeo de vigilância veicular. O funcionamento desse protótipo ocorre da seguinte maneira:

- Inicialmente é realizada a captura das imagens através de câmeras ao redor do veículo.
- No momento em que os sensores identificarem um evento, esse período (antes e após a detecção do evento) de filmagem é rotulado para indicar que o mesmo é relevante.
- Inicia-se o processo de detecção de Placas de veículos e face humana, sendo que após a localização de uma face a mesma deve ser comparada com a imagem do motorista (previamente cadastrada no sistema).
- Os dados são exibidos para os usuários.

Caracterizam-se como os principais requisitos funcionais necessários para o atendimento dos objetivos propostos por esse projeto, os itens abaixo:

- Destacar os períodos do vídeo ao detectar um evento.

- Localizar e destacar placar de veículo nas imagens do vídeo que forem destacadas
- Localizar e destacar pessoas nas imagens do vídeo que forem destacadas
- Permitir o cadastro da imagem da face do motorista do veículo
- Reconhecer o motorista do veículo
- Permitir a definição do tempo de destaque do vídeo a cada evento

A estrutura do protótipo desenvolvido nesse projeto pode ser dividida em módulos, são eles:

- Módulo de Infraestrutura: inclui os hardwares (câmeras e sensores) utilizados no sistema, desde a captura das imagens, detecção de eventos, até a geração dos momentos relevantes do vídeo.
 - 1) câmera board Raspberry Pi com resolução nativa de 5 megapixels, e lente de foco fixo onboard. Essa câmera é capaz de fornecer imagens estáticas 2592 x 1944 pixels, e também suporta 1080p30, 720p60 e 640x480p60 / 90 vídeo.
 - 2) Sensor ultrassônico de distância HY-STRO5 cuja distância de detecção é entre: 2cm e 450cm, possui alta precisão: até 0,3 cm, ângulo sensor de não mais do que 15 graus, tensão de funcionamento de 5V (DC) e corrente de estática de menos de 2mA.
 - 3) Módulo GPS (*Global Positioning System*)
 - 4) Cartão de Memória 32 G
 - 5) Placa Raspberry Pi B+
- Módulo de Aplicação: responsável pela captura e processamento das imagens obtidas dos dispositivos externos (câmeras), controle dos dispositivos que detectam eventos (sensores) e implementação do algoritmo de reconhecimento de face humana e detecção de padrões placas de veículos.

Além disso, o protótipo oferece as funções de localizar placas de veículos e reconhecer face humana que possam estar presentes nessas imagens. E dessa forma, gerar um vídeo de menor duração contendo os momentos relevantes da filmagem. Os momentos relevantes são reconhecidos a partir da detecção de um evento. Neste projeto, são considerados como "eventos" as seguintes condições:

- A identificação de algum obstáculo a frente do veículo feita através do uso de sensores de distância. É importante ressaltar que, para este evento, considera-se como condição de ocorrência a velocidade na qual um veículo se aproxima de um obstáculo.
- Alterações bruscas de velocidade do veículo. Por exemplo: quando o veículo está prestes a colidir com algo, geralmente, ocorre uma redução brusca de velocidade, instantes antes, provocada pelo motorista na tentativa de evitar o choque. Essa situação caracterizaria um momento do vídeo para análise.
- Detecção e identificação da face do motorista. Por exemplo, a identificação de um motorista diferente do

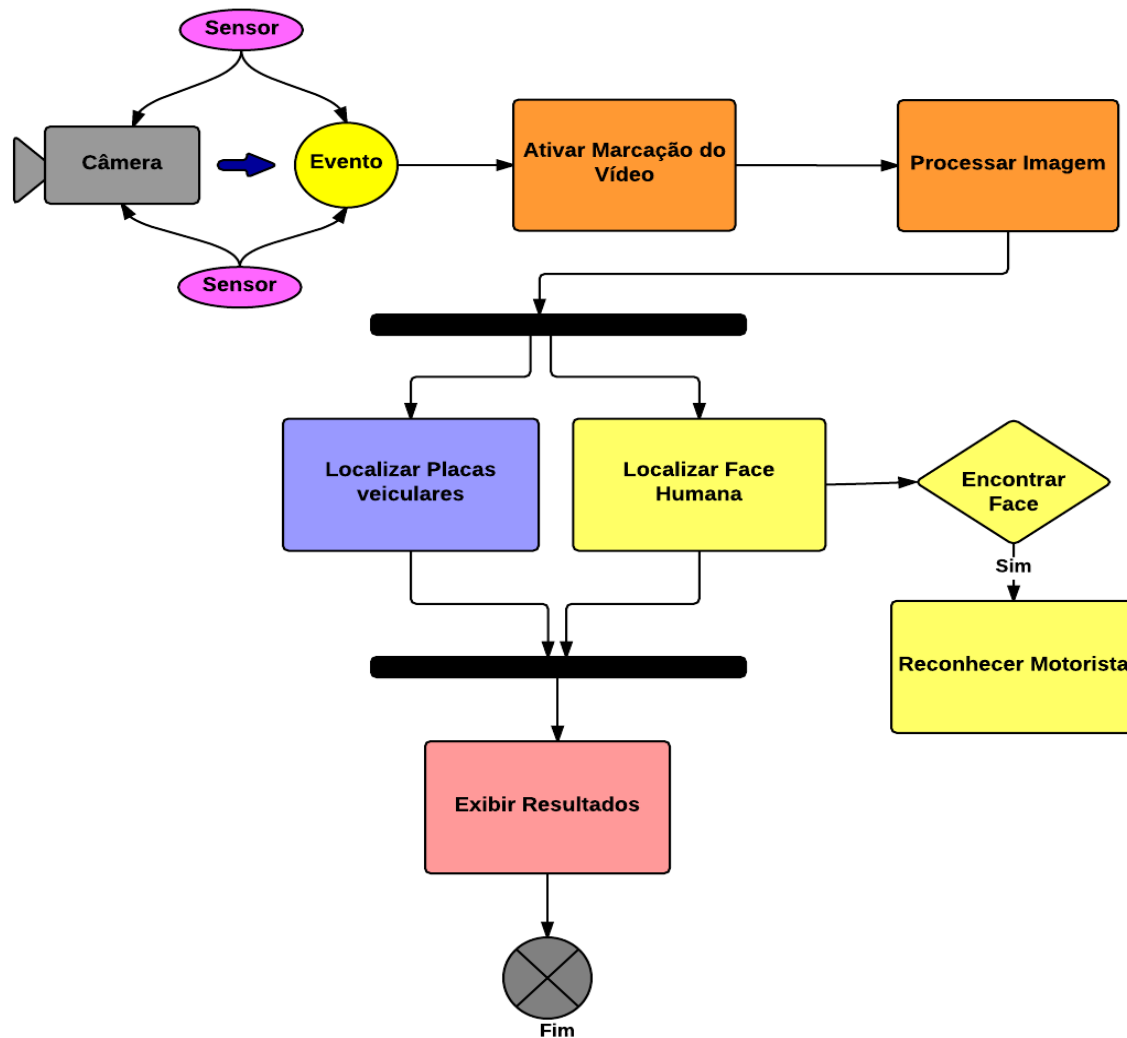


Figura 6. Etapas do funcionamento do sistema

pré-cadastrado dispararia um evento e consequentemente a marcação no vídeo.

As etapas de funcionamento do protótipo desenvolvido é representado pela Figura 6 que especifica as funções de detecção de eventos, localização de placas de veículos e de pessoas oferecidas pelo sistema. Também é representado na figura os dispositivos utilizados neste trabalho, como por exemplo a câmera para captura de imagens e os sensores para identificação de eventos.

Enquanto que o fluxo de execução do mesmo pode ser melhor compreendido pelo diagrama de sequência apresentado pela Figura 7. Nele é possível identificar a ordem em que cada atividade é realizada e as dependências entre elas. Inicialmente são capturadas as imagens, enquanto se verifica a existência de um evento. Ao detectar um dos eventos descrito anteriormente nesse texto, as imagens são processadas para redução de ruídos e detecção de face e bordas. Por fim, são exibidos os instantes detectados como relevantes para o vídeo.

Esse sistema visa garantir principalmente desempenho, portabilidade e mobilidade. Sabe-se que sistemas que processam imagens demandam um elevado poder de processamento e que também é necessário uma qualidade mínima das imagens capturadas para que possam ser utilizadas na identificação de padrões. No entanto, para reduzir os custos, foram utilizados materiais com poder de processamento não tão grande, como o raspberry e câmera board Raspberry Pi com resolução nativa de 5 megapixels, mas suficientes para um bom desempenho em um sistema não comercial. A portabilidade está implícita no algoritmo desenvolvido em linguagem de programação Python e a biblioteca OpenCV que fornecem essa qualidade ao sistema. A mobilidade é adquirida com os dispositivos móveis utilizados.

Além disso, é uma característica desse sistema a sensibilidade ao contexto, necessária para a identificação dos eventos que permitem a detecção de um momento relevante no vídeo. Isso é possível devido a utilização de sensores conectados ao raspberry, os sensores de distância, que detectam obstáculos

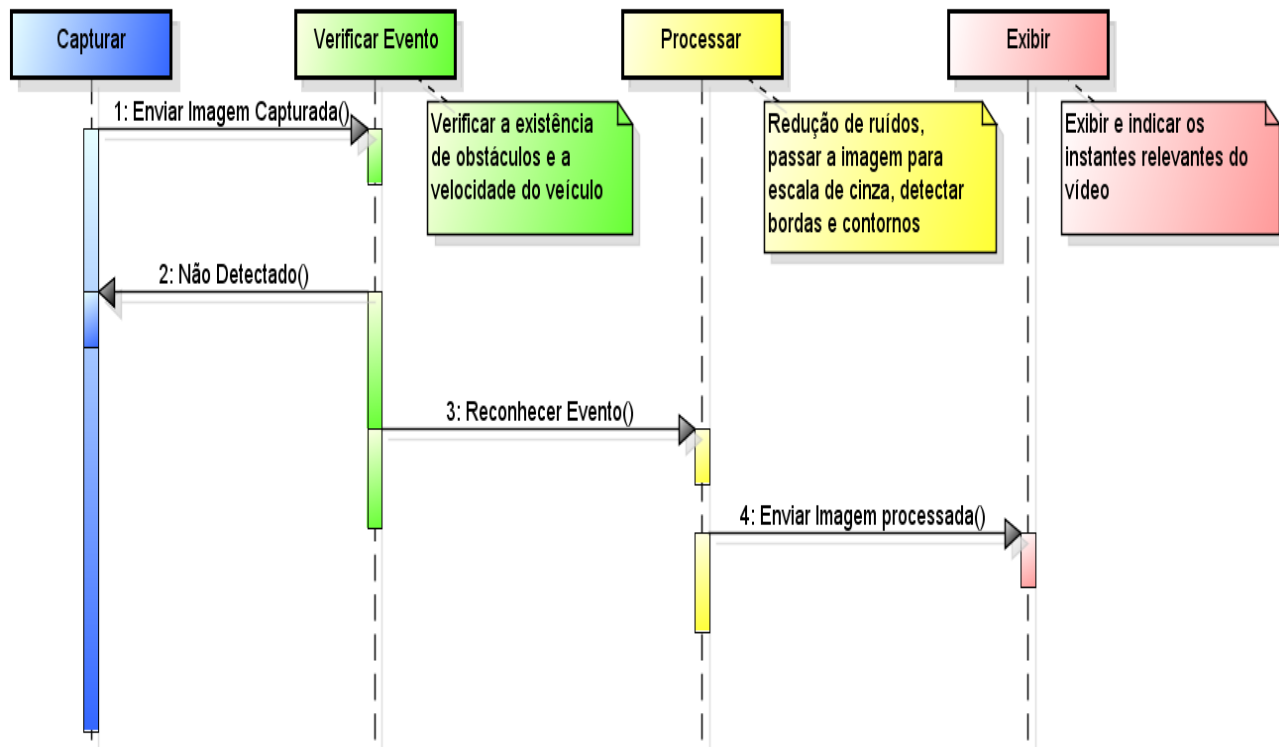


Figura 7. Diagrama de Sequência

ao redor do veículo e indica a velocidade do mesmo através do cálculo apresentado no próximo tópico.

V. IMPLEMENTAÇÃO

A implementação desse projeto pode ser dividida em cinco etapas, são elas: captura das imagens, detecção de eventos, processamento de imagens, detecção face e placas de veículos e obtenção do período relevante do vídeo. A linguagem de programação utilizada para a implementação do algoritmo foi Python, utilizando a biblioteca de visão computacional OpenCV em sua versão 2.4.10. A Figura 8 representa essas etapas de implementação, destacando as principais atividades exercidas em cada etapa.

A. Captura das Imagens

Neste projeto, algumas imagens foram obtidas por diferentes dispositivos de captura para possibilitar os testes que foram aplicados, tais como: câmera de smartphone e webcam. É possível, no entanto, utilizar a câmera raspberry board que possui dimensões 25mm x 20mm x 9mm e pesa cerca de 3g. Essa câmera está conectada ao processador BCM2835 do raspberry Pi através do CSI (Câmera Interface Serial) uma ligação de largura de banda mais elevada que transporta dados de pixel da câmara para o processador.

As imagens capturadas tiveram seu tamanho pré-fixado em 640x480. Imagens maiores teriam o tempo de processamento muito alto e fotos com resolução menor oferecem um insuficiente detalhamento de imagem, por isso, a resolução foi mantida como definida para esse projeto.

A função `picâmera` do Python é utilizada no algoritmo 1 que mostra como foi configurada a resolução da câmera para 640x480 e como o vídeo é inicializado para gravação para um arquivo em disco chamado `video.h264`. Na linha seguinte, verifica-se a utilização da função `"wait_encoding(time)"` para identificar a existência de possíveis erros durante a gravação (como por exemplo: sem espaço em disco) e gera uma exceção caso isso ocorra. Por fim, a gravação é encerrada a partir da função `"wait_encoding()"`.

Algoritmo 1. Gravar vídeo

```

def gravar():
    with picamera.PiCamera() as camera:
        camera.resolution = (640, 480)
        camera.start_recording('video.h264')
        camera.wait_recording(10)
        camera.stop_recording()
  
```

B. Detecção de Evento

Para detectar os eventos que indicam momentos relevantes no vídeo de vigilância do veículo foi utilizado o sensor ultrassônico de distância HY-SRF05, combinado ao Raspberry Pi B+ para possibilitar a interação do sistema com o contexto do mundo real. A distância do veículo em relação a um obstáculo foi obtida com o auxílio da biblioteca `LibsensorPy`.

O sensor de distância HY-SRF05 possui cinco pinos por onde se dá a transmissão de sinais para detecção de um obstáculo, VCC responsável por fornecer a alimentação de 5v para o sensor, Trig recebe o sinal de que a detecção será iniciada, Echo: retorna a distância entre o sensor e um obstáculo, Out que não foi utilizada para esse protótipo e GND

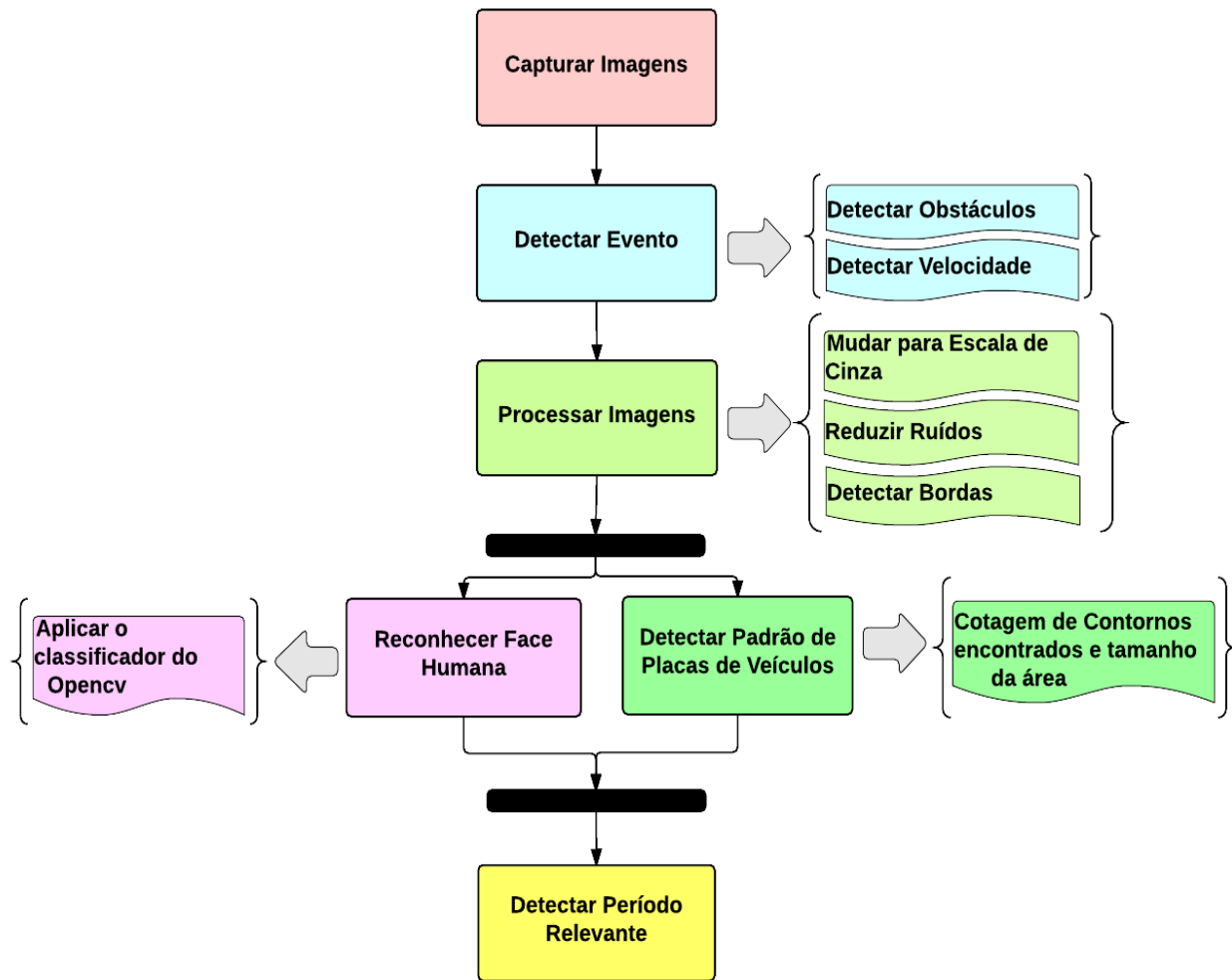


Figura 8. Etapas da Implementação

(ground). Os pinos Vcc e GND foram utilizados nas entrada 2 e 6 do raspberry, enquanto que os pinos Tring e Echo foram acoplados as entradas 12 e 13 respectivamente.

O funcionamento do sensor consiste, basicamente na emissão de uma onda sonora de alta frequência que, ao atingir um obstáculo, volta para o sensor, e com base no tempo entre o envio e recebimento do eco produzido quando esta onda se choca com um objeto capaz de refletir o som, é calculada a distância entre o sensor e o obstáculo. Mais especificamente o sensor funciona da seguinte maneira:

- 1) Um sinal com duração de 10 microssegundos é enviado ao pino trigger, indicando que a medição terá início.
- 2) O sensor emite pulsos de 40 KHz e aguarda o retorno do sinal pelo receptor.
- 3) Quando um objeto reflete estes pulsos, o eco resultante é recebido e convertido em um sinal elétrico.
- 4) Caso haja um retorno de sinal, é determinada a distância entre o sensor e o obstáculo utilizando a seguinte equação: $Distancia = (\text{pulso} \times \text{velocidade do som}) / 2$.

A LibsensorPy facilita a utilização de sensores em sistemas embarcados/ubíquos. Usando a biblioteca, não é necessária a configuração no sistema das conexões entre o GPIO e o sensor, bem como torna-se dispensável fazer a leitura dos dados enviados pelo sensor, que são geralmente dados binários ou hexadecimal, fazer a conversão, interpretação e validação desses dados lidos. A biblioteca fornece para o usuário os dados de forma legível, pois é mais fácil entender 25°C do que "1x0 0x0 2x2 0x0 1x0", por exemplo [31].

O algoritmo apresentado em 2 refere-se a implementação do método que obtém a distância do sensor a um determinado obstáculo utilizando a LibsensorPy. Essa biblioteca abstrai para o desenvolvedor a necessidade de configuração dos pinos do raspberry para o sensor no algoritmo.

Algoritmo 2. Detectar Obstáculo

```

def detectarObstaculo():
    srf05 = UltrasonicSensorFactory
    .createSensor("HYSRF05")
    result = srf05.distance_in_cm()
    return result
  
```

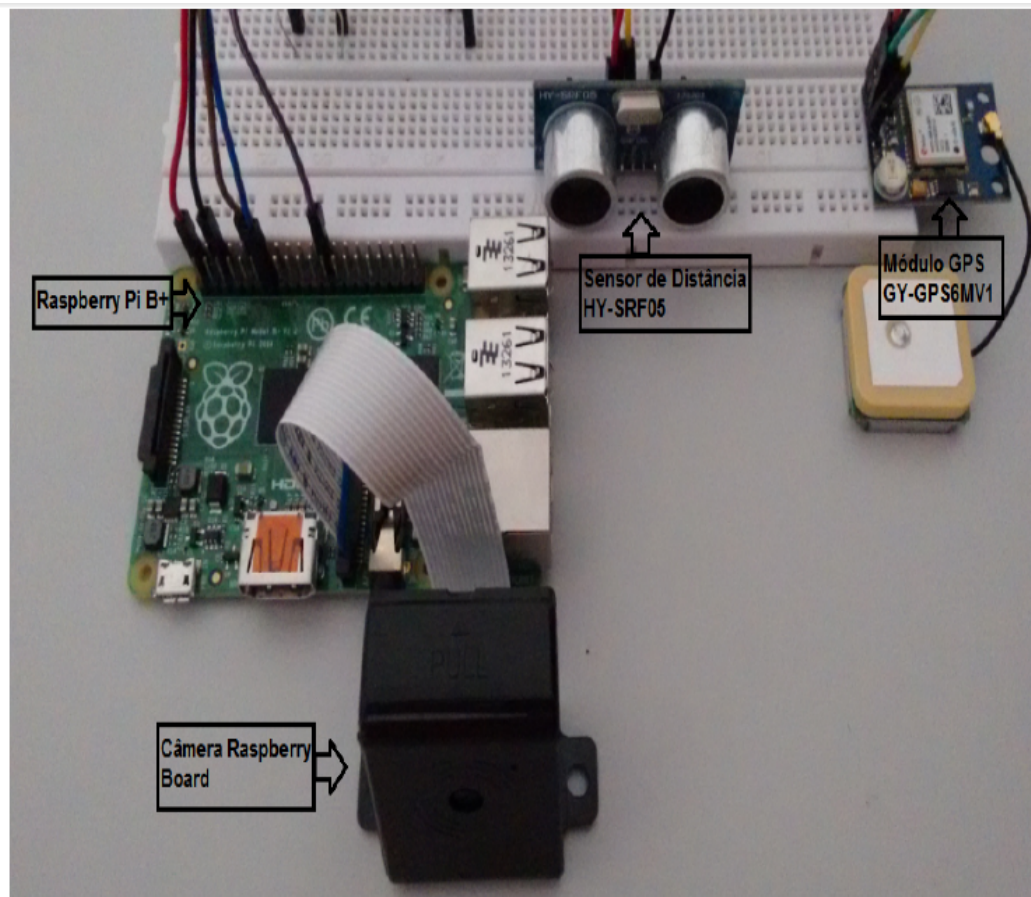


Figura 9. Protótipo para o Monitoramento Veicular

Inicialmente, testou-se a captura da velocidade média a partir dos dados obtidos pelo sensor de distância. Considerando que, quando o sensor, detecta um obstáculo a menos de 3m do veículo o sistema poderia armazenar essa distância inicial e começar uma contagem de tempo. Assim, com a diferença de tempo inicial para o tempo atual e a diferença de distância inicial e atual seria possível obter a velocidade de deslocamento, representada pela equação:

$$Vm = \frac{\Delta d}{\Delta t}. \text{ Onde:}$$

Vm = Velocidade Média

Δd = Diferença de deslocamento

Δt = Diferença de Tempo

No entanto essa teoria não seria verdadeira para um veículo em movimento, considerando que o obstáculo tido como referência também poderia estar em movimento, como, por exemplo, no caso de outro veículo. Por isso, optou-se pela utilização do GPS para se obter a velocidade de veículo. O módulo GPS utilizado foi o modelo GY-GPS6MV1 que é composto por uma antena ativa cerâmica para captura de sinal forte, uma bateria e um chip EEPROM para armazenamento de dados. A comunicação do módulo é feita por meio de uma porta serial, possui 4 pinos, VCC, RX, TX e GND.

A imagem 9 apresenta o protótipo desenvolvido nesse projeto para o monitoramento de veículos. É possível verificar

a utilização do sensor de distância, responsável por indicar um momento relevante durante a gravação do vídeo de vigilância, além da câmera que captura as imagens.

C. Processamento de imagens

Geralmente as atividades de visão computacional necessitam de uma etapa de pré-processamento. Neste projeto as imagens foram processadas para reduzir ruídos da imagem, converter o formato, tamanho e cores das imagens, entre outras funções. São considerados ruídos, além da qualidade da imagem, a interferência de variáveis como: iluminação do ambiente, condições climáticas no momento da aquisição da imagem, posição entre o objeto de interesse e a câmera, etc.

Essas alterações nas imagens foram necessárias para aumentar o desempenho da aplicação e facilitar a identificação de padrões de placa de veículos, sendo útil também para facilitar a detecção de face humana. Essas atividades foram desenvolvidas utilizando a biblioteca OpenCV que disponibiliza funções que permitem a realização do processo de captura de imagens e vídeos, adição de filtros, tratamento de imagens, entre outros.

A figura 10 apresenta as imagens, de um veículo após o processamento descrito nessa seção. É possível verificar a diferença entre a imagem original e a resultante da aplicação de operadores e filtros. Inicialmente a imagem original é convertida para escala de cinza com o intuito de reduzir a quantidade

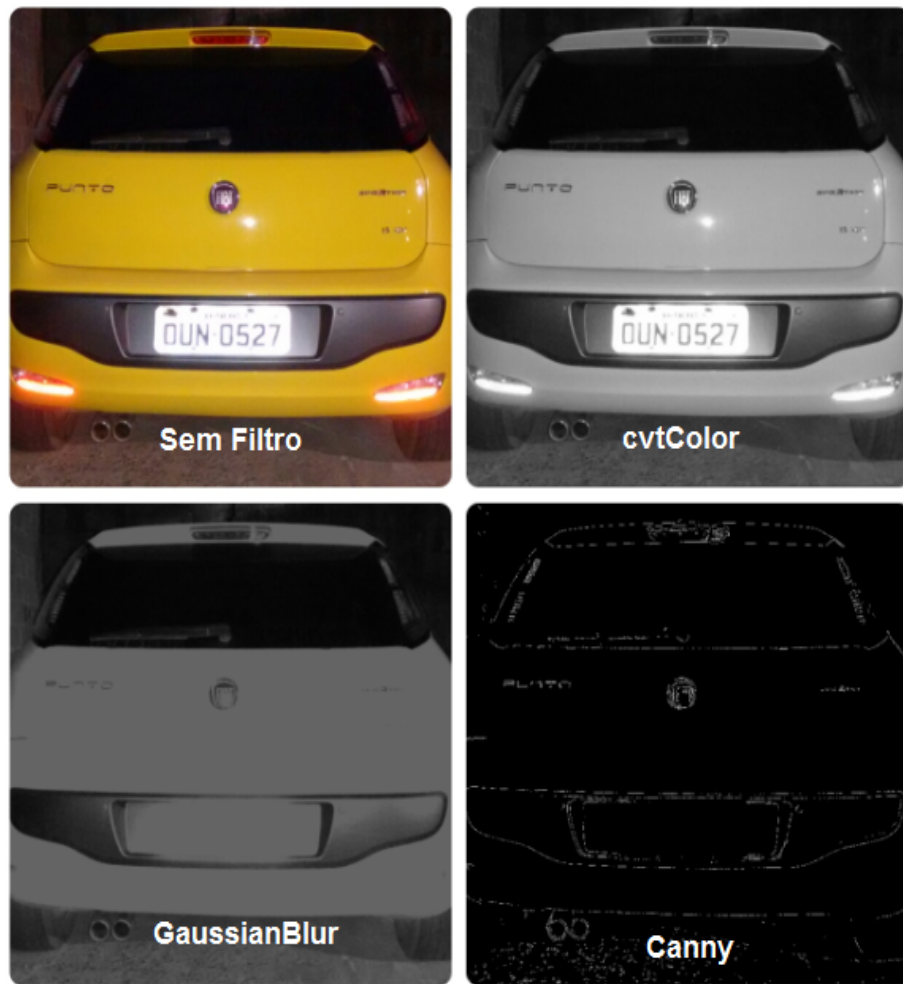


Figura 10. Imagens resultantes dos operadores aplicados na etapa de processamento de imagens

de canais de cores de três (RGB) para um canal (Grayscale). Essa última se utiliza apenas da iluminação, eliminando-se assim a matiz e a saturação da imagem e consequentemente melhorar o desempenho do sistema. Isso é realizado através da função "*cv2.cvtColor(frame, cv2.COLOR_BGR2GRAY)*". Essa função possui como parâmetros a imagem capturada inicialmente, e a constante que indica a conversão da imagem para escala de cinza.

Após a conversão para escala de cinza, a imagem é processada para a correção de ruídos utilizando o filtro Gaussiano implementado pela função "*GaussianBlur*" fornecida pelo OpenCV. O filtro Gaussiano tem a finalidade diminuir a diferença de tons entre os pixels, deixando a imagem aparentemente "borrada" para descartar imperfeições, ruídos e linhas pouco evidentes, facilitando a identificação de padrões. Os parâmetros informados nessa função foram: a imagem em escala de cinza, a variável que recebe a nova imagem, o tipo de filtro para suavizar, que nesta implementação foi indicado pela constante "*cv_gaussian*".

A partir da imagem resultante da aplicação do filtro Gaussiano, foi realizada a detecção de bordas com o auxílio da função Canny. O operador Canny exerce, basicamente, duas funções: suavizar a imagem e detectar bordas. São definidas

três metas para os algoritmos de detecção de bordas, as quais são atingidas de maneira simples utilizando o OpenCV. Tais metas determinam o que é necessário para detecção de bordas, como por exemplo:

- Baixa probabilidade de erros.
- O algoritmo deve ser preciso na detecção da borda, ou seja, a borda detectada deve estar bem próxima da borda real.
- Evitar a detecção de bordas duplas.

O algoritmo 3 apresenta as funções do OpenCV descritas anteriormente e que foram utilizadas para o processamento de imagens nesse projeto.

Algoritmo 3. Processar Imagens

```
def processarImagem(frame):
    imgCinza = cv2.cvtColor(frame,
        cv2.COLOR_BGR2GRAY)
    imgSemRuido = cv2.GaussianBlur(imgCinza, (
        5, 5), 0)
    imgBordas = cv2.Canny(imgSemRuido, 100,200)
    return imgBordas
```

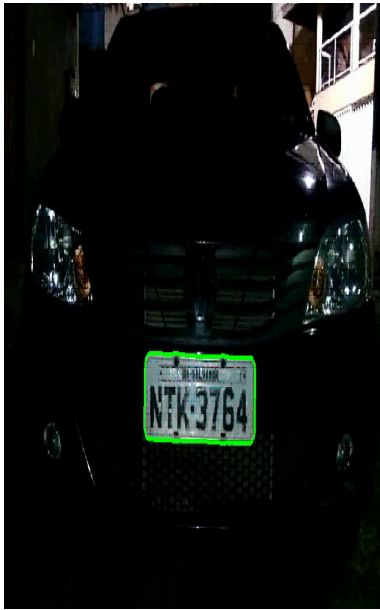


Figura 11. Localização de placa em vídeo

D. Detecção de Placas de Veículos

Para detecção de padrões de placas de veículos foi aplicado nas imagens a segmentação por bordas. Uma borda em uma imagem é caracterizada por uma mudança consideravelmente grande entre o nível de intensidade dos pixels. Os detectores de borda são definidos para encontrar este tipo de variação nos pixels e quando estes pixels estão próximos eles podem ser conectados formando uma borda e dessa forma definindo uma região ou objeto [15].

Considerando que as placas veiculares seguem um padrão de tamanho e de formato retangular foi desenvolvido o método 4. Nesse método foi utilizado a função `findContours` e `approxPolyDP` do OpenCv para buscar contornos nas bordas das imagens e aproximar os pontos encontrados nesses contornos, respectivamente. Em apresenta o resultado da aplicação do algoritmo desenvolvido nesse projeto em um vídeo.

Algoritmo 4. Detectar Placa

```
def detectarPlaca(frame, imgBordas):
    contornos, hierarquia = cv2.findContours(
        imgBordas,
        cv2.RETR_TREE, cv2.CHAIN_APPROX_SIMPLE)
    if contornos != None:
        for c in contornos:
            peri = cv2.arcLength(c, True)
            approx = cv2.approxPolyDP(c, 0.02 *
                peri, True)
            area = abs(cv2.contourArea(c));
            if len(approx) == 4 and area > 2000:
                cv2.drawContours(frame, [c], -1,
                    (0, 255, 0), 3)
            return frame
        break
```

A Figura 11 apresenta o resultado do algoritmo aplicado a um vídeo que exibe um veículo. A detecção foi realizada em condições ideais para o algoritmo, que são de pouca

iluminação e a menos de três metros de distância entre a câmera e o veículo.

E. Detecção e Reconhecimento de Pessoas

Neste projeto foi desenvolvido um algoritmo para detecção e reconhecimento de pessoas. O objetivo dessa parte da implementação é permitir a identificação do motorista do veículo, e em um trabalho futuro notificar quando uma pessoa não autorizada for identificada no vídeo de vigilância. As etapas de desenvolvimento desse módulo do projeto são apresentadas pela Figura 12.

Para detecção e reconhecimento da face humana nos vídeos de vigilância foi utilizado o algoritmo de Viola-Jones através da biblioteca `Opencv`. O método Viola-Jones foi desenvolvido baseando-se no detector Wavelets e é dividido em unidades básicas denominadas features retangulares. Cada feature possui um valor calculado por uma equação específica, através dessa representação calcula-se o resultado do somatório dos valores dos pixels que ocupam determinada área retangular na imagem. O algoritmo Viola-Jones utiliza a técnica Haar para identificar nas imagens as características de uma face humana. A partir do resultado obtido, as áreas da imagem são classificadas.

Uma cascata de classificadores avaliam as características, todos treinados pelo algoritmo de aprendizagem supervisionada AdaBoost sobre um conjunto de imagens contendo exemplos positivos e negativos. A partir da estrutura de cascata é possível identificar as regiões de fundo da imagem e descartá-las, focando no processamento em áreas com maior chance de conter o objeto que está sendo procurado [8].

Para o reconhecimento de face é utilizado o método Eigenfaces do `opencv` que aplica a técnica de Análise de Componente Principal (PCA) para resumir a dimensionalidade de uma imagem transformando um conjunto de variáveis correlacionadas em um conjunto menor de variáveis não correlacionadas a fim de diminuir o custo computacional e melhorar a precisão. O Eigenface permite reconhecer características da face humana em imagens, e aplicando esse método a uma base de treinamento é possível identificar pessoas.

A imagem 12 apresenta o método utilizado neste projeto para identificar a face de uma pessoa. Uma pasta com imagens do condutor do veículo, previamente adicionada na aplicação, é percorrida para realizar o treinamento do identificador de padrões do `Opencv` (Eigenface). Nesse projeto, um total de quatrocentas fotos foram processadas para identificação de características. Em seguida os padrões treinados são comparados com a face capturada no vídeo e é realizado o reconhecimento ou não da pessoa.

Algoritmo 5. Detectar Pessoa

```
def reconhecerFace():
    for filename
        in walk_files(POSITIVE_DIR, '*.pgm'):
            faces.append(prepare_image(filename))
            labels.append(config.POSITIVE_LABEL)
            pos_count += 1

    for filename
        in walk_files(NEGATIVE_DIR, '*.pgm'):
            faces.append(prepare_image(filename))
            labels.append(config.NEGATIVE_LABEL)
```

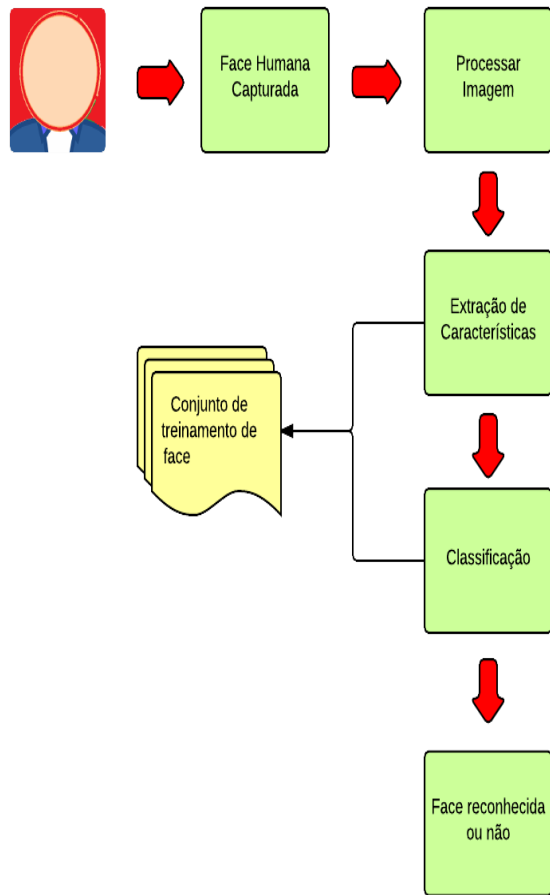



Figura 12. Etapas para o reconhecimento de pessoas

```

neg_count += 1

model = cv2.createEigenFaceRecognizer()
model.train(np.asarray(faces),
            np.asarray(labels))
face_cascade = cv2.CascadeClassifier(
    'haarcascade_frontalface_alt_tree.xml')
  
```

Em 13 é apresentado o resultado do reconhecimento de pessoas realizado pelo algoritmo aqui desenvolvido.

F. Teste Funcional

Os testes aplicados no protótipo desenvolvido nesse projeto foram realizados de acordo com as etapas de implementação do protótipo. Nas etapas de captura de imagens e detecção de eventos os testes se resumiram ao atendimento da funcionalidade necessária para o sistema, que são a obtenção de imagens na resolução definida e armazenamento no raspberry, a detecção de obstáculos a partir do sensor de distância e a obtenção da velocidade pelo GPS.

Os testes referentes a detecção de placas e face humana foram realizados a partir de imagens capturadas aleatoriamente

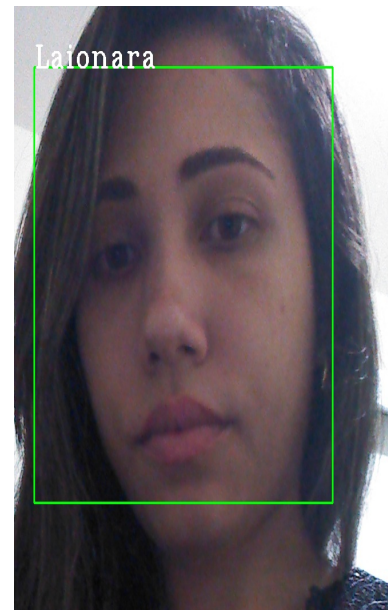


Figura 13. Face reconhecida pelo algoritmo

de placas de veículos e de pessoas para verificar a detecção a partir do sistema. Os resultados apresentados referem-se a quantidade de detecções de padrões de placas de veículos e de face humana que foram realizados corretamente pelo sistema.

A Tabela III apresenta o resultado de testes realizados para verificar a eficiência do método de reconhecimento de pessoas descrito anteriormente nesse projeto. Foram definidas cinco situações em que pessoas aparecem no vídeo, sendo a resolução do vídeo a mesma em todas as situações (640x480). A identificação de face humana foi testada diante de variações de iluminação, número de pessoas em presentes nas imagens, pessoas fazendo gestos como bocejar e espreguiçar e ocultando os olhos (usando óculos escuros, por exemplo).

TABELA III. TESTES DE DETECÇÃO E RECONHECIMENTO DE FACE HUMANA

Vídeo	Nº de Pessoas	Circunstância	Resultado
1	1	Muita iluminação	Bom
2	1	Pouca iluminação	Muito Bom
3	2	Face lateral do rosto (perfil)	Razoável
4	1 e 2	Gestos como bocejar e espreguiçar	Razoável
5	1 e 2	ocultando os olhos	Razoável

Concluiu-se com os testes que o algoritmo implementado apresentou bons resultados para detecção de face. A detecção foi prejudicada, em alguns momentos, devido ao posicionamento do rosto, sendo mais difícil por exemplo, a

localização da lateral da face. Imagens em que os olhos são ocultados também não foram obtidos bons resultados. Isso se dá principalmente devido a característica da técnica PCA utilizada nesta implementação que busca a identificação de face a partir da premissa que devem encontrar em todas as faces olhos, nariz e boca. Contudo, em situações de variação de iluminação do ambiente a detecção da face foi pouco afetada.

Enquanto que com relação ao reconhecimento de pessoas, notou-se a dependência da quantidade de imagens adicionadas para treinamento do algoritmo, quanto maior o número de imagens, mais preciso é o reconhecimento da face. Em imagens com mais de uma pessoa diferente, o resultado apresentado em alguns momentos, considerava as duas pessoas como sendo iguais, e em imagens com expressões como bocejar e espreguiçar o reconhecimento também foi prejudicado.

Testes também foram realizados para verificar a eficiência do algoritmo desenvolvido neste projeto para a detecção de placas. A tabela IV apresenta os resultados dos testes em situações diferentes. Mantendo a resolução do vídeo em 640x480, foi analisada a identificação de placas diante de variações de iluminação, e distância.

TABELA IV. TESTES DE DETECÇÃO PLACAS

Vídeo	Nº de Veículos	Circunstância	Resultado
1	1	Muita iluminação	Ruim
2	1	Pouca iluminação	Muito Bom
4	1 e 2	Distância de até 3m	Bom
5	1 e 2	Distância maior que 3m	Ruim

O algoritmo de detecção de placas de veículo apresenta ótimos resultados em imagens em que o veículo encontra-se a menos de três metros da câmera e paralelo a mesma. Enquanto que em imagens mais distantes, o algoritmo tende a considerar outros objetos como sendo uma placa devido a grande quantidade de informações contidas nas imagens. Além disso, percebeu-se que a detecção de placas é favorecida pela pouca iluminação do ambiente, apesar do tratamento aplicado pelo algoritmo na imagem antes da busca por padrões de placas de veículos, em ambientes muito iluminados ainda é difícil a identificação da placa.

É importante ressaltar que não é o objetivo desse projeto apresentar um algoritmo excelente para o reconhecimento de pessoas e localização de placas veiculares. Essas intervenções nos vídeos foram implementadas para ajudar o usuário na análise dos vídeos e consequentemente reduzir ainda mais o tempo gasto nessa atividade.

VI. LIMITAÇÕES

Não se pode garantir que os eventos definidos (já especificados anteriormente nesse texto) como indicadores de um momento relevante no vídeo serão realmente importantes para o usuário no momento da análise dos vídeos. Os eventos são identificados pelo algoritmo desenvolvido nesse projeto e pelos sensores conectados ao sistema. Portanto, a qualidade dos componentes de hardware (câmera e sensores) utilizados nesse sistema influenciam a precisão na identificação de um momento relevante do vídeo.

A detecção de velocidade, especialmente, é comprometida em função do local em que se esteja o módulo GPS. Em locais fechados, principalmente, o envio de sinais do dispositivo para os satélites é bastante comprometida. Por este motivo, os testes foram realizados com dados de velocidade aleatórios.

Além disso, a localização de placas de veículos e o reconhecimento de pessoas podem ser comprometidas em imagens de baixa qualidade. Para alcançar bons resultados na extração de informações, é necessário obter como entrada do sistema uma imagem de boa qualidade [4]. Ou seja, o bom funcionamento do sistema aqui proposto depende também da qualidade das imagens que serão processadas. Por esse motivo foi incluída nesse sistema a etapa de processamento de imagem, representada na Figura 6, que tem como objetivo diminuir os ruídos presentes nas imagens.

Outras limitações estão relacionadas ao meio de obtenção das imagens. Fatores como: luminosidade, oscilações na pista, entre outros podem tornar difícil a identificação de um evento e a obtenção das imagens. Outro fator que deve ser levado em consideração é a velocidade de processamento do computador. O sistema deve ser ajustado para trabalhar com uma capacidade de processamento equivalente a velocidade de gravação da câmera para possibilitar a detecção em tempo real [4].

Existem ainda limitações impostas pelo algoritmo utilizado para interpretação das imagens. Cada algoritmo possui um nível de eficiência específico, portanto a eficiência na localização das placas de veículos e reconhecimento de face humana dependem da eficácia do algoritmo utilizado. Nesse projeto, algumas limitações proveniente do algoritmo desenvolvido foram identificadas na etapa de testes, podemos citar por exemplo a dificuldade em identificar uma placa em imagens com muitos objeto semelhantes.

VII. MELHORIAS E TRABALHOS FUTUROS

Embora o principal objetivo desse trabalho tenha sido alcançado, a análise dos resultados do estudo de caso e os testes aplicados ao protótipo que foi desenvolvido revelaram que esse projeto pode ser melhorado e permite o desenvolvimento de trabalhos futuros. Entre as funcionalidade já oferecidas e que podem ser melhoradas, destacam-se as seguintes:

- 1) Utilizar outros sensores para ampliar os tipos de eventos a serem detectados.
- 2) Utilizar a localização do veículo como um evento que indique um momento relevante do vídeo. Por exemplo, um automóvel que esteja em uma rota que não é aquela que costuma seguir poderia ser considerado como um momento relevante do vídeo.

- 3) Aplicar outros testes de eficiência em detecção de placa de veículos e face humana ao algoritmo desenvolvido neste trabalho.
- 4) Buscar um maior número de consumidores de sistemas de vídeo vigilância veicular para apresentar o protótipo.

Existem funcionalidades que podem ser implementadas com o intuito de atender outras necessidades do usuário, e dessa forma gerando novos trabalhos, tais como:

- 1) Controlar e obter informações do protótipo através de uma aplicação para smartphone.
- 2) Integrar com outros sistemas de segurança, por exemplo: sistemas de controle de estacionamentos, sistema de fiscalização de veículo em vias de trânsito, entre outros.
- 3) Notificar ao usuário via smartphone em caso de desvio de rota do veículo.
- 4) Reconhecer os caracteres da placa do veículo e comparar com um banco de dados para obter informações sobre a situação legal do mesmo.

VIII. CONCLUSÃO

Este trabalho apresenta uma revisão sistemática de estudos que seguiu a seguinte estratégia de pesquisa: foram selecionados estudos pertencentes a três áreas distintas que são aplicadas neste projeto, são elas: detecção de momentos relevantes em vídeos de vigilância (especialmente vigilância de veículos), detecção de face humana em imagens, e detecção de placas de veículos em imagens. As pesquisas foram realizadas no Google Scholar com restrições de período de publicação, entre 2000 e 2014, e que apresentem as palavras chaves previamente definidas dando preferência a estudos publicados no IEEE e ACM Portal.

A partir dos resultados dessa revisão sistemática constatou-se que existem estudos referentes a detecção de melhores momentos em quase todos os anos considerados nessa pesquisa, com exceção do ano 2000. A grande maioria dos estudos relacionados a detecção de melhores momentos foi aplicado a vídeos de esportes, como, por exemplo, em vídeos de jogos de futebol e Basquete. Os eventos mais utilizados para a detecção dos melhores momentos nesses vídeos foram a voz do narrador do jogo e o áudio dos torcedores. Já para a localização de objetos em imagens foram encontradas nos estudos a utilização de técnicas como o Viola-Jones, Filtro Gabor, AdaBoost, entre outros.

Além disso, foi realizado um estudo de casos com aplicação de um formulário de pesquisa para obter informações de como são utilizados os sistemas de vídeo vigilância em empresas localizadas em Salvador. A pesquisa foi realizada em três tipos de público, são eles: o "Grupo A" composto por pessoas que não utilizam sistemas de vídeo vigilância, o "Grupo B" pessoas que utilizam sistemas de vídeo vigilância como ferramenta de trabalho, mas não exclusivamente para o monitoramento de veículos, e o "Grupo C" formado por pessoas que utilizam sistemas de vídeo vigilância como ferramenta de trabalho e exclusivamente para o monitoramento de veículos.

O resultado obtido a partir das respostas do formulário permite concluir que as tecnologias para vigilância em veículos

são pouco conhecidas e possuem valor de mercado pouco acessível a maioria das pessoas. Enquanto que análise dos vídeos é realizada em sua grande maioria através da observação de monitores e em poucos lugares são utilizados artifícios que facilitam a análise dos mesmos, tais como: detecção de movimento, sensor de calor para imagens noturnas, detecção de pessoas, etc.

Nesse projeto foi desenvolvido também um protótipo de para vídeo vigilância em veículos. O objetivo desse sistema é detectar os momentos mais importantes em um vídeo de vigilância veicular através da observação de eventos específicos. Além de oferecer para o usuário, dados que facilitem a análise do vídeo, tais como: detecção de padrões de placas veiculares e de face humana. Esses objetivos foram alcançados, como mostrado nos tópicos anteriores.

Portanto espera-se que esse projeto contribua para os estudos na área de vídeo vigilância veicular e que também facilite o processo de análise dos mesmos a partir da detecção de momentos relevantes. Os testes aplicados as funcionalidades oferecidas pelo protótipo que foi desenvolvido constataram a eficácia do mesmo. No entanto, são necessárias algumas melhorias que podem ser desenvolvidas em trabalhos futuros, tais como: acessar as informações fornecidas pelo protótipo a partir de smartphones, notificar o usuário sobre alguns eventos, utilizar outros tipos de sensores e dessa forma ampliar a quantidade de eventos que auxiliam na identificação de um momento relevante no vídeo, entre outros.

REFERÊNCIAS

- [1] Paul Viola, Michael Jones. Rapid Object Detection using a Boosted Cascade of Simple Features. 2001. IEEE
- [2] Barbara Kitchenham , Rialette Pretorius, David Budgen , O. Pearl Brereton, Mark Turner, Mahmood Niazi , Stephen Linkman. Systematic literature reviews in software engineering – A tertiary study. 2010.IEEE.
- [3] Gabriel Matos Araujo. Algoritmo para Reconhecimento de Características Faciais Baseado em Filtros de Correlação. Fevereiro de 2010, Rio de Janeiro. Dissertação de Mestrado apresentada a Universidade Federal do Rio de Janeiro.
- [4] Jean Dias do Nascimento. Detecção e Reconhecimento de Placa Automotiva com Baixo Custo. Brasília, Dezembro 2012.
- [5] WEBB, A. R. Statistical Pattern Recognition. John Wiley e Sons, 2002.
- [6] Leandro L. G. Oliveira, Eduardo S. de Albuquerque, Iwens Gervásio Sene Júnior. Avaliação de características para o reconhecimento de faces por meio de Wavelets. VI Workshop de Visão Computacional
- [7] Larissa Natalia das Virgens Carneiro. Reconhecimento de Face Invariante a Iluminação Baseado em uma Abordagem Supervisionada. Universidade Federal de Ouro Preto 2012.
- [8] Francisco Adriano Nobre Freire, José Everardo Bessa Maia. Localização Automática de Placas de Veículos em Imagens Baseada no Detector de Viola-Jones. Universidade Estadual do Ceará - UECE. Fortaleza - CE 2013.
- [9] Maria Géssica dos Santos Aragão, Yuri Tavares dos Passos, Francisco Rodrigues Santos. Reconhecimento e segmentação de caracteres em placas de veículos usando análise de projeção e redes neurais. Erbase 2014.
- [10] Ícaro Ribeiro, Giovani Chiachia, Aparecido Nilceu Marana. Reconhecimento de Faces Utilizando Análise de Componentes Principais e a Transformada Census. VI Workshop de Visão Computacional. Universidade Estadual Paulista - UNESP 2010.
- [11] Yo-Ping Huang, Ching-Lin Chiou, Frode Eika Sandnes. An intelligent strategy for the automatic detection of highlights in tennis video recordings. Agosto de 2009.

- [12] Yi-Ling Chen, Tse-Shih Chen, Tsiao-Wen Huang, Liang-Chun Yin, Shiou-Yaw Wang, Tzi-cker Chiueh. Intelligent Urban Video Surveillance System for Automatic Vehicle Detection and Tracking in Clouds. 2013.
- [13] Christos-Nikolaos E. Anagnostopoulos, Ioannis E. Anagnostopoulos, Ioannis D. Psoroulas, Vassili Loumos, Eleftherios Kayafas. Plate Recognition From Still Images and Video Sequences: A Survey. 03 de Setembro de 2008.
- [14] B. Hongliang and L. Changping, "A hybrid license plate extraction method based on edge statistics and morphology," in Proc. ICPR, 2004, pp. 831–834.
- [15] Rodolfo F. Nunes , André C. Hernandez , Marcelo Beker, "Detecção de Faixas com Câmera Monocular e Otimização Através de Algoritmo Genético", Anais do XIX Congresso Brasileiro de Automática, CBA 2012.
- [16] P. Comelli, P. Ferragina, M. N. Granieri, and F. Stabile, "Optical recognition of motor vehicle license plates," IEEE Trans. Veh. Technol., vol. 44, no. 4, pp. 790–799, Nov. 1995.
- [17] S. Draghici, "A neural network based artificial vision system for license plate recognition," Int. J. Neural Syst., vol. 8, no. 1, pp. 113–126, Fevereiro 1997.
- [18] H.-J. Lee, S.-Y. Chen, and S.-Z. Wang, "Extraction and recognition of license plates of motorcycles and vehicles on highways," in Proc. ICPR, 2004, pp. 356–359.
- [19] Felipe Moreira Lopes Ribeiro. Detecção de Pontos Fiduciais em Faces Usando Filtragem Linear. Rio de Janeiro Fevereiro de 2014.
- [20] Wellington da Rocha Gouveia. Detecção de Face Humanas em Imagens Coloridas Utilizando Redes Neurais Artificiais. Dissertação de Mestrado. Universidade de São Paulo 2010.
- [21] Fereshteh Falah Chamasemani, Lilly Suriani Affendey. Systematic Review and Classification on Video Surveillance Systems. Faculty of Computer Science and Information Technology, University Putra Malaysia, Malaysia. Junho de 2013.
- [22] Gian Luca Foresti, Christian Micheloni, Lauro Snidaro, Paolo Remagnino, and Tim Ellis. Active Video-Based Surveillance System. 2008 IEEE.
- [23] H.Venkateswara Reddy, B.Suresh Kumar. Automated Video Surveillance System With Sms Alert. 03 de Dezembro de 2013. International Journal Of Engineering And Computer Science ISSN.
- [24] Maia, J. G. R. Detecção e Reconhecimento de Objetos usando Descritores Locais. Tese de Doutorado. Universidade Federal do Ceará, Fortaleza, 2010.
- [25] Jian WU, Zhiming CUI, Yuehui ZHANG, Jianming CHEN. A Clustering-based Motion Understanding Method for Traffic Vehicle. Journal of Computational Information Systems 2011.
- [26] Hao Tang, Vivek Kwatra, Mehmet Emre Sargin, Ullas Gargi. Detecting Highlights in Sports Videos: Cricket as a Test Case. IEEE 2011.
- [27] H. Pun, P. van Beek, M. I. Sezan. Detection of Slow-Motion Replay Segments in Sports Video for Highlights Generation. IEEE 2001.
- [28] S. Dagtas, M. Abdel-Mottaleb. Extraction of TV Highlights Using Multimedia Features. IEEE 2001.
- [29] Peng Chang, Mei Han, Yihong Gong. Extract Highlights From Baseball Game Video With Hidden Markov Models. IEEE 2002.
- [30] Yong Rui, Anoop Gupta, e Alex Acero. Automatically Extracting Highlights for TV Baseball Programs. ACM Digital Library 2000.
- [31] Edivaldo M. F. de Jesus Jr, Manoel C. M. Neto. An Extendable Python Library To Manipulate Sensors Coupled To The Raspberry Pi. Instituto Federal da Bahia. Salvador/Bahia 2015