

CENTRO UNIVERSITÁRIO FEI

FABIANO BARBOSA LEITE

**USO DE ALGORITMOS DE CLASSIFICAÇÃO DE IMAGENS PARA DETECÇÃO DE
FORMAS HUMANAS EM CENAS AÉREAS DE DESASTRES**

**São Bernardo do Campo
2015**

FABIANO BARBOSA LEITE

**USO DE ALGORITMOS DE CLASSIFICAÇÃO DE IMAGENS PARA DETECÇÃO DE
FORMAS HUMANAS EM CENAS AÉREAS DE DESASTRES**

Dissertação apresentada ao Centro Universitário da FEI como parte dos requisitos para obtenção do título de Mestre em Engenharia na área de Inteligência Artificial. Orientador: Prof. Dr. Reinaldo Augusto da Costa Bianchi.

São Bernardo do Campo
2015

Leite, Fabiano Barbosa.

USO DE ALGORITMOS DE CLASSIFICAÇÃO DE IMAGENS
PARA DETECÇÃO DE FORMAS HUMANAS EM CENAS AÉREAS
DE DESASTRES / Fabiano Barbosa Leite. São Bernardo do Campo,
2016.

132 p. : il.

Dissertação - Centro Universitário FEI.

Orientador: Prof. Dr. Reinaldo Augusto da Costa Bianchi.

1. CENTRIST. 2. Máquina de vetores de Suporte. 3. Árvore de
Decisão. 4. Rede Neural Artificial. 5. Rede Bayesiana.. I. Bianchi,
Reinaldo Augusto da Costa, orient. II. Título.



CENTRO UNIVERSITÁRIO DA FEI

APRESENTAÇÃO DE DISSERTAÇÃO ATA DA BANCA EXAMINADORA

Programa de Pós-Graduação Stricto Sensu em Engenharia Elétrica

Mestrado

PGE-10

Aluno: Fabiano Barbosa Leite

Matrícula: 112102-9

Título do Trabalho: Uso de algoritmos de classificação de imagens para detecção de formas humanas em cenas aéreas de desastres.

Área de Concentração: Inteligência Artificial Aplicada à Automação

Orientador: Prof. Dr. Reinaldo Augusto da Costa Bianchi

Data da realização da defesa: 26/11/2015

ORIGINAL ASSINADA

Avaliação da Banca Examinadora:

São Bernardo do Campo, / / .

MEMBROS DA BANCA EXAMINADORA

Prof. Dr. Reinaldo Augusto da Costa Bianchi Ass.: _____

Prof. Dr. Paulo Eduardo Santos Ass.: _____

Prof.^a Dr.^a Anna Helena Reali Costa Ass.: _____

A Banca Julgadora acima-assinada atribuiu ao aluno o seguinte resultado:

APROVADO

REPROVADO

VERSÃO FINAL DA DISSERTAÇÃO

**APROVO A VERSÃO FINAL DA DISSERTAÇÃO EM QUE
FORAM INCLUÍDAS AS RECOMENDAÇÕES DA BANCA
EXAMINADORA**

Aprovação do Coordenador do Programa de Pós-graduação

Prof. Dr. Carlos Eduardo Thomaz

Dedico este trabalho a Deus e a Nossa Senhora Aparecida pela fé, à minha amada esposa pela paciência, minha mãe pelo incentivo e em memória de meu amado pai.

AGRADECIMENTOS

Acima de tudo a Deus e a Nossa Senhora Aparecida pela força interior e pela superação de minhas dificuldades. À minha querida esposa Silvana, pela paciência, compreensão e auxílio em muitas dúvidas. Ao meu orientador, Prof. Dr. Reinaldo Augusto da Costa Bianchi, pelo suporte no aperfeiçoamento deste trabalho, ao Prof. Dr. Paulo Sérgio e ao Prof. Dr. Paulo Santos por suas sugestões e observações no enriquecimento deste trabalho.

Aos colegas de mestrado e doutorado, Isaac Jesus da Silva e Danilo Hernani Périco, pelo auxílio na construção do banco de imagens para testes e em minha iniciação como usuário no sistema operacional LINUX, em que meu trabalho está estruturado.

À minha mãe Aldinete e minha irmã Fabíola, pelo incentivo e fé.

À FEI por ter me selecionado como aluno do mestrado, me mostrando novos caminhos em minha evolução profissional e dando-me oportunidade de retorno ao ambiente do conhecimento, após tantos anos.

Ao universo que conspirou para que tudo isso acontecesse para a realização dos meus sonhos. A todos o meu Muito Obrigado!

“Não existem métodos fáceis para resolver problemas difíceis.”

René Descartes

"Se você quer transformar o mundo, experimente primeiro promover o seu aperfeiçoamento pessoal e realizar inovações no seu próprio interior."

O Caminho da Tranquilidade (Dalai-Lama)

RESUMO

A proposta deste trabalho é a detecção de corpos humanos em cenas pós-desastres sob o ponto de vista aéreo, ou seja, de cima para baixo. No intuito de alcançar este objetivo procurou-se maximizar a exatidão na detecção minimizando o número de falsos positivos e negativos, levando em consideração o menor tempo possível entre detecções. Como parte integrante desta solução foi utilizado como descritor de atributos o CENTRIST, normalmente usado para descrever a topologia de cenas, e aqui, empregado para detectar o contorno humano e suas partes e que tem como virtude, a captura da estrutura global e não a textura da imagem, além de, processar em um tempo linear. Estas propriedades são relevantes, já que, o tempo de processamento é um dos parâmetros de avaliação do algoritmo como um todo. Porém o reconhecimento do corpo humano em uma cena de pós-desastre não é uma tarefa de soluções bem estudadas, como na detecção de um pedestre (posição vertical), pois, o corpo humano nesta situação sofre deformação, rotação e oclusão. A solução aqui então é uma associação de métodos diversos onde para detecção do corpo completo, caso seja possível detectá-lo em uma imagem, é realizado pela Máquina de vetores de Suporte (núcleo Linear e HIK) em cascata e tendo o algoritmo real Adaboost como alternativa ao núcleo HIK, para teste de desempenho. No caso de não haver uma detecção de corpo completo na imagem foram criados comitês de especialistas na detecção dos membros do corpo humano (cabeça, braços e pernas), formados dos seguintes classificadores: Máquina de vetores de Suporte (Linear), Árvore de Decisão e Rede Neural Artificial. A tomada de decisão, se há um corpo humano em uma imagem, é então avaliada pela concordância booleana entre comitês e pela avaliação final por uma Rede Bayesiana, que são as principais contribuições deste trabalho. Os resultados deste trabalho indicam o potencial de detecção de formas humanas em situações generalizadas de desastres.

Palavras-chave: CENTRIST. Máquina de vetores de Suporte. Árvore de Decisão. Rede Neural Artificial. Rede Bayesiana.

ABSTRACT

The purpose of this study is the human detection in post-disaster situations from the aerial view, i.e., from top to bottom. In order to achieve this goal the detection accuracy was maximized and number of false positives and negatives was minimized, taking into account the shorter time between detections. The descriptor CENTRIST was applied in this work to get the human shape because captures contours and also requires linear time, but it is commonly used to describe topology from scenes. These properties are important and linear time is relevant to get advantage in performance in the algorithm. However the human shape detection in adverse situations of positioning, direction, rotation and occlusion of body is not easy task like in pedestrian detection (vertical position). The method proposed in this work is a combining pattern classifiers to detect the human body. In case of full body made through the Support Vector Machine (Linear and HIK kernels) classifier in cascade and were also made with the Adaboost classifier instead of the Support Vector Machine (HIK). In case there is no full body detection the expert committee detector human body parts (heads, arms and legs) are made with theses classifiers: Support Vector Machine (Linear), Decision tree e Neural Network. The decision-making, if there is a human body in image, is first evaluated by the Boolean agreement between committees and finally a Bayesian Network, which are the main contribution this work. The results of this work show the potential of human bodies' detection in general situations of disaster.

Keywords: CENTRIST. Support Vector Machine. Decision Tree. Neural Network. Bayesian Network.

LISTA DE FIGURAS

Figura 1 - Cena de desastre	21
Figura 2 - Passos para computação do HOG	25
Figura 3 - Imagem integral no bloco D	26
Figura 4 - Obtenção do descritor LBP	27
Figura 5 - Representação humana com contrastes diferentes	28
Figura 6 - CENTRIST: Matriz de transformação	29
Figura 7 - CENTRIST: Correlação de valores	30
Figura 8 - Reconstrução de imagem usando o CENTRIST	30
Figura 9 - Ilustração dos níveis 2, 1 e 0 de uma imagem dividida	31
Figura 10 - Exemplo de separação feita por SVM.....	33
Figura 11 - SVM aplicação do hiperplano	34
Figura 12 - Interpretação do erro de classificação ε_i	35
Figura 13 - Árvore de decisão dos riscos da pressão arterial alta	40
Figura 14 - Árvore de Decisão: medidas de aleatoriedade.....	42
Figura 15 - Representação básica de um neurônio.....	43
Figura 16 - Transformação afim produzido pela presença de <i>bias</i>	44
Figura 17 - Função de ativação limiar	45
Figura 18 - Função de ativação Sigmóide	45
Figura 19 - Rede Neural de simples camada	46
Figura 20 - Rede Neural de multi-camadas	46
Figura 21 - Rede Neural recorrente	46
Figura 22 - Rede Bayesiana com a tabela de probabilidade condicional	47
Figura 23 - Implementação da proposta de Soni	50
Figura 24- Visão do mapa na interface com o robô	52
Figura 25 - A guia Snaps com a vítima	52
Figura 26 - Imagem em escala de cinza	53
Figura 27- Aplicação do filtro Sobel	53
Figura 28 - Superfície de exatidão de classificação	54
Figura 29 - Exatidão de classificação de 100 diferentes HOGs	54
Figura 30 - O sistema proposto por Soni.....	55
Figura 31- Resultados obtidos da esquerda para a direita respectivamente: HOG, SURF e HOG com IL	57

Figura 32 - Resultados obtidos da esquerda para a direita respectivamente: HOG, SURF e HOG com IL	57
Figura 33 - Imagem original à esquerda e a imagem após a projeção de volta à direita.....	60
Figura 34 - Resultados de várias detecções obtidas dos modelos apresentados (amarelo <i>true positive</i> (+) e vermelho falso positivo(-))......	61
Figura 35 - Gráfico de desempenho dos modelos com e sem escala a priori.....	62
Figura 36 - Avaliação entre os diferentes modelos e combinações.....	62
Figura 37- Algoritmo completo de detecção de formas humanas em cenas de desastres.....	66
Figura 38 - Ilustração da ponderação dos gradientes (G_X e G_Y) no valor absoluto de 0.012...68	68
Figura 39 - Ilustração da ponderação dos gradientes (G_X e G_Y) no valor absoluto de 0.009...68	68
Figura 40 - Ilustração da ponderação dos gradientes (G_X e G_Y) no valor absoluto de 0.003...68	68
Figura 41- Forma generalizada de calcular o vetor de atributos.....	69
Figura 42 - Forma otimizada para obter o limiar de trabalho do classificador	69
Figura 43 - Divisão da janela de detecção em blocos.....	70
Figura 44 - Aquisição de imagem, CENTRIST e determinação do limiar.....	71
Figura 45 - Estrutura de classificação em cascata.....	73
Figura 46 - Estrutura de formação dos blocos classificadores.....	76
Figura 47 - Grafo direcionado e acíclico de dependências das variáveis no modelo do corpo humano	77
Figura 48 - Imagem de corpo completo com iluminação artificial.....	86
Figura 49 - Imagem de corpo completo com iluminação natural	86
Figura 50 - Imagem de parte do corpo: Braço.....	86
Figura 51 - Imagem de parte do corpo: Perna.....	86
Figura 52 - Imagem de ambiente interno 1.....	86
Figura 53 - Imagem de ambiente interno 2.....	86
Figura 54 - Quadrotor ARDrone da Parrot SA.....	87
Figura 55 - Imagem feita pelo quadrotor (FEI).....	87
Figura 56 - Imagem feita pelo quadrotor (FEI).....	87
Figura 57 - Imagem real de desastre.....	87
Figura 58 - Imagem real de desastre.....	87
Figura 59 - Matriz de confusão.....	88
Figura 60 - Espaço da curva ROC, com pontos A e C com boa classificação, ponto B com classificação aleatória e ponto C péssima classificação.....	90
Figura 61 - Processo de treinamento dos classificadores.....	91

Figura 62 - Curva ROC do classificador em cascata (SVM-Adaboost) na condição $\Delta = 3$ e $\sigma = 0.9$ para o conjunto A.....	99
Figura 63 - Curva ROC do classificador em cascata (SVM-Adaboost) na condição $\Delta = 5$ e $\sigma = 0.8$ para o conjunto A.....	99
Figura 64 - Saída do classificador SVM-Adaboost (Em vermelho detecção de corpo completo em verde detecção de partes do corpo) para o conjunto A.....	100
Figura 65 - Curva ROC do classificador em cascata (SVM-HIK) na condição $\Delta = 3$ e $\sigma = 0.9$ para o conjunto A.....	101
Figura 66 - Curva ROC do classificador em cascata (SVM-HIK) na condição $\Delta = 5$ e $\sigma = 0.8$ para o conjunto A.....	102
Figura 67 - Saída do classificador SVM-HIK (Em vermelho detecção de corpo completo em verde detecção de partes do corpo) para o conjunto A.....	102
Figura 68 - Curva ROC do classificador em cascata (SVM-Adaboost) na condição $\Delta = 3$ e $\sigma = 0.9$ para o conjunto B.....	104
Figura 69 - Curva ROC do classificador em cascata (SVM-Adaboost) na condição $\Delta = 5$ e $\sigma = 0.8$ para o conjunto B.....	104
Figura 70 - Saída do classificador SVM-Adaboost (Em vermelho detecção de corpo completo em verde detecção de partes do corpo) para o conjunto B.....	105
Figura 71 - Saída do classificador SVM-Adaboost (Em verde detecção de partes do corpo) para o conjunto B.....	106
Figura 72 - Curva ROC do classificador em cascata (SVM-HIK) na condição $\Delta = 3$ e $\sigma = 0.9$ para o conjunto B.....	107
Figura 73 - Curva ROC do classificador em cascata (SVM-HIK) na condição $\Delta = 5$ e $\sigma = 0.8$ para o conjunto B.....	108
Figura 74 - Saída do classificador SVM-HIK (Em vermelho detecção de corpo completo em verde detecção de partes do corpo) para o conjunto B.....	108
Figura 75 - Curva ROC do classificador em cascata (SVM-Adaboost) na condição $\Delta = 3$ e $\sigma = 0.9$ para o conjunto C.....	110
Figura 76 - Curva ROC do classificador em cascata (SVM-Adaboost) na condição $\Delta = 5$ e $\sigma = 0.8$ para o conjunto C.....	110
Figura 77- Saída do classificador SVM-Adaboost (Em vermelho detecção de corpo completo) para o conjunto C.....	111
Figura 78 - Curva ROC do classificador em cascata (SVM-HIK) na condição $\Delta = 3$ e $\sigma = 0.9$ para o conjunto C.....	112

Figura 79 - Curva ROC do classificador em cascata (SVM-HIK) na condição $\Delta = 5$ e $\sigma = 0.8$ para o conjunto C.....	113
Figura 80 - Saída do classificador SVM-HIK (Em vermelho detecção de corpo completo em verde detecção de partes do corpo) para o conjunto C.....	113

LISTA DE TABELAS

Tabela 1 - Exatidões dos classificadores usadas na Rede Bayesiana.....	74
Tabela 2 - Tabela de Probabilidade Condicional (TPC) para a variável Braço.....	79
Tabela 3 - Tabela de Probabilidade Condicional (TPC) para a variável Cabeça.....	79
Tabela 4 - Tabela de Probabilidade Condicional (TPC) para a variável Perna.....	80
Tabela 5 - Tabela de Probabilidade Condicional (TPC) da presença do corpo P(HB) ou não P(\neg HB), dadas as probabilidades determinísticas Braço (B), Cabeça(C) e Perna(P) (em vermelho) + uso do OU-Ruidoso.....	82
Tabela 6 - Tamanho total das amostras.....	92
Tabela 7 - Indicadores para a amostra de corpo completo.....	92
Tabela 8 - Indicadores para a amostra Braço.....	92
Tabela 9 - Indicadores para a amostra Cabeça.....	92
Tabela 10 - Indicadores para a amostra Perna.....	93
Tabela 11 - TPC aplicada para o nó Braço usando OU-Ruidoso	94
Tabela 12 - TPC aplicada para o nó Cabeça usando OU-Ruidoso.....	94
Tabela 13 - TPC aplicada para o nó Perna usando OU-Ruidoso.....	95
Tabela 14 - TPC aplicada para o nó (Corpo Completo HB) usando OU-Ruidoso.....	96
Tabela 15 - Parâmetros do algoritmo	97
Tabela 16 - Indicadores (SVM-Adaboost) para o conjunto (A) com $\Delta = 3$ e $\sigma = 0.9$	98
Tabela 17 - Indicadores (SVM-Adaboost) para o conjunto (A) com $\Delta = 5$ e $\sigma = 0.8$	98
Tabela 18 - Tempo de processamento para o conjunto (A) com (SVM-Adaboost).....	98
Tabela 19 - Indicadores (SVM-HIK) para o conjunto (A) com $\Delta = 3$ e $\sigma = 0.9$	101
Tabela 20 - Indicadores (SVM-HIK) para o conjunto (A) com $\Delta = 5$ e $\sigma = 0.8$	101
Tabela 21 - Tempo de processamento para o conjunto (A) com (SVM-HIK).....	101
Tabela 22 - Indicadores (SVM-Adaboost) para o conjunto (B) com $\Delta = 3$ e $\sigma = 0.9$	103
Tabela 23-Indicadores (SVM-Adaboost) para o conjunto (B) com $\Delta = 5$ e $\sigma = 0.8$	103
Tabela 24 - Tempo de processamento para o conjunto (B) com (SVM-Adaboost).....	103
Tabela 25 - Indicadores (SVM-HIK) para o conjunto (B) com $\Delta = 3$ e $\sigma = 0.9$	107
Tabela 26 - Indicadores (SVM-HIK) para o conjunto (B) com $\Delta = 5$ e $\sigma = 0.8$	107
Tabela 27 - Tempo de processamento para o conjunto (B) com (SVM-HIK).....	107
Tabela 28 - Indicadores (SVM-Adaboost) para o conjunto (C) com $\Delta = 3$ e $\sigma = 0.9$	109
Tabela 29 - Indicadores (SVM-Adaboost) para o conjunto (C) com $\Delta = 5$ e $\sigma = 0.8$	109
Tabela 30 - Tempo de processamento para o conjunto (C) com (SVM-Adaboost).....	109

Tabela 31 - Indicadores (SVM-HIK) para o conjunto (C) com $\Delta = 3$ e $\sigma = 0.9$	112
Tabela 32 - Indicadores (SVM-HIK) para o conjunto (C) com $\Delta = 5$ e $\sigma = 0.8$	112
Tabela 33 - Tempo de processamento para o conjunto (C) com (SVM-HIK).....	112
Tabela 34 - Comparação dos indicadores para os conjuntos (A, B, e C) com $\Delta = 3$ e $\sigma = 0.9$	114
Tabela 35 - Comparação dos indicadores para os conjuntos (A, B, e C) com $\Delta = 5$ e $\sigma = 0.8$	115
Tabela 36 - Comparação de trabalhos correlatos.....	115
Tabela 37 - Erros de classf. (SVM-Adaboost) para o conjunto (A) com $\Delta = 3$ e $\sigma = 0.9$	129
Tabela 38 - Erros de classf. (SVM-Adaboost) para o conjunto (A) com $\Delta = 5$ e $\sigma = 0.8$	129
Tabela 39 - Erros de classf. (SVM-HIK) para o conjunto (A) com $\Delta = 3$ e $\sigma = 0.9$	129
Tabela 40 - Erros de classf. (SVM-HIK) para o conjunto (A) com $\Delta = 5$ e $\sigma = 0.8$	129
Tabela 41 - Erros de classf. (SVM-Adaboost) para o conjunto (B) com $\Delta = 3$ e $\sigma = 0.9$	130
Tabela 42 - Erros de classf. (SVM-Adaboost) para o conjunto (B) com $\Delta = 5$ e $\sigma = 0.8$	130
Tabela 43 - Erros de classf. (SVM-HIK) para o conjunto (B) com $\Delta = 3$ e $\sigma = 0.9$	130
Tabela 44 - Erros de classf. (SVM-HIK) para o conjunto (B) com $\Delta = 5$ e $\sigma = 0.8$	130
Tabela 45 - Erros de classf. (SVM-Adaboost) para o conjunto (C) com $\Delta = 3$ e $\sigma = 0.9$	131
Tabela 46 - Erros de classf. (SVM-Adaboost) para o conjunto (C) com $\Delta = 5$ e $\sigma = 0.8$	131
Tabela 47 - Erros de classf. (SVM-HIK) para o conjunto (C) com $\Delta = 3$ e $\sigma = 0.9$	131
Tabela 48 - Erros de classf. (SVM-HIK) para o conjunto (C) com $\Delta = 5$ e $\sigma = 0.8$	131

LISTA DE ABREVIATURAS

ACC	Accuracy.
AdaBoost	Adaptive Boosting.
AUC	Area Under the Curve.
CENTRIST	CENSus TRansform HISTogram.
C-HOG	Circular Histogram of Oriented Gradients.
CN	Condição Negativa.
CP	Condição Positiva.
CPT (TPC)	Conditional Probability Table (Tabela de Probabilidade Condicional).
CT	Census Transform.
DPM	Discriminatively trained multiscale Deformable part Model.
Dtree	Decision Tree.
ERR	Equal Error Rate.
FDR	False Discorey Rate.
FN	False Negative.
FNR	False Negative Rate.
FOR	False Omission Rate.
FPR	False Positive Rate.
GAD	Grafos Acíclicos Dirigidos .
HB	Human Body.
HE	HOG Ensemble.
HIK	Histogram Intersection Kernel.
HOG	Histogram of Oriented Gradients.
IL	Incremental Learnig.
K-NN	K-Nearest Neighbours.
LBP	Local Binary Pattern.
MRF	Marcov Radom Fields.
NNT	Neural Network.
NPV	Negative Predictive Value.
NR-LBP	Non Redundant Local Binary Pattern.
OPENCV	Open Source Computer Vision Library.
PBD	Poselet-Based.
PPV	Positive Predictive Value.

PS	Pictorial Structures.
R-HOG	Rectangular Histogram of Oriented Gradients.
ROC	Receiver Operating Characteristic.
RT	Result Test.
RTN	Resultado de Teste Negativo.
RTP	Resultado de Teste Positivo.
Sb_i	Sensor de braço índice i .
Sc_i	Sensor de cabeça índice i
SE	SURF Ensemble.
SLAM	Simultaneous Localization and Mapping.
SPC	Specificity.
SPD	Simple Parts Detector.
SPD-R	Simple Parts Detector with Rotation.
Sp_i	Sensor de perna índice i .
SURF	Speeded Up Robust Features.
SVM	Support Vector Machine.
TN	True Negative.
TNR	True Negative Rate.
TP	True Positive.
TPP	Total Population.
TPR	True Positive Rate.
UAV	Unmanned Aerial Vehicle.

LISTA DE SÍMBOLOS

L_i	Norma L_1, L_2 .
τ	Matriz transposta.
v	Vetor.
$\ v\ $	Norma de um vetor.
e	Constante
Σ	Somatória.
$\&$	Lógica E.
(\cdot)	Produto escalar.
w	Vetor normal para o hiperplano.
Λ	Matriz diagonal
$\varphi(\cdot)$	Função de ativação.
Π	Produto.
$P(x_i p_i)$	Probabilidade condicional X_i dado P_i .
G_X	Gradiente na direção X.
G_Y	Gradiente na direção Y.
I'	Transformação de imagem.
λ	Coefficiente do modelo de treinamento.
U	União.
β	Ajuste de ponderação dos gradientes (G_X, G_Y).
Δ	Passo do deslocamento da janela de detecção.
σ	Passo de redimensionamento da janela de detecção.
Φ_1	Limiar de trabalho do SVM-Linear.
Φ_2	Limiar de trabalho do SVM-HIK.

SUMÁRIO

1	INTRODUÇÃO.....	19
1.1	MOTIVAÇÃO	20
1.2	OBJETIVO	20
1.3	DEFINIÇÃO DO PROBLEMA	20
1.4	JUSTIFICATIVA E CONTRIBUIÇÕES.....	21
1.5	ORGANIZAÇÃO DO TRABALHO	22
2	REVISÃO TEÓRICA.....	23
2.1	DESTRITORES DE ATRIBUTOS	23
2.1.1	Histograma de Gradientes Orientados	23
2.1.2	Descritor Haar-Like	25
2.1.3	Padrão Binário Local	27
2.1.4	Centrist	29
2.2	CLASSIFICADORES	32
2.2.1	Máquina de Vetores de Suporte	32
2.2.2	Máquina de Vetores de Suporte Núcleo HI.....	36
2.2.3	AdaBoost.....	38
2.2.4	Árvore de Decisão	39
2.2.5	Rede Neural Artificial	42
2.2.6	Rede Bayesiana	46
2.3	CONSIDERAÇÕES FINAIS	48
3	TRABALHOS RELACIONADOS.....	49
3.1	DETECÇÃO E LOCALIZAÇÃO DE VÍTIMASEM DESASTRES URBANOS.....	49
3.2	USO DO HOG NADETECÇÃO DE VÍTIMAS EM ÁREAS DE DESASTRES....	53
3.3	CONJUNTO DE CLASSIFICADORES COM APRENDIZADO INCREMENTAL PARA DETECÇÃO DE VÍTIMAS DE DESASTRES.....	55
3.4	DETECÇÃO DE VÍTIMA BASEADO EM VISÃO A PARTIR DE UM QUADRUTOR	58
3.5	CONSIDERÇÕES FINAIS	63
4	DETECÇÃO DE FORMAS HUMANAS EM CENAS AÉREAS DE DESASTRES PELO USO DE COMITÊ DE ESPECIALISTAS E REDE BAYESIANA.....	64

4.1	AQUISIÇÃO DA IMAGEM E TRANSFORMAÇÕES	64
4.2	DETECÇÃO DE CORPO COMPLETO	72
4.3	DETECÇÃO DE PARTES DO CORPO	74
4.4	MODELO BAYESIANO DE DETECÇÃO DO CORPO	77
4.5	CONSIDERAÇÕES FINAIS	83
5	EXPERIMENTOS E RESULTADOS	84
5.1	BANCO DE IMAGENS E CONJUNTOS DE TESTES	84
5.2	INDICADORES DE AVALIAÇÃO DE DESEMPENHO	88
5.3	RESULTADO DOS INDICADORES DE DESEMPENHO INDIVIDUAL DE CADA CLASSIFICADOR	90
5.4	RESULTADO DO OU-RUIDOSO NA REDE BAYESIANA.....	94
5.5	EXPERIMENTOS	97
5.5.1	Experimentos com o Conjunto de Imagens Indoor (A)	98
5.5.2	Experimentos com o Conjunto de Imagens feitas pelo Drone (B).....	103
5.5.3	Experimentos com o Conjunto de Imagens de Desastres (C)	109
5.5.4	Discussão dos Resultados	114
5.6	CONSIDERAÇÕES FINAIS	116
6	CONCLUSÃO E TRABALHOS FUTUROS	117
	REFERÊNCIAS	119
	APÊNDICE A - Estratégia de concordância entre classificadores.....	126
	APÊNDICE B - Erros de classificação dos conjuntos (A), (B) e (C) para as cascatas svm-adaboost e svm-hik	128

1 INTRODUÇÃO

A detecção da forma humana tem sido objeto de intensa pesquisa, em razão da ampla gama de aplicações, que vão desde a área de segurança em locais de grandes movimentos, como estações de metrô e aeroportos, identificação de situações suspeitas, a sistemas de residências inteligentes e aplicações militares. Porém todas estas aplicações são basicamente voltadas á detecções da forma humana na posição vertical, ou seja, pedestres. Poucas pesquisas são realizadas na identificação da forma humana em situações adversas, como no caso de pós-desastres, e número menor ainda do ponto de vista aéreo, de cima para baixo, onde a utilização de quadrotoros podem de uma maneira mais ágil, localizar e informar o posicionamento de vítimas às equipes de resgates.

Diferentemente da pesquisa, já bem estudada, porém, ainda não totalmente solucionada, relativa á detecção de pedestres. A detecção da forma humana em uma situação de desastre tem um nível de dificuldade superior, pois, a estrutura do corpo é extremamente flexível e nestas condições, os conhecidos fatores de iluminação, oclusão (destrosos e outros corpos) e vestimentas humanas, dificultam a adoção dos métodos já utilizados na detecção de pedestre, ou seja, detecção de corpo completo.

Na perspectiva de obtenção de resultados mais confiáveis vários métodos tem sido empregados, desde a tentativa tradicional da detecção do corpo completo, à separação da detecção de partes do corpo. Entre as técnicas utilizadas esta a abordagem do modelo de parte deformável, multi-escala, treinado discriminativamente, “*discriminatively trained, multiscale, deformable partmodel*” (DPM) (FELZENSZWALB et al., 2008) DPM, que tenta reunir as diversas partes integrantes do corpo humano, como um conjunto, em torno da raiz, no caso o corpo completo, utilizando métodos tradicionais de classificação binária pelo uso do algoritmo de máquina de vetores de suporte, Support Vector Machine (SVM) (VAPNIK; CHERVONENKIS, 1963) e o histograma de gradientes orientados, “*histogram of gradients oriented*” (HOG) (DALAL ; TRIGGS, 2005), na identificação das partes do corpo. Outras técnicas usam o atributo Robusto acelerado (Speeded Up Robust Features) (SURF) (BAY et al., 2006) e o HOG com conjuntos de classificadores, tais como: SVM, (*k-nearest neighbours*) (*K-NN*) (FIX; HODGES, 1951) e classificador AdaBoost (FREUND; SCHAPIRE, 2003), com o uso da técnica de votação por maioria e também no histórico de classificação para determinar o resultado. Também há técnicas baseadas na regra de Bayes em que um corpo é representado como uma configuração posterior L de partes flexíveis dada a evidenciada existência do corpo E , ou seja, $p(L|E) \propto p(L) p(E|L)$,

(FELZENSZWALB; HUTTENLOCHER, 2007) e em correspondência de modelos (Castillo et al.,2005). No estudo destes trabalhos, no entanto, não foi identificado nenhum que tenha reunido técnicas de detecção de corpo completo e partes, atuando de forma independente e complementar. Este trabalho realiza esta proposta, como uso de comitês de especialistas na detecção dos membros do corpo humano (cabeça, braços e pernas) e rede bayesiana para calcular a probabilidade da existência de um corpo em uma imagem.

1.1 MOTIVAÇÃO

Auxílio na localização e resgate de pessoas em situações de pós-desastres.

1.2 OBJETIVO

O objetivo deste trabalho é o estudo, a análise e o desenvolvimento de um algoritmo, reconhecimento de formas humanas em eventos de desastres naturais ou não, sob o ponto de vista aéreo, ou seja, de cima para baixo, tendo como metas: maximizar a exatidão na detecção, minimizando o número de falsos positivos e negativos. O tempo entre detecções também é levado em consideração, pois, vidas dependem de quão ágil estas são.

1.3 DEFINIÇÃO DO PROBLEMA

O estudo de detecção de formas humanas em cenas de desastres, do ponto de vista aéreo, de cima para baixo, é uma importante parte na ação de resgate. Devido à própria situação do desastre, a cena, como um todo, é um complicador, envolvendo um plano de fundo muitas vezes confuso. Um agravante da iluminação local, somado às condições em que o corpo humano foi submetido, tais como deformações, rotações, leva a situações não favoráveis à detecção e ao reconhecimento da forma humana, por meio de apenas um único método como no caso da detecção de pedestres, posição vertical.

Em função dessas limitações, a proposta neste trabalho é utilizar associações de métodos diferentes que unam conceitos de detecções do corpo humano completo, já testadas e aprovadas, e a detecção de partes essenciais (cabeça, pernas e braços) como forma de contornar os problemas de oclusões e deformações da estrutura humana em cenários de desastres. Na Figura 1, a seguir, é apresentada uma cena com alguns desses complicadores.

Figura 1 - Cena de desastre.



Fonte: Galveston disaster, Texas, 1900.

1.4 JUSTIFICATIVA E CONTRIBUIÇÕES

Um dos grandes desafios na implementação de sistemas com aplicações práticas na detecção da forma humana é desenvolver algoritmos que consigam capturar e interpretar as melhores características que venham a representar e identificar essas formas. Entre os mais conhecidos tem-se: o Histograma de Gradientes Orientados, Histogram of Gradients Oriented (HOG) (DALAL; TRIGGS, 2005) e o Padrão Binário Local, Local Binary Pattern (LBP) (OJALA et al., 1994). Os algoritmos citados são considerados o estado da arte, porém, ainda apresentam um tempo de processamento pesado, devido a normalizações e parametrizações. Diante dessas exigências e limitações, buscou-se um algoritmo que, em primeiro lugar, reduzisse o tempo de processamento e a partir de uma simples característica representasse o corpo humano. Então não se optou pelo uso de outras características humanas, tais como cor e textura, pois, pela própria condição de desastre estas são modificadas.

O algoritmo CENTRIST, proposto em "*CENTRIST: A Visual Descriptor for Scene Categorization*" (WU; REHG, 2011), como descritor usado no reconhecimento da topologia de cenas de forma global, foi escolhido para uso neste trabalho. Apesar de esse algoritmo ter como principal uso a descrição de cenas, ele também foi utilizado na descrição do contorno da forma humana, como no artigo "*Real-Time Human Detection Using Contour Cues*" (WU; REHG, 2011). Segundo os autores, o algoritmo atingiu uma velocidade de detecção de até 20 FPS. Esse artigo, naturalmente, não é voltado para a detecção de formas humanas em cenas de desastres, e sim de pedestres; porém, a estrutura geral para uso com as imagens aéreas foi utilizada.

O CENTRIST foi, então, o descritor aqui utilizado, por apresentar um rápido processamento e capturar do contorno humano; entretanto, não foi a única solução para todos

os problemas já definidos de oclusões parciais ou quase totais da forma humana. Como tentativa de amenizar esses problemas, foi empregado um conjunto de soluções:

- a) para a detecção de corpo completo: a Máquina de Vetores de Suporte, Support Vector Machine (SVM) (VAPNIK; CHERVONENKIS, 1963), com os núcleos Linear e Intersecção de Histogramas (HIK) (SWAIN; BALLARD, 1991), em configuração de cascata, e também o classificador AdaBoost (FREUND; SCHAPIRE, 2003), como alternativa ao núcleo HIK, para avaliação de desempenho;
- b) para a detecção parcial (cabeças, pernas e braços): os classificadores SVM-Linear, Árvore de Decisão (binária), Decision Tree (DT) (BREIMAN et al., 1984) e Rede Neural Artificial, Neural Network (NNT) (MCCULLOCH; PITTS, 1943);
- c) a decisão final da presença ou não de uma forma humana, no caso de oclusões, é dada por uma Rede Bayesiana, Bayesian Network (BN) (PEARL, 1988).

A justificativa da escolha destes métodos de detecções são:

- a) no caso da detecção de corpo completo, o uso do SVM-linear (pelo rápido processamento), do SVM-HIK (exatidão de detecção) e real Adaboost como alternativa ao HIK.
- b) na detecção de partes do corpo humano, a escolha estabaseda no simples fato de empregar classificações distintas, para obter resultados também distintos para partes iguais de um corpo. Cujo objetivo é evitar vícios de resultados iguais.
- c) A rede bayesiana foi escolhida por associar o comitês de especialistas (classificadores de partes), para obter a probabilidade de existência de um corpo. Que é com a detecção de partes uma das principal contribuição deste trabalho.

1.5 ORGANIZAÇÃO DO TRABALHO

Este trabalho está organizado da seguinte forma: no capítulo 2, apresenta-se uma breve revisão teórica dos principais algoritmos usados na detecção da forma humana e os utilizados neste trabalho. No capítulo 3, alguns trabalhos correlatos que utilizam os algoritmos abordados no capítulo anterior, na detecção de formas humanas, e artigos relacionados à detecção de pessoas em cenas de desastres para contextualizar a abordagem desta dissertação. No capítulo 4, a descrição do comitê de especialistas. No capítulo 5, os experimentos realizados e os resultados obtidos. O capítulo 6, a conclusão dos resultados e a discussão de propostas de possíveis trabalhos futuros.

2 REVISÃO TEÓRICA

Neste capítulo, será feita uma breve revisão da teoria dos principais métodos usados na detecção da forma humana e vistos nos trabalhos correlatos e os utilizados neste trabalho.

2.1 DESCRITORES DE ATRIBUTOS

Descritores de atributos são algoritmos responsáveis pela extração das características essenciais, que venham melhor representar um objeto em estudo. No caso deste trabalho o contorno da forma humana é a característica utilizada em todo o desenvolvimento.

2.1.1 Histograma de Gradientes Orientados

O Histograma de Gradientes Orientados, Histogram of Oriented Gradients (HOG) (DALAL; TRIGGS, 2005) tem, como objetivo inicial o reconhecimento de pessoas na posição vertical estática; porém, atualmente, há inúmeras variações deste algoritmo na detecção de inúmeros objetos e situações. O HOG tem como base o algoritmo de histograma de orientação de borda (FREEMAN et al., 1995), transformado de atributo de escala invariante, *Scale-Invariant Feature Transform* (SIFT) (LOWE, 2004), e contexto de forma (BELONGIE et al., 2001). Todavia, os histogramas são calculados em uma grade de células densas, uniformemente espaçadas, e para melhor desempenho, o HOG é realizado na superposição de blocos de células, com contraste normalizado.

A utilização do HOG sobre uma imagem é feita pela divisão dessa em pequenas regiões conectadas, conhecidas como células. Dentro dessas células, é então formado um histograma com as direções dos gradientes, ou seja, o descritor final é a concatenação desses histogramas. Como forma de melhorar o desempenho do descritor, menos susceptível à mudança de luminosidade, faz-se uma normalização do contraste local no histograma, calculando-se a medida de intensidade, por meio de uma larga região da imagem, conhecida como bloco, e usa-se esse valor para normalizar todas as demais células dentro do bloco.

Os passos para a computação do HOG seguem descritos adiante:

- a) computação do Gradiente: aplicar uma filtragem com máscara 1-D derivativa, centrada sobre a imagem nas direções verticais e horizontais, com o filtro:

$$[-1 \ 0 \ 1] \ e \ [-1 \ 0 \ 1]^T, \quad (1)$$

- b) retenção da Orientação: cada pixel dentro de uma célula fornece uma ponderação, magnitude do gradiente, individual para uma divisão do histograma orientado, que está baseado nos valores achados na computação do gradiente. O formato das células pode ser retangular ou radial e as divisões são uniformemente distribuídas entre 0° a 180° ou 0° a 360°, dependendo se o gradiente considera o sinal ou não. Nos experimentos (DALAL; TRIGGS, 2005), os gradientes sem sinal usados em conjunção com histogramas de nove divisões tiveram melhor desempenho na detecção de formas humanas;
- c) blocos Descritores: os blocos estão normalmente sobrepostos, o que significa que cada célula contribui mais do que uma vez na formação do descritor final. Existem dois formatos de blocos, o retangular R-HOG e o circular C-HOG. O R-HOG foi o que melhor apresentou resultados na detecção da forma humana (DALAL; TRIGGS, 2005). Como já mencionado, esse tem formato de grade retangular, representado por três parâmetros: o número de células por bloco, de pixels por célula e de divisões por histograma; ou seja, blocos formados por 3x3 células de 6x6 pixels, com histogramas de nove divisões;
- d) normalização dos Blocos: a normalização dos blocos pode ser feita através de uma das seguintes fórmulas:

$$L2 - norm : f = \frac{v}{\sqrt{\|v\|_2^2 + e^2}}, \quad (2)$$

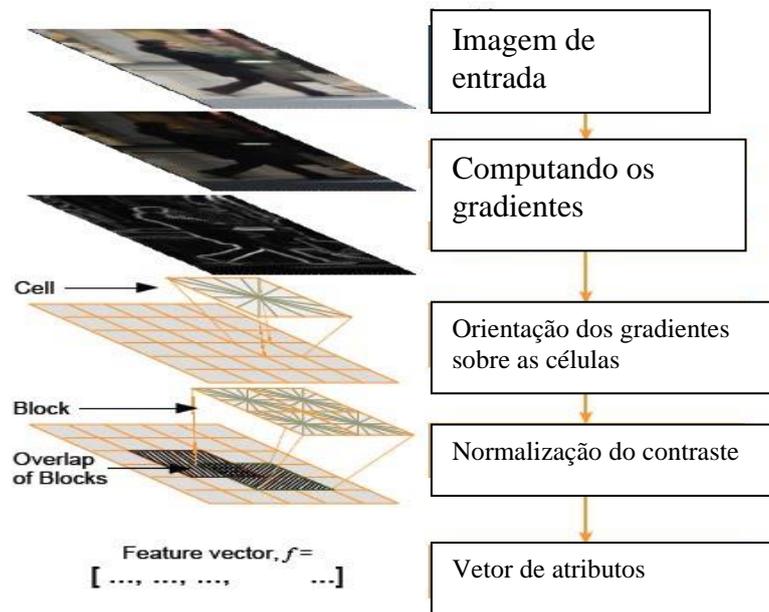
$$L1 - norm : f = \frac{v}{(\|v\|_1 + e)}, \quad (3)$$

$$L1 - sqrt : f = \sqrt{\frac{v}{(\|v\|_1 + e)}}, \quad (4)$$

em que v é o vetor de normalização que contém todos os histogramas de um dado bloco, $\|v\|_k$ é sua k-norm com $k = 1, 2$ e e é uma constante. Os passos para a computação do HOG é apresentada na Figura 2.

O HOG é um dos descritores mais eficientes para uso na detecção da forma humana, contudo, apresenta um alto custo computacional, pelo uso da sobreposição dos blocos descritores e pela normalização destes, logo não será usado neste trabalho.

Figura 2 – Passos para a computação do HOG.



Fonte: Autor “adaptado de” Pete Barnum, 2006.

2.1.2 Descritor Haar-Like

Um descritor *Haar-Like* (VILA; JONES, 2001) é uma ideia adaptada da característica *Haar wavelet*. Uma característica *Haar-like* considera regiões retangulares adjacentes de uma específica localização, em uma janela de detecção, como a soma de todas as intensidades dos pixels ali presentes em cada região e calcula a diferença entre elas. Essa diferença é, então, usada para classificar subseções de uma imagem. Por exemplo, no caso de detecção de faces humanas, é comum a observação da região dos olhos ser mais escura do que as das bochechas. Portanto, quando esse descritor é utilizado na detecção de faces com *Haar-Like*, são usados dois conjuntos de retângulos adjacentes, que se posicionam sobre a região dos olhos e das bochechas. A posição desses retângulos é definida pela posição da janela de detecção, que atua como uma caixa limitadora para o objeto alvo. Para cada subseção da imagem, um descritor *Haar-Like* é calculado. A diferença entre as somas, já definida, é comparada a um limiar, *threshold*, que separa o objeto do não-objeto. Pelo fato de o descritor *Haar-Like* ser insuficiente, ou seja, fraco para determinação da classificação, um grande número de características para determinar um objeto é necessário. Esse grande número de características, porém, impacta diretamente no tempo de processamento. A solução

encontrada foi o uso do conceito de Imagem Integral (CROW, 1984), utilizado em visão computacional no artigo “*Robust Real-Time Face Detection*”(VILA; JONES, 2001).

A definição de *Integral Image* é a soma de todos os pixels dentro de cada retângulo, isto é, uma nova representação da imagem. Segue adiante a relação matemática para a Imagem Integral e o par de recorrências usado para obter o resultado final.

Seja a imagem integral dada por $ii(x,y)$ (5),

$$ii(x, y) = \sum_{x' \leq x, y' \leq y} i(x', y'), \quad (5)$$

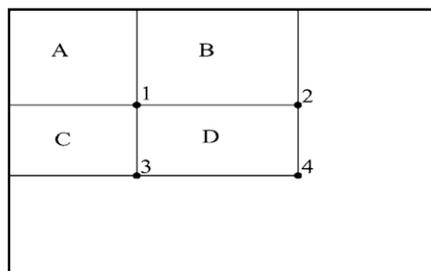
As equações (6) e (7) são pares de recorrências e $s(x, y)$ é a soma acumulativa da linha que, no início, é $s(x, -1) = 0$ e $ii(-1, y) = 0$. A Imagem Integral pode, então, ser calculada sobre a imagem original:

$$s(x, y) = s(x, y - 1) + i(x, y), \quad (6)$$

$$ii(x, y) = ii(x - 1, y) + s(x, y), \quad (7)$$

Na Figura 3, há um exemplo de como é feito o cálculo da integral no retângulo D. Primeiramente, é calculada a soma no ponto 1 (todos os pixels do bloco A). O ponto 2 é igual à soma dos blocos A e B. No ponto 3, a soma corresponde aos blocos A e C, e finalmente, o ponto 4 é a soma dos blocos A, B, C e D. Logo, a imagem integral no bloco D é igual a $4 + 1 - (2 + 3) \rightarrow A + B + C + D + A - A - B - A - C = D$:

Figura 3 - Imagem integral no bloco D.



Fonte: (VIOLA; JONES, 2003).

A ideia da imagem integral para acelerar o processo de contabilização do total de pixels em uma imagem é, aqui neste trabalho, utilizada para reduzir o tempo entre detecções no modelo de corpo completo.

2.1.3 Padrão Binário Local

O Algoritmo Padrão Binário Local, Local Binary Pattern (LBP) (OJALA; PIETIKÄINEN; HARWOOD et. al., 1994), foi originalmente desenvolvido para descrever e classificar a textura, pois é robusto a variações de iluminação na imagem, simples implementação, e tem alto poder de discriminação. Matematicamente, sua definição é:

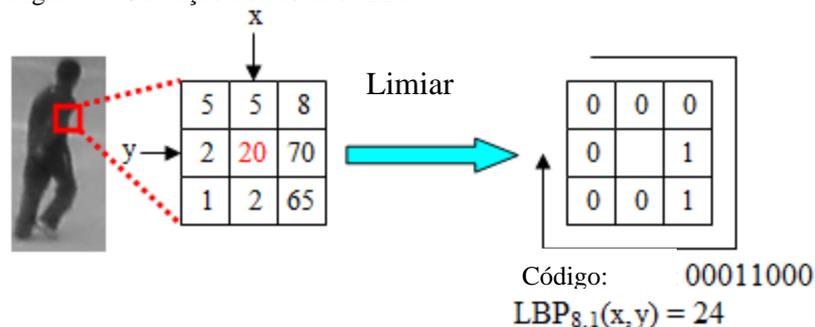
$$LBP_{P,R}(x_c, y_c) = \sum_{p=0}^{P-1} S(g_p - g_c) \cdot 2^p, \quad (8)$$

em que $c = (x_c, y_c)$ é o valor de LBP do pixel c na posição x_c e y_c ; P é o número do pixel amostrado (vizinho do pixel c), cuja distância não ultrapasse a R . Os valores g_p e g_c são as intensidades dos correspondentes pixels na posição p e do pixel central c . Então a relação S (pixel central c com seus vizinhos) é definida como:

$$s(x) = \begin{cases} 1, & \text{se } x \geq 0 \\ 0, & \text{ao contrário} \end{cases}, \quad (9)$$

Na Figura 4, há um exemplo de como é obtido o valor do LBP. O valor da intensidade do pixel central, $C = 20$, em vermelho, corresponde ao pixel retirado da imagem e tem a intensidade comparada com as intensidades dos pixels vizinhos $P \{1,2,2,5,5,8,70,65\}$, no raio R de vizinhança igual a 1. Caso o valor de P seja maior ou igual a C , o valor assumido por P será 1; caso contrário, será 0. O valor final do LBP é então calculado pelo movimento circular, de cima para baixo, e da esquerda para direita, e convertido para base 10. A partir dos valores dos $LBP_{P,R}$ é construído o histograma que tem 2^P divisões.

Figura 4 - Obtenção do descritor LBP.



Fonte: (Nguyen T. Ducet al., 2010).

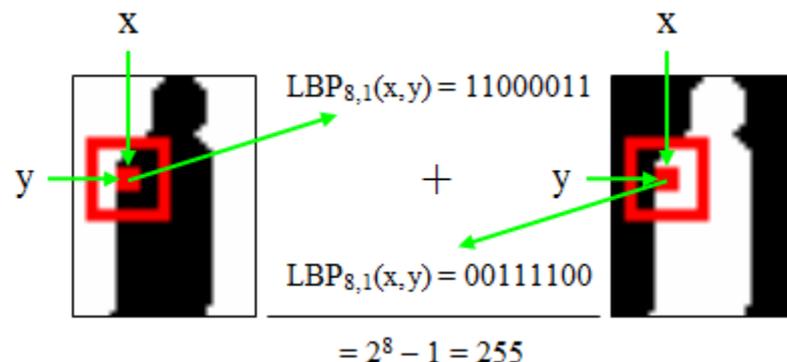
Pela descrição do LBP, seguem as características que podem levar a vantagens e desvantagens:

- caso haja a uniformidade definida como tendo o máximo de duas transições de mudança de bits, ou seja, de 0 para 1 ou de 1 para 0 dentro da sua representação binária circular, é assegurado maior detalhamento, o que leva a representar estruturas primitivas dentro de uma estrutura;
- caso não haja uniformidade, irão surgir ruídos inesperados, pois é menos discriminativo.

Embora o LBP seja bastante robusto para discriminar texturas no reconhecimento de objetos, não obteve um desempenho aceitável na captura da forma humana, já que apresenta algumas desvantagens:

- como os números de divisões no histograma são dependentes de 2^P , haveria a necessidade de uma grande quantidade de memória para armazenamento e posterior processamento;
- pelo fato de o LBP ser susceptível a mudanças nos bits, ou seja, a variações no plano de fundo e no primeiro plano, deve-se tomar cuidado com a região interna da forma humana, que depende da intensidade dos pixels locais e do vestuário utilizado. Na figura 5, há duas imagens que apresentam valores de LBPs diferentes, mas, representam também a mesma estrutura em situações de contrastes diferentes.

Figura 5 - Representação humana com contrastes diferentes



Fonte: (Nguyen T. Duc et al., 2010)

Diante dessas desvantagens, o Não-Redundante LBP (NRLBP) foi proposto, que tem como definição matemática:

$$NRLBP_{P,R}(x_c, y_c) = \min \{ NRLBP_{P,R}(x_c, y_c), 2^P - 1 - LBP_{P,R}(x_c, y_c) \}, \quad (10)$$

Intuitivamente, essa definição mostra que o NRLBP considera o LBP gerado e seu complemento com o mesmo valor. No exemplo mostrado anteriormente, o código binário “11000011” e seu complemento “00111100” são considerados como sendo iguais, ou seja, são contabilizados uma única vez. O que é obtido é um descritor mais robusto, discriminativo e adaptativo, que reflete as mudanças entre o plano de fundo e o primeiro plano. Somado a isso, também o tamanho do histograma no NRLBP é a metade do original LBP. Como o LPB foi desenvolvido para aplicações de textura e é mais suscetível a ruídos não optou-se pelo seu uso neste trabalho.

2.1.4 Centrist

O algoritmo CENSus TRansform HISTogram, (CENTRIST) (WU; REHG, 2011) é um descritor usado no reconhecimento da topologia de lugares ou cenas de uma forma holística, ou seja, global. Esse algoritmo tem como propriedade capturar e codificar a estrutura da imagem, suprime efeitos de textura, além de fácil implementação e não ser paramétrico. Seu funcionamento é baseado na comparação do valor de intensidade do pixel central em uma matriz 3x3 em relação aos seus oito pixels vizinhos. Caso o pixel central seja menor do que o vizinho a ser comparado, o vizinho assumirá valor 0 (zero). Caso contrário, assumirá valor 1 (um). Após essa etapa, os valores obtidos das comparações são posicionados em uma matriz também de mesma dimensão e organizados em uma sequência adotada a partir do canto superior esquerdo para o direito, e de cima para baixo, formando uma sequência binária. Essa, convertida para decimal, resultará no valor do CT entre [0 255] figura 6:

Figura 6 - CENTRIST: Matriz de transformação.

$$\begin{pmatrix} 32 & 64 & 96 \\ 32 & \mathbf{64} & 96 \\ 32 & 32 & 96 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 1 & 0 \\ 1 & \mathbf{CT} & 0 \\ 1 & 1 & 0 \end{pmatrix} \rightarrow (11010110)_2 \rightarrow CT = 214_{10}$$

Fonte: Autor.

Pode-se observar que o procedimento anterior se assemelha ao algoritmo LBP, mas o resultado é o inverso. Aqui também o CENTRIST é invariante à mudança de iluminação.

Uma importante observação na transformação como ilustra a, figura 7, é a propriedade de que os valores dos *CTs* são altamente correlacionados. Cabe observar as matrizes apresentadas, com pixels centrais 36 e 37, em que as transformações do algoritmo *Census Transform* (ZABIH; WOODFILL, 1994) são aplicadas. Ademais, que os bits em negrito são estados que dependem dos seus valores em pixels (36 e 37), porém, comparados em

ordens diferentes (35→31→32→33→**37** e 34→33→32→**36**). Desse modo, os dois bits em negrito devem ser complementares, se os seus valores em pixel forem diferentes; ou seja, o bit 5 do CT (x,y) deve ser complementar do bit 4 do $CT(x+1,y)$, se os pixels $(x,y) \neq (x+1,y)$. O exposto anteriormente é, então, propagado à toda rede de bits na imagem durante as comparações, propriedade transitiva, o que, implicitamente, contém informações para a descrição global da estrutura. Assim conclui-se que o número de pixels, em toda imagem, cujo bit 5 vale 1, deve ser igual ou maior do que os pixels cujo valor do bit 4 é igual a 0.

Figura 7 - CENTRIST: Correlação de valores

$$\begin{array}{c|cccc} 31 & 32 & 33 & 34 \\ 35 & \mathbf{36} & \mathbf{37} & 38 \\ 39 & 40 & 41 & 42 \end{array} \rightarrow \begin{array}{l} 1111\mathbf{0}000 \quad (36) \\ 1111\mathbf{0}000 \quad (37) \\ \text{bit } 12345678 \end{array}$$

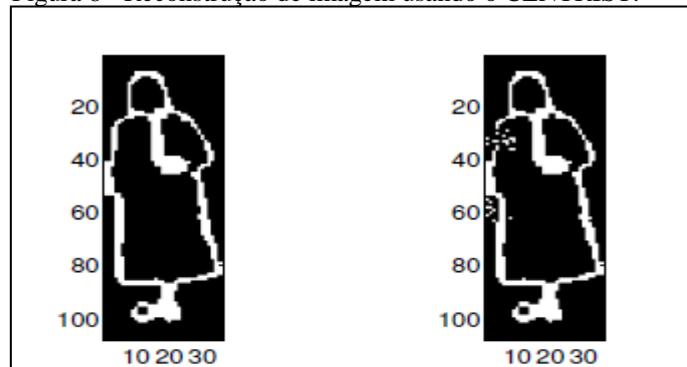
Fonte: Autor.

Após o cálculo de todos os CTs , usando uma matriz 3x3 e obtendo um valor entre [0 e 255] que corresponde a uma intensidade local, verifica-se então que há uma estrutura de valores indexada, porém, não há uma ordem entre esses valores.

A figura 8 é um exemplo de reconstrução de imagem de uma forma humana. A primeira imagem, da esquerda, representa a original, exemplo, a forma humana sobel. A imagem da direita é a imagem obtida a partir dos valores CENTRISTs, assim, retornando a imagem original.

Esse experimento mostra que a imagem é estruturada, codificada e que pode ser recriada a partir do CENTRIST. No entanto, fica a ressalva de que somente é válido para pequenas imagens, poucos pixels, e não para uma imagem ampla.

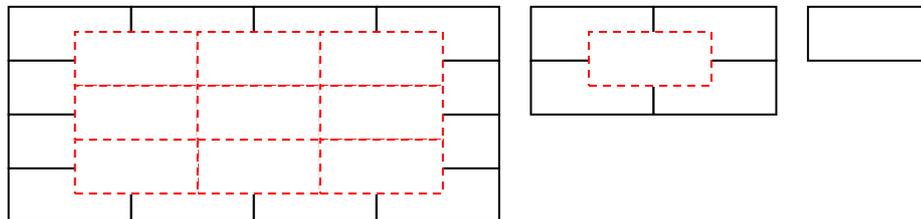
Figura 8 - Reconstrução de imagem usando o CENTRIST.



Fonte: Autor "adaptado de" CENTRIST Wu J., Rehg M., 2011.

Como forma de driblar a limitação do algoritmo CENTRIST, na codificação em imagens amplas, o conceito de Correspondência Espacial de Pirâmide é utilizado, dividindo a imagem em sub-regiões e integrando os resultados de correspondência nessas mesmas sub-regiões. Na figura 9, há um exemplo de pirâmide assim desenvolvida: o nível 2 é formado por uma grade com 16 blocos, divisões, seguido de nove blocos, em tracejado, deslocados para não sobrepor os anteriores. Esses nove blocos são necessários para garantir a captura dos pixels na região entre os 16 blocos que não se sobrepõem. Nos níveis 1 e 0, respectivamente, apresentam-se cinco e um blocos.

Figura 9 - Ilustração dos níveis 2, 1 e 0 de uma imagem dividida.



Fonte: Autor “adaptado de” CENTRIST (WU J. REHG M., 2011).

Durante a varredura da imagem sobre a janela de detecção, que contém a pirâmide, a imagem é redimensionada, de forma que sempre terá todos os seus pixels inclusos na pirâmide de blocos, em qualquer um dos níveis assumidos. O vetor de características será, então, formado pela concatenação de todos os CENTRIST extraídos de cada bloco em cada nível.

Consoante o exposto, se for alterada a ordem dos bits na matriz *Census Transform*, durante a comparação, ou seja, onde é 1, passa a ser 0, e vice-versa, tem-se o descritor LBP. Entretanto, a principal diferença entre os dois algoritmos é que o LBP deve trabalhar com valores uniformes, LBP uniforme, no processo de detecção de objetos, o que leva a agrupar os valores não uniformes. Como consequência, não se consegue fazer a recuperação da imagem, pois não é definida a estrutura global, devido à perda dos valores, causada pela não-uniformidade. Paratentar contornar o problema, é feita uma interpolação; porém, o resultado obtido é uma versão borrada da imagem do contorno original.

Pelas vantagens de simples operação, processar em um tempo linear, ou seja, $O(1)$ e também capturar a estrutura global de uma cena, no caso deste trabalho à forma humana, o algoritmo CENTRIST foi escolhido como o descritor de atributos.

2.2 CLASSIFICADORES

Classificadores são algoritmos usados para separação de dados de classes distintas. No caso deste trabalho, os dados são as características extraídas pelos descritores de atributos (item 2.1), e que representam as formas humanas em cenas de pós-desastres ou não, ou seja, as diferentes classes. Nesta seção são abordados todos os classificadores utilizados neste trabalho, e também a rede bayesiana como forma de associar os resultados concordantes, obtidos dos classificadores, e obter uma probabilidade de existência de um corpo em uma detecção.

2.2.1 Máquina de Vetores de Suporte

Máquina de Vetores de Suporte, Support Vector Machine (SVM) (VAPNIK; CHERVONENKIS, 1963), é um algoritmo de aprendizado supervisionado, usado para a classificação de dados e análises de regressão. O funcionamento desse algoritmo é baseado na formulação de um hiperplano separador de classes de dados; ou seja, dado um conjunto de dados de treinamentos, cada conjunto identificado por uma classe, o algoritmo computa um modelo, que classificará novos exemplos como sendo pertencentes a uma ou outra classe. Essa separação entre classes, margens entre o hiperplano, é o que define o quão eficiente é o algoritmo. O SVM pode operar com dados em um plano linear ou não-linear, em que esses são classificados em um espaço de característica de alta dimensionalidade, usando o truque do núcleo, *kernel*.

Mais detalhadamente, o algoritmo SVM cria hiperplanos em um espaço de alta dimensão, cujo objetivo é separar os dados de entrada de forma fácil e clara, isto é, que a distância seja a maior possível para os dados de treinamento muito próximos entre classes, função marginal. A função marginal, como já explanado, acima, deve ser a mais ampla possível, para obter uma boa separação e, conseqüentemente, um baixo erro de classificação. Os hiperplanos em espaço de alta dimensão são definidos com um conjunto de pontos, cujo produto escalar com um vetor naquele espaço é constante. Como os vetores são os que definem os hiperplanos, pode-se estabelecer uma combinação linear com parâmetros α_i de vetores de atributos x_i que ocorrem em uma base de dados. Na definição de um hiperplano, os pontos x no espaço de atributos desse são definidos conforme equação:

$$\sum_i \alpha_i k(x_i, x) = cte, \quad (11)$$

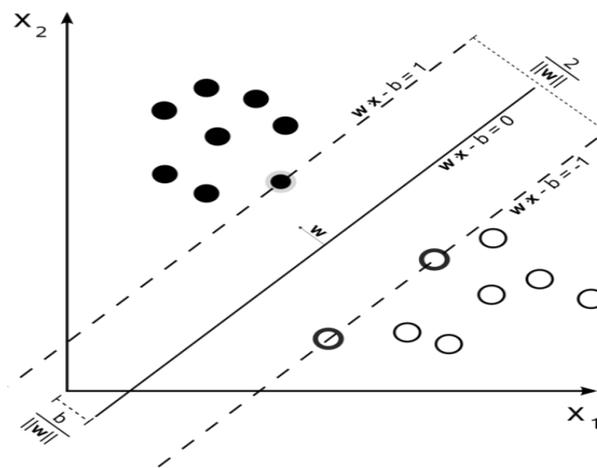
A relação anterior é usada para medir a proximidade de cada ponto de teste com os pontos da base de treinamento.

No SVM, qualquer hiperplano é definido como o conjunto de pontos x que satisfaça à equação:

$$w \cdot x - b = 0, \quad (12)$$

Em que o ponto $(.)$ denota o produto escalar e w o vetor normal para o hiperplano. O parâmetro $\frac{b}{\|w\|}$ determina o compensador (offset) do hiperplano a partir da origem ao longo do vetor normal w , conforme a figura 10:

Figura 10 - Exemplo de separação feita por SVM.



Fonte: Wikipédia, 2013.

$$w \cdot x - b = 1, \quad (13)$$

$$w \cdot x - b = -1, \quad (14)$$

As equações (13) e (14) representam os hiperplanos de separação dos conjuntos de pontos das classes lineares de treinamento, e a região entre os hiperplanos é chamada “margem”. Pelas relações de geometria, a distância entre os dois planos é definida como $\frac{2}{\|w\|}$. Deste modo, para maximizar a distância, deve-se minimizar $\|w\|$.

Para não permitir a entrada de pontos na região da margem, são então impostas as restrições (15) ou (16) a seguir:

$$w \cdot x - b > 1, \text{ para a classe 1,} \quad (15)$$

$$w \cdot x - b \leq -1, \text{ para a classe 2,} \quad (16)$$

Seja $\{x_1 \dots x_n\}$ um conjunto de dados de uma classe e $y_i \in \{1, -1\}$.

Assim, o limite de decisão deve classificar todos os pontos corretamente:

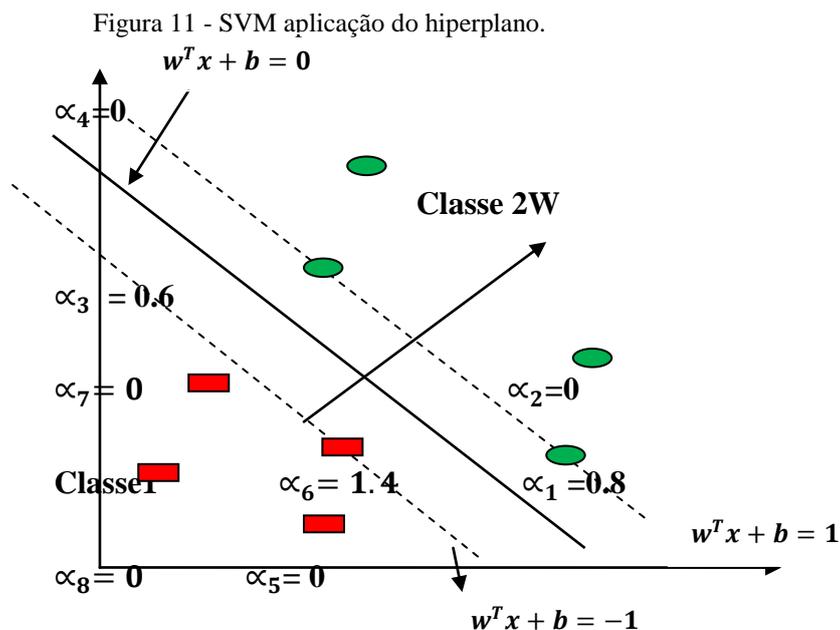
$$y_i(w \cdot x - b) \geq 1, \quad \forall_i, \quad (17)$$

Para a otimização do problema apresentado anteriormente, é difícil resolver, porque depende de $\|w\|$, ou seja, a norma de w , que envolve a raiz quadrada. Para facilitar, substitui-se $\|w\|$ por $\frac{1}{2}\|w\|^2$, conveniência matemática. Como w e b são os mesmos parâmetros nas duas formas, então, é um problema de otimização e programação quadrática em que minimize $\frac{1}{2}\|w\|^2$ sujeito a $y_i(w \cdot x - b) \geq 1, \forall_i$. Destarte, após a aplicação das técnicas de programação quadrática, tem-se (18):

$$w = \sum_{i=1}^n \alpha_i y_i X_i, \quad (18)$$

em que somente uns poucos α_i serão maiores do que zero e o correspondente X_i é o vetor de suporte, o qual deita sobre a margem e satisfaz $y_i(w \cdot x - b) = 1$, figura 11. E b é calculado como segue:

$$b = \frac{1}{N_{SV}} \sum_{i=1}^{N_{SV}} (w \cdot X_i - y_i), \quad (19)$$



Se abordado o problema com sua forma dual, obtém-se que a máxima margem entre os hiperplanos, isto é, a própria classificação, é somente função dos vetores de suporte

Ao considerar que $\|w\|^2 = w \cdot w$ e substituindo na equação (16) e maximizando em α_i , tem-se:

$$w(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1, j=1}^n \alpha_i \alpha_j y_i y_j X_i^T X_j, \quad (20)$$

sujeito a:

$$\alpha_i \geq 0, \sum_{i=1}^n \alpha_i y_i = 0, \quad (21)$$

em que:

$$X_i^T X_j = \text{kernel} = k(X_i, X_j) = X_i \cdot X_j, \quad (22)$$

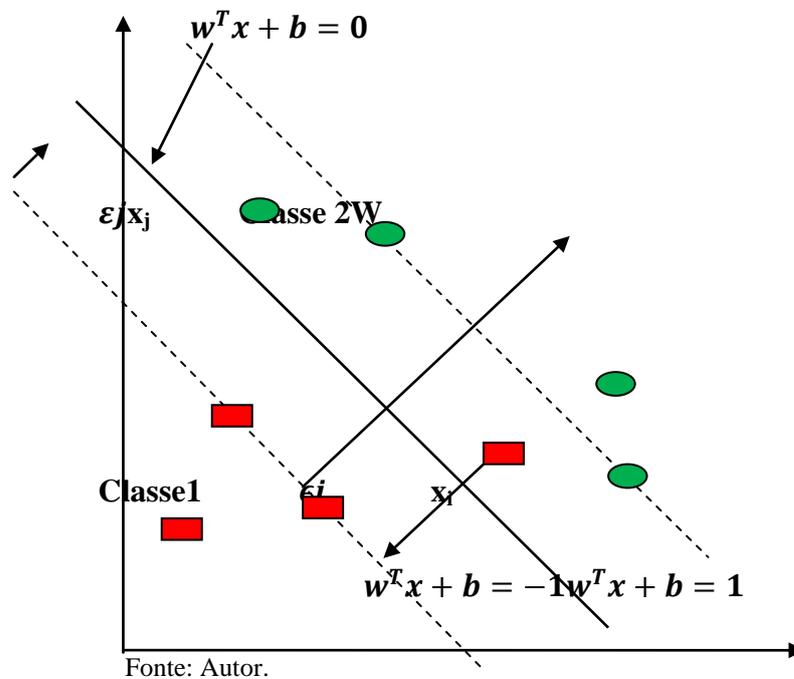
isso é um problema quadrático (QP), no qual os máximos globais podem sempre ser encontrados, e W pode ser calculado como em (18).

Hiperplanos com passagem na origem do sistema de coordenadas são chamados “*unbiased*” (23), enquanto os que não passam são chamados “*biased*”.

$$b = 0 \text{ e } \sum_{i=1}^n \alpha_i y_i = 0, \quad (23)$$

Caso não existam hiperplanos que dividam as classes, há um método de margem suave (VAPNIK; CORTES, 1995), que escolherá um hiperplano para separar e maximizar a distância. O método apresenta a variável ϵ_i , que mede o grau de erros de classificação dos dados x_i , figura 12.

Figura 12 - Interpretação do erro de classificação ϵ_i .



$$y_i(w \cdot X_i - b) \geq 1 - \epsilon_i \quad (24)$$

$$y_i(w \cdot X_i + b) \leq -1 + \epsilon_i \quad (25)$$

$$\varepsilon_i \geq 0, \quad (26)$$

No caso de $\varepsilon_i = 0$, não existe erro de classificação para X_i . Para valores de $\varepsilon_i \neq 0$, as funções anteriores são penalizadas. Por sua vez, a otimização é uma relação de ganhos e perdas entre largura de margem e erro de classificação pequeno. Caso o erro seja linear, a otimização é conforme segue:

$$\text{minimizar } \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n \varepsilon_i, \quad (27)$$

sujeito a $y_i(w \cdot X_i - b) \geq 1 - \varepsilon_i$, $\varepsilon_i \geq 0$. Da forma dual, deve maximizar-se:

$$w(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j k(x_i x_j), \quad (28)$$

Sujeito a $C \geq \alpha_i \geq 0$, $\sum_{i,j} \alpha_i y_i = 0$. As figuras 11 e 12 apresentam a interpretação dos conceitos, mas de forma geométrica.

2.2.2 Máquina de Vetores de Suporte Núcleo HI

O Núcleo de Intersecção de Histograma, Histogram Intersection Kernel (HIK) (SWAIN; BALLARD, 1991), é um dos núcleos usados com o SVM em medida de similaridade. Sua definição matemática é:

$$K_{HI}(x, x') = \sum_{i=1}^m \min\{x_i, x'_i\}, \quad (29)$$

Em que $x = (x_1, \dots, x_m)$ e $x' = (x'_1, \dots, x'_m)$ são dois histogramas de cores, com m divisões no espaço \mathbb{R}^m . Basicamente, conta a quantidade de número de pixels comuns de uma mesma cor entre histogramas de cores x e x' ; então, a soma global de todas as cores é normalizada pelo tamanho da imagem.

A propriedade fundamental do HIK é que ele aceita apenas valores positivos, ou seja, dados n exemplos $(x_1, \dots, x_m) \in \mathbb{R}_+^d$, e qualquer vetor $\alpha \in \mathbb{R}^n$ tem-se:

$$\alpha^T K \alpha \geq 0, \quad (30)$$

em que $K_{ii'} = K_{HI}(x_i, x_{i'})$,

isto é, dados n real números $(x_1, \dots, x_m) \in \mathbb{R}_+$, e qualquer vetor $\alpha \in \mathbb{R}^n$ tem-se:

$$\alpha^T K \alpha \geq 0, \quad (31)$$

em que $K_{ii'} = \min(x_i, x_{i'})$.

A prova (WU, 2010) é demonstrada adiante. Suponha que $0 \leq x_1 \leq \dots \leq x_n$, e então, definem-se:

$$R_{ij} = \begin{cases} 1 & \text{if } i = j \\ -1 & \text{if } i = j - 1 \text{ e} \\ 0 & \text{ao contrário} \end{cases} \quad \Lambda_{ij} = \begin{cases} x_1 & \text{if } i = j = 1 \\ x_i - x_{i-1} & \text{if } i = j > 1 \\ 0 & \text{ao contrário} \end{cases} \quad (32)$$

assim, tem-se que:

$$\Lambda = R^T K R, \quad (33)$$

desde que a matriz diagonal Λ seja semidefinitiva positiva e R é *rank* completo, K será também semidefinitivo positivo. Logo, é então definido o teorema da condição de valor positivo. Dadas n amostras $(x_1, \dots, x_m) \in \mathbb{R}^d_+$, um vetor $C \in \mathbb{R}^n$ com $\sum_{i=1}^n C_i = 0$, tem-se:

$$C^T K C \geq 0, \quad (34)$$

em que $K_{ii'} = K_{HI}(x_i, x_{i'})$.

quando há *bias* em SVM tem-se:

$$\sum_{i=1}^n \alpha_i y_i = 0, \quad (35)$$

Conclui-se que HIK pode ser usado com SVM desde que mantenha a propriedade (34). Neste trabalho, o HIK leva em consideração não a cor, e sim os valores dos *CTs* do *Census Transform* (ZABIH; WOODFILL, 1994).

2.2.3 AdaBoost

O algoritmo Adaptive Boosting, *AdaBoost* (FREUND; SCHAPIRE, 1995), foi o primeiro e, ainda hoje, é o mais conhecido e prático implementado de um algoritmo *Boosting* (KEARNS; VALIANT, 1989). A ideia do algoritmo *Boosting* está baseada na combinação de vários classificadores de aprendizagem fracos e com baixa exatidão, que, ao final, formarão um novo classificador, mas com maior exatidão na classificação.

Um dos grandes problemas na classificação ou reconhecimento de um objeto, no caso deste trabalho, a forma humana, é justamente o número de características ou atributos que irá determinar o tamanho do vetor de atributos. O AdaBoost, ao contrário do SVM ou da Rede Neural Artificial, durante o processo de treinamento do classificador, seleciona somente os atributos conhecidos por realmente proporcionarem um desempenho melhor durante a criação do modelo de predição. Com isso, além de diminuir a dimensionalidade do vetor de atributos, também reduz o tempo de processamento, o que é fundamental para a ocorrência prática no processo de reconhecimento.

O pseudocódigo do *AdaBoost* é apresentado no algoritmo 1: em um exemplo de treinamento $(x_1, y_1), \dots, (x_m, y_m)$ onde os x_i s estão em um domínio X , e os rótulos $y_i \in \{-1, +1\}$. Em cada passo de $t = 1, \dots, T$, uma distribuição é computada sobre os exemplos de treinamento, e um aprendizado fraco é realizado para achar uma classificação fraca. O objetivo do aprendizado fraco é achar uma fraca classificação com erro de baixa ponderação ϵ_t relativa à distribuição. A combinação final das classificações fracas, resulta em um sinal dessa combinação ponderada.

Algoritmo 1- Algoritmo AdaBoost.

Algoritmo 1: Pseudocódigo do *AdaBoost* .

- 1 Obtém N exemplos de imagens $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$, onde x corresponde a matriz de pixels de uma imagem, e $y = 0$, ou 1 para exemplos negativos e positivos, respectivamente.
 - 2 Inicializa os pesos $\omega_{1,i} = \frac{1}{2m}, \frac{1}{2l}$ para $y_i = 0$ e 1 respectivamente, onde m é o total de exemplos negativos e l o total de positivos.
 - 3 **para** $t=1, \dots, T$ **faça**
 - 4 Normalize os pesos: $\omega_{t,i} \leftarrow \frac{\omega_{t,i}}{\sum_{j=1}^N \omega_{t,j}}$, onde ω_t é uma distribuição de probabilidade.
 - 5 **para cada característica** j **faça**
 - 6 Treine um classificador h_j restrito ao uso de uma única característica.
 - 7 Avalie o erro de acordo com: $\omega_j, \epsilon_j = \sum_i \omega_i |h_j(x_i) - y_i|$.
 - 8 Escolha o classificador h_t com o menor erro ϵ_t .
 - 9 Atualize os pesos: $\omega_{t+1,i} = \omega_{t,i} \beta_t^{(1-e_i)}$, onde $e_i = 0$, se o exemplo x_i for classificado corretamente, $e_i = 1$ caso contrário, e $\beta_t = \frac{\epsilon_t}{1-\epsilon_t}$.
 - 10 O classificador forte é definido por: $H(x) = \begin{cases} 1, & \text{se } \sum_{t=1}^T \alpha_t h_t(x) \geq \frac{1}{2} \sum_{t=1}^T \alpha_t \\ 0, & \text{caso contrário.} \end{cases}$,
onde $\alpha_t = \log \frac{1}{\beta_t}$.
-

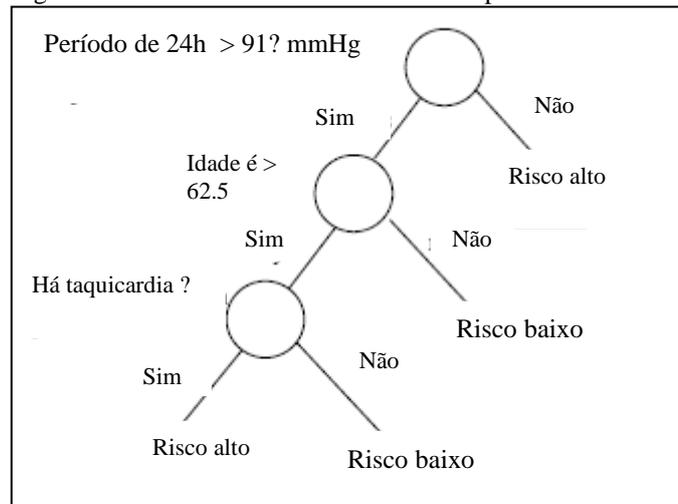
Fonte: Autor "adaptado de" ExplainigAdaBoost(SCHAPIRE).

2.2.4 Árvore de Decisão

Um dos classificadores mais populares conhecidos é o de árvore de decisão, *Decision Tree*, (DT) (BREIMAN et al., 1984), representado como uma partição recursiva da instância espaço. Basicamente, consiste de nós que formam uma árvore enraizada, *rooted tree*, e tem um nó principal de entrada, chamado raiz ou *root*. Todos os outros nós têm exatamente uma entrada, e o nó com uma saída, interno ou nó de teste. Todos os outros são chamados de folhas, ou nós terminais ou de decisão. Em uma árvore de decisão, cada nó interno divide o espaço dentro de dois ou mais sub-espacos, conforme uma função discreta dos valores atribuídos na entrada. No modo mais simples e frequente caso, cada nó teste considera um atributo individual, tal que o espaço é dividido de acordo com os valores dos atributos. No caso de atributos numéricos, a condição refere-se ao intervalo de entrada.

Cada folha é atribuída a uma classe, que representa o valor alvo mais apropriado. Alternativamente, a folha pode manter um vetor de probabilidade, indicando a probabilidade de o atributo de alvo ter certo valor. Exemplos são classificados pela navegação dos nós a partir do nó raiz da árvore para baixo, para uma folha, conforme o resultado dos testes ao longo do caminho. A figura 13 é a representação de uma árvore de decisão, em que pacientes são classificados como de alto ou baixo risco, com relação à pressão arterial sistólica.

Figura 13 - Árvore de decisão dos riscos da pressão arterial alta.



Fonte: “Adaptado de”, Wikipédia pressão arterial, 2015.

Ao observar a figura 13, o valor mínimo da pressão sistólica do sangue é monitorado nas primeiras 24 horas, e então, determina-se se ela está acima de 91 mmHg. Se a resposta for não, o paciente é classificado como de alto risco. Logo, não há necessidade de observar outras medidas para esse paciente. Se a resposta for sim, não se deve tomar uma decisão imediata. Nesse caso, a árvore de decisão deverá ser analisada em outros nós, ou seja, em pacientes com idade maior que 62,5 anos. Se a resposta for não, o paciente é classificado como de baixo risco. Entretanto, se o paciente está acima da idade, ainda não é possível determinar a decisão, e é verificado o próximo nó. Nesse nó, é avaliado se taquicardia sinusal está presente. Caso a resposta seja não, o paciente é classificado como de baixo risco, caso contrário, de alto risco.

Em casos de atributos numéricos, a árvore de decisão pode ser interpretada geometricamente como uma coleção de hiperplanos, cada um ortogonal ao outro em relação aos eixos. Cabe notar que o nível de complexidade da árvore tem efeito crucial na sua exatidão, e essa complexidade é explicitamente controlada pelo critério de parada, usado no método de poda empregado. Geralmente, a complexidade é medida por um dos seguintes parâmetros: número total de nós, de folhas, profundidade da árvore e número de atributos usados. No caso deste trabalho, o número de atributos corresponde ao tamanho do vetor de atributos.

O algoritmo árvore de decisão é definido como uma métrica para estimar a aleatoriedade dos dados em todos os nós da árvore, ou seja, a aleatoriedade da variável a prever (classe). Por exemplo, quando usada a regressão para ajustar uma função, pode-se usar

a diferença quadrática entre os valores verdadeiros e os previstos. Dessa forma, minimiza-se a soma das diferenças e os erros, em cada nó da árvore. Os rótulos das classes são definidos como uma medida mínima, quando há mais valores em um nó de uma mesma classe. Três medidas comuns usadas são: entropia, índice de *Gini* e erro de classificação (Bradski; Kaebler, 2008). Uma vez obtida uma dessas medidas, uma árvore de decisão binária busca, através do vetor de atributos, achar a característica, combinada com o limiar que mais minimiza a aleatoriedade dos dados. Por convenção, é adotado que as características acima do limiar são verdadeiras, ou seja, são classificadas e encaminhadas para o ramo esquerdo da árvore, caso contrário, para o ramo direito. Esse procedimento então é usado recursivamente para baixo na árvore em cada ramo, até os dados estarem suficientemente minimizados na aleatoriedade ou até o número de dados em um nó alcançarem conjunto mínimo.

A equação de minimização do erro para um nó, é a diferença dos quadrados entre o valor do nó y e o valor do dado x :

$$i(N) = \sum_j (y_j - x_j)^2, \quad (36)$$

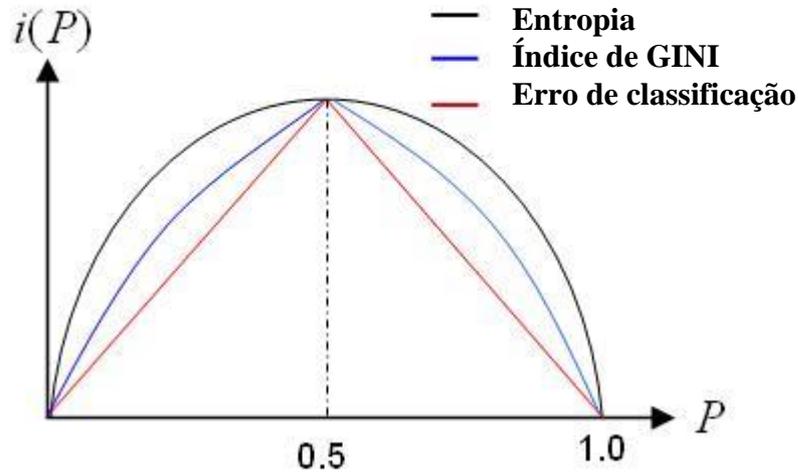
Para a classificação e como já mencionado, pode ser adotado um dos três métodos: entropia, índice de *Gini* e erro de classificação. Para esses métodos, é usada a notação $P(W_j)$ para denotar a fração padrão do nó N que está na classe W_j . Cada uma dessas medidas tem um leve efeito de diferença sobre a decisão de divisão; porém, todos têm como objetivo minimizar a aleatoriedade do nó, figura 14. A seguir, observam-se as definições matemáticas do exposto: entropia (37), índice de *Gini* (38) e erro de classificação (39).

$$i(N) = - \sum_j P(W_j) \log P(W_j), \quad (37)$$

$$i(N) = \sum_{j \neq i} P(W_j) P(W_j), \quad (38)$$

$$i(N) = 1 - \max P(W_j), \quad (39)$$

Figura 14 - Árvore de Decisão medidas de aleatoriedade.



Fonte: “Computer Vision”, Bradski; Kaebler, 2008, página 487.

2.2.5 Rede Neural Artificial

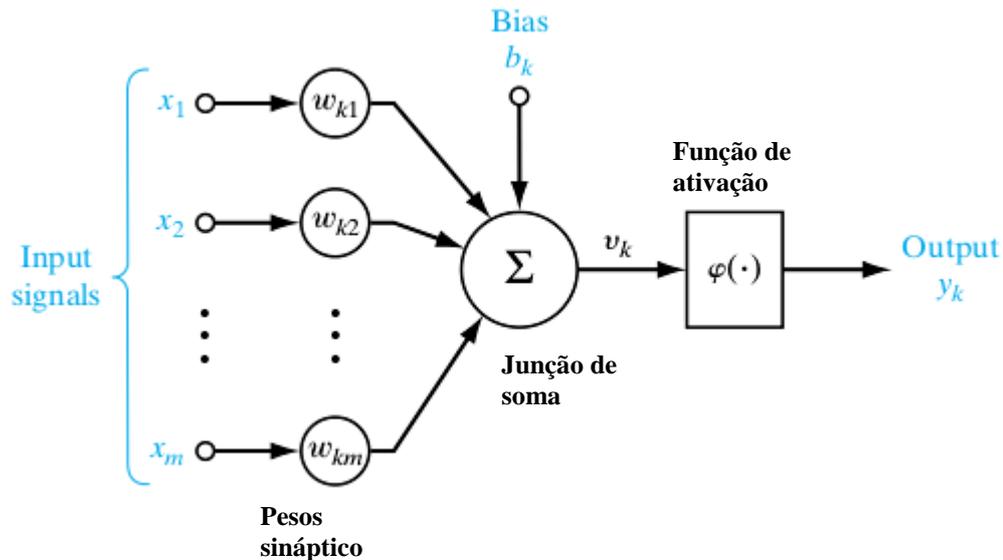
Rede neural artificial, *Neural Network*, (NNT) (MCCULLOCH; PITTS, 1943), tem sua concepção motivada no modo como o cérebro humano operacionaliza o reconhecimento, que é totalmente diferente dos computadores digitais. O cérebro é altamente complexo, não-linear, e computa de forma paralela. Sua estrutura de organização é constituída por neurônios, que reconhecem padrões, percepções e controle do movimento, de uma forma extremamente rápida e mais eficiente do que qualquer computador digital. Nesse sentido, o conceito de rede neural foi desenvolvido para modelar o jeito com o qual o cérebro executa uma particular tarefa ou função de interesse; a rede é usualmente implementada em um programa de aprendizagem e treinamento que, para, alcançar um bom desempenho, emprega massivas interconexões de simples células, chamadas de neurônios ou unidades de processamento. A partir do exposto, pode-se definir acerca da rede neural artificial:

- a) é um processo distribuído massivamente em paralelo, mas realizado por um só processador e que retém a informação para uso posterior;
- b) o conhecimento é adquirido pelo processo de aprendizagem;
- c) as interconexões com os neurônios, que, no cérebro, são conhecidos como sinapses, ocorrem na rede neural, por pesos sinápticos, que são usados para reter o conhecimento adquirido.

O procedimento usado para executar o processo de aprendizagem é chamado de algoritmo de aprendizagem, cuja função consiste em modificar os pesos sinápticos da rede de forma ordenada, de modo a atingir um objetivo desejado.

Na figura 15, é apresentada a unidade básica de uma operação de uma rede neural, ou seja, o modelo do neurônio, seguida de considerações explicativas.

Figura 15 - Representação básica de um neurônio.



Fonte: Haykin,2009, p.937.

O conjunto de conexões ou sinapses é caracterizado por um peso próprio. Especificamente, um sinal de entrada x_j é conectado à sinapse j , ponderado pelo peso w_{kj} , que por sua vez, é conectado ao neurônio k . Uma observação importante é que o peso sináptico da rede neural pode trabalhar em um intervalo, que vai desde valores negativos até positivos.

Os sinais ponderados das sinapses são encaminhados a um somador que faz uma combinação linear desses.

A função de ativação para limitar a amplitude da saída do neurônio é aplicada para manter a saída dentro do intervalo finito de trabalho. Os valores típicos normalizados deste intervalo estão entre $[0,1]$ ou $[-1,1]$. Observa-se também a aplicação de um sinal externo *bias* b_k , que tem como efeito acrescentar ou decrementar a entrada da rede da função de ativação, dependendo se é positiva ou negativa, respectivamente.

A representação matemática do exposto acima está nas equações seguintes: combinação linear u_k (40) e a saída do neurônio y_k (41).

$$u_k = \sum_{j=1}^m w_{kj} x_j, \quad (40)$$

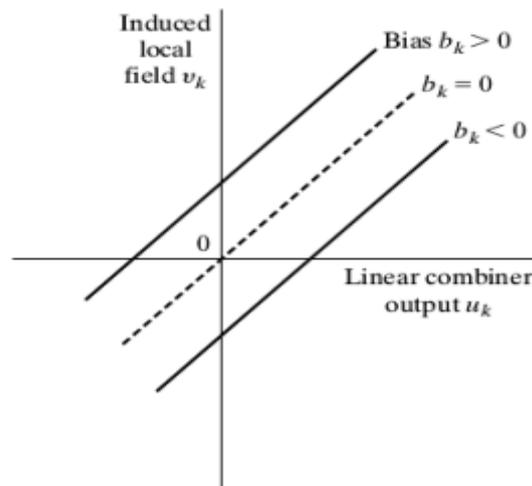
$$y_k = \varphi(u_k + b_k), \quad (41)$$

Em que x_1, x_2, \dots, x_m são as entradas de sinais, $w_{k1}, w_{k2}, \dots, w_{km}$ são os respectivos pesos das sinapses do neurônio K , u_k é a saída de combinação linear, devido aos sinais de entrada, b_k é o *bias*, $\varphi(\bullet)$ é a função de ativação e y_k é o sinal de saída do neurônio. O b_k (*bias*) tem o efeito de aplicar uma transformação afim para a saída u_k :

$$v_k = u_k + b_k, \quad (42)$$

a figura 16 é a representação gráfica do efeito de b_k (*bias*) sobre a saída.

Figura 16 - Transformação afim produzido pela presença de *bias*.



Fonte: Haykin, 2009, p.937.

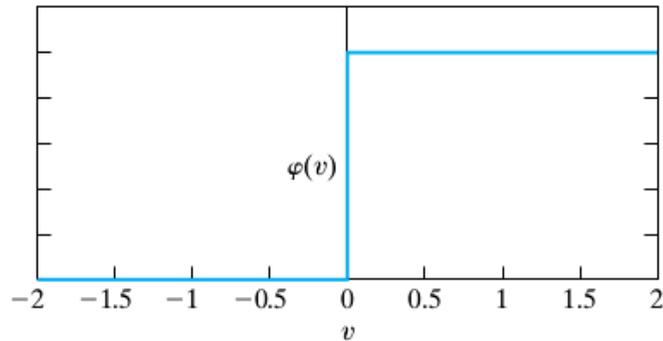
Pelas equações (40) e (42), é obtida a equação (43), e a partir da (41) e (43), obtém-se (44):

$$v_k = \sum_{j=0}^m w_{kj} x_j, \quad (43)$$

$$y_k = \varphi(v_k), \quad (44)$$

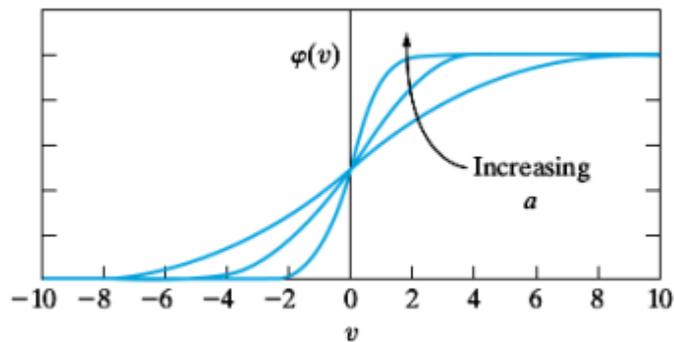
A função de ativação $\varphi(v_k)$ que define a saída do neurônio em termos de v_k é apresentada com dois tipos básicos figuras 17 e 18:

Figura 17- Função de ativação limiar.



Fonte: Haykin, 2009, p.13.

Figura 18 - Função de ativação Sigmóide.



Fonte: Haykin, 2009, p.13.

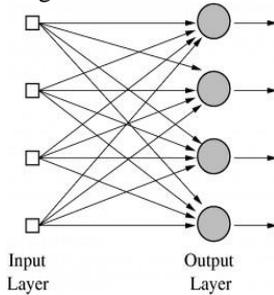
As funções de ativações são apresentadas como seguem: limiar (45) e Sigmóide (46):

$$\varphi(v) = \begin{cases} 1, & v \geq 0 \\ 0, & v < 0 \end{cases} \quad (45)$$

$$\varphi(v) = \frac{1}{1 + \exp(-av)}, \quad (46)$$

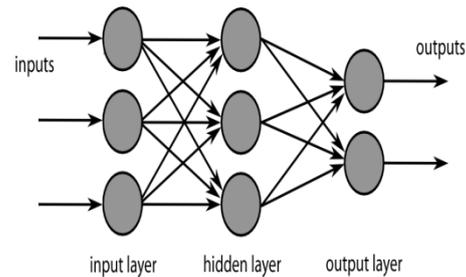
Existem três tipos básicos de arquitetura das redes neurais: simples camadas de alimentação direta, que são uma simples camada de neurônios figura 19; rede de multi-camadas de alimentação direta, com uma ou mais camadas escondidas (usadas neste trabalho) figura 20, e rede recorrente, que apresenta pelo menos uma realimentação figura 21.

Figura 19 - Rede Neural de simples camada.



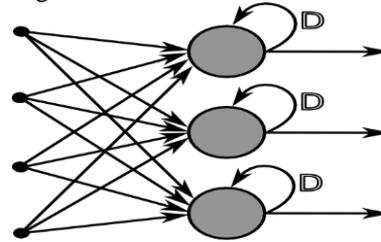
Fonte: Haykin,2009, página 22.

Figura 20 - Rede Neural de multi-camadas.



Fonte: Haykin,2009, página 22.

Figura 21 - Rede Neural recorrente.



Fonte: Haykin,2009, página 23.

2.2.6 Rede Bayesiana

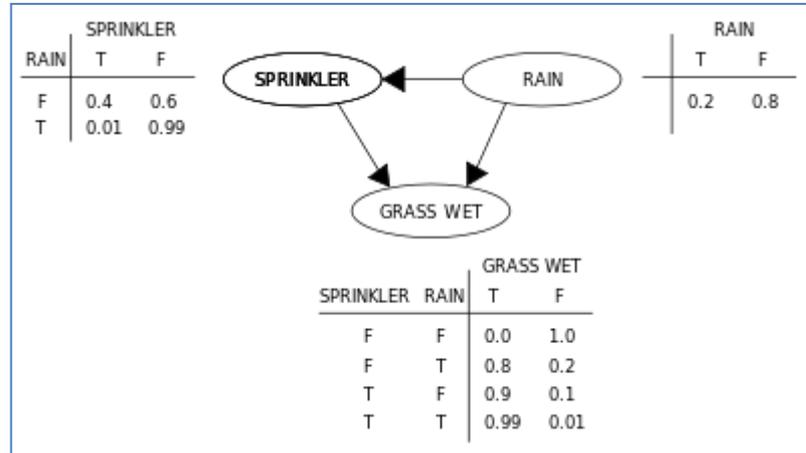
Rede Bayesiana (PEARL, 1988) figura 22, também conhecidas como redes de crenças, rede causais ou rede probabilísticas, são representações do conhecimento no domínio da incerteza. As redes Bayesianas são grafos acíclicos dirigidos (GADs), em que os nós representam variáveis aleatórias, e os arcos, as dependências condicionais entre as variáveis aleatórias. Essas variáveis aleatórias podem ser contínuas ou discretas. Neste trabalho, as variáveis são discretas. A rede Bayesiana é um grafo que apresenta então as seguintes características:

- um conjunto de variáveis que compõe os nós das redes;
- um conjunto de pares de nós conectados por arcos direcionados. O significado intuitivo de uma seta que parte de um nó X para um nó Y é que X tem uma influência direta sobre Y;
- cada nó tem uma tabela de probabilidade condicional (*conditional probability table*) (TPC), que quantifica os efeitos que os pais (*parents*) têm sobre os nós. Os pais de um nó X são todos aqueles nós que têm setas direcionadas para X;
- o grafo não tem ciclos direcionados, ou seja, é um grafo acíclico direcionado (*directed acyclic graph*) (GAD).

Uma rede bayesiana representa a distribuição de probabilidade conjunta (DPC) de tamanho exponencial, de uma maneira compacta. Cada entrada na DPC pode ser computada a partir da informação na rede bayesiana, pela regra da cadeia:

$$P(x_1, \dots, x_n) = \prod_{i=1}^n P(x_i | \text{Parents}(x_i)), \quad (47)$$

Figura 22 - Rede Bayesiana com a tabela de probabilidade condicional.



Fonte: “Adaptado de”, Guo ;Hsu, 2002.

A inferência em uma rede bayesiana é usada para achar uma atualização de conhecimento de um estado em um sub-conjunto de variáveis, quando outras variáveis, as evidências, são observadas, o que é chamado de inferência probabilística, ou seja, é a aplicação do teorema de Bayes. As inferências têm duas abordagens:

- a) inferência exata: $\begin{cases} \text{Enumeração} \\ \text{Eliminação de variável} \end{cases}$
- b) inferência aproximada: $\begin{cases} \text{Simulação estocástica} \\ \text{Cadeia de Markov métodos de Monte Carlo} \end{cases}$

Como a rede bayesiana é formada de variáveis em que arcos representam a relação de dependência condicional entre estas, foi então utilizado esta característica para estabelecer a dependência entre as variáveis nós pais, obtidas pela exatidão de detecção dos classificadores (SVM, rede neural artificial e árvore de decisão binária), e a presença de uma determinada parte do corpo. Na sequência o corpo, variável objetivo da detecção, estabelece relações de dependência entre as partes do corpo.

2.3 CONSIDERAÇÕES FINAIS

Ao longo deste capítulo, os métodos mais utilizados na extração de atributos da forma humana, em aplicações geralmente voltadas à captura de imagens de pedestres, foram descritos. Nessa breve descrição, também foram apresentados os modelos de classificadores empregados na identificação dos atributos, com seus pontos positivos e suas deficiências. A abordagem do atributo CENTRIST, foi apresentada como a solução utilizada neste trabalho e a rede bayesiana como fator de decisão da presença de um corpo foi então utilizadas como parte integrante do conjunto de informações da base teórica, para a melhor compreensão dos resultados obtidos nos experimentos realizados. No próximo capítulo, alguns trabalhos relacionados à detecção da forma humana em situações de desastres serão abordados.

3 TRABALHOS RELACIONADOS

A detecção de pessoas em uma cena tem sido um tema de grande desafio em muitas pesquisas; porém, têm ocorrido grandes avanços, principalmente quando as poses são fixas, como nos casos de detecção de pedestres, ou seja, posição vertical. Todavia, a detecção de pessoas em uma cena de desastre tem um nível de dificuldade superior, por alguns dos motivos a seguir relacionados:

- a) diversidade de poses: como o corpo humano tem uma flexibilidade grande, pode assumir qualquer posicionamento, direção, rotação e deformação, com variados pontos de vista em relação à observação;
- b) vestuário diversificado: não há um padrão de vestimenta humana, já que se pode revelar todo o formato do corpo, ou também, ocultá-lo completamente;
- c) oclusões diversas: pela própria situação da cena do desastre, qualquer objeto pode levar a oclusões parciais ou quase totais do corpo humano;
- d) condições da imagem: um dos pontos mais críticos é a condição de iluminação, que afeta diretamente a aparência do corpo humano.

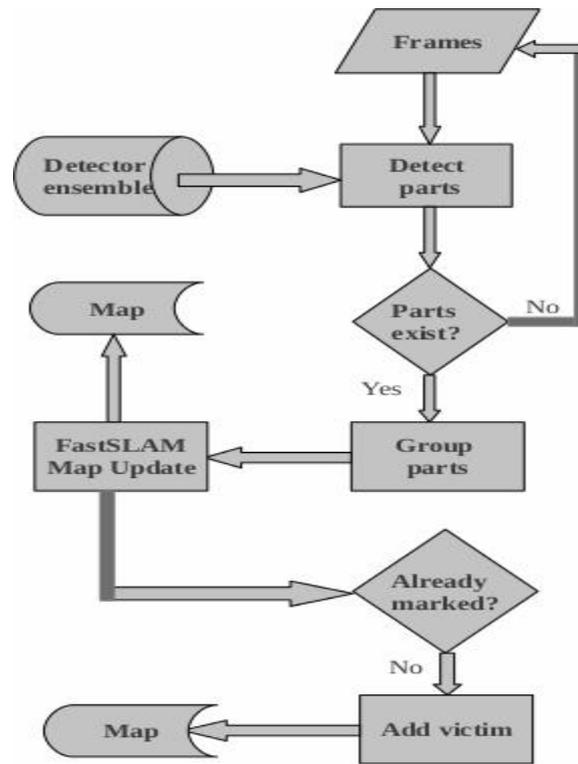
Neste capítulo, alguns trabalhos relacionados à detecção de vítimas de desastres serão apresentados, mas, pela diversidade do tema, nem todos são voltados a cenas do ponto de vista aéreo, ou seja, de cima para baixo.

3.1 DETECÇÃO E LOCALIZAÇÃO DE VÍTIMAS EM DESASTRES URBANOS

O artigo “*Victim Detection and Localisation in na Urban Disaster Site*” (SONI; SOWMYA, 2013) trata, como o próprio título revela, da detecção e localização de vítimas em desastres urbanos com o uso de robô. Neste trabalho, foi utilizado um descritor de características HOG com o classificador SVM e o Modelo de parte deformável, multi-escala, treinado discriminativamente (DPM) (FELZENSZWALB et al.,2008), para a detecção de partes do corpo humano. Quanto à localização da vítima, foi usado o algoritmo FastSLAM (MILSTEIN et al.,2011) para o mapeamento e rastreamento, enquanto a informação é disponibilizada e apresentada por meio de um guia baseado em interface do usuário.

Nos parágrafos que se seguem, são descritos de forma sintetizada os principais pontos dessa proposta, conforme o fluxograma da figura 23:

Figura 23- Implementação da proposta de Soni.



Fonte: Soni; Sowmya, 2013.

A proposta consiste nos seguintes blocos descritivos:

- a) o detector de vítimas: dois métodos de detecção são abordados; cima-para-baixo, que foca, primeiramente, na detecção do corpo completo, e depois, na localização das partes desse corpo com o uso algoritmo de modelo de parte deformável, multi-escala, treinado discriminativamente (DPM); e a abordagem, baixo-para-cima, que busca detectar as partes do corpo e, então, agrupá-las no detector de partes simples, *simple parts detector* (SPD) (BOURDEV; MALIK, 2009). O DPM trata o corpo humano como uma coleção de partes, tal que a raiz base é o corpo completo, para o qual todas as outras são conectadas no formato de uma estrela. Durante o treinamento, não é necessário rotular cada uma das partes, e sim, somente a raiz, e então, o algoritmo DPM cria um registro das regiões salientes recorrentes dentro da raiz. O detector de partes simples, SPD, é um conjunto de detectores de partes, em que cada detector é capaz de detectar uma parte específica do corpo humano (face, tronco e mãos). Cada parte é então modelada como uma classificação binária, através do algoritmo SVM, tendo como descritor de atributos o HOG. Uma

- b) vez que as partes do corpo são detectadas, o SPD agrupa todas, usando uma função heurística algoritmo 2;
- c) localização da vítima: a localização faz uso do algoritmo FastSLAM, que cria um mapa ao mesmo tempo em que um robô se desloca dentro da planta, ou seja, simultânea localização e mapeamento, *Simultaneous Localization and Mapping* (SLAM). Quando o robô detecta um objeto de interesse, a posição desse, relativa para cada partícula de localização, é transmitida para o algoritmo FastSLAM. Baseado na informação já presente no mapa, o sistema pode determinar se ele é uma nova vítima ou não. Caso essa informação, vítima, já tenha sido detectada anteriormente, então esta informação é considerada em duplicidade e rejeitada. A atualização do mapa ocorrerá se existirem outras vítimas localizadas além da distância de um metro da última localização.

Algoritmo 2 - Algoritmo de agrupamento. $P = Parent, N = Neighbour, T = threshold$.

```

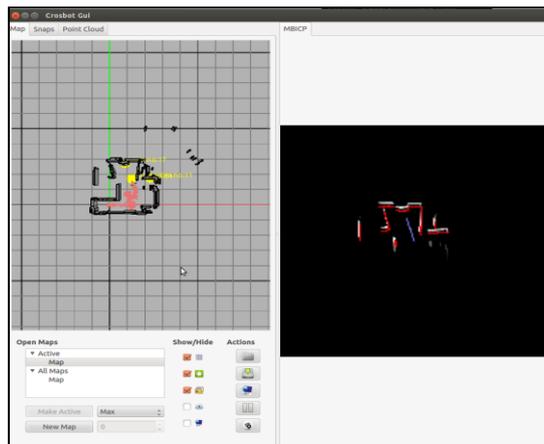
1: function GROUP (bodypart)
2:   while (part := bodyparts)! = NULL do
3:     P := part
4:     if P.group := NULL then
5:       P.group ← newGroup
6:     end if
7:     while (N := bodyparts)! = NULL do
8:       if distance(P,N) < T then
9:         if posAvail(P.group) then
10:          N.group ← Parent.group
11:        else
12:          N.group ← newgroup
13:        end if
14:      end if
15:    end while
16:  end while
17: endfunction

```

Fonte: Autor "adaptado de": Soni; Sowmya, 2013.

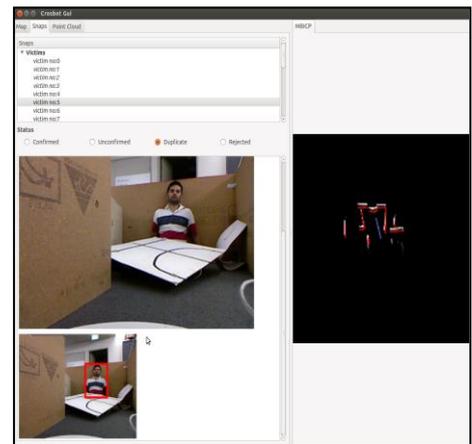
- d) componente da interface do usuário, HRI (KADOUS et al.,2006): é o responsável pela comunicação entre a equipe de resgate e o robô, ou seja, representa a informação recolhida pelo robô, via guia baseado em interface do usuário, que inclui mapas-guias intitulados *maps snap* (imagens de objetos conhecidos) e pontos de nuvens figuras 24 e 25.

Figura 24 - Visão do mapa na interface com o robô.



Fonte: Soni; Sowmya, 2013.

Figura 25 -A guia Snapscom a vítima.



Fonte: Soni; Sowmya, 2013.

Segundo o autor a taxa de falsos positivos durante o experimento produz um maior conteúdo de vítimas registradas do que realmente estavam presentes nas cenas simuladas de desastre. O autor propõe para o futuro uma redução nesses valores.

3.2 USO DO HOG NA DETECÇÃO DE VÍTIMAS EM ÁREAS DE DESASTRES

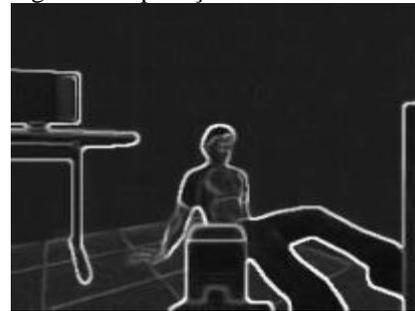
O trabalho de (UZUN et al., 2013) foi feito exclusivamente com o uso do algoritmo HOG em cenas de desastres, foi verificada a eficiência na detecção das várias poses e condições de iluminação. Segundo os autores, 95% de sucesso foi alcançado na detecção automática de vítimas em simulação de ambiente de tempo real, USARSim. Nas figuras 26 e 27, adiante, é apresentada a sequência da aplicação do algoritmo HOG em uma cena de simulação de desastre.

Figura 26-Imagem em escala de cinza.



Fonte: Uzun et al., 2013.

Figura 27-Aplicação do filtro Sobel.



Fonte: Uzun et al., 2013.

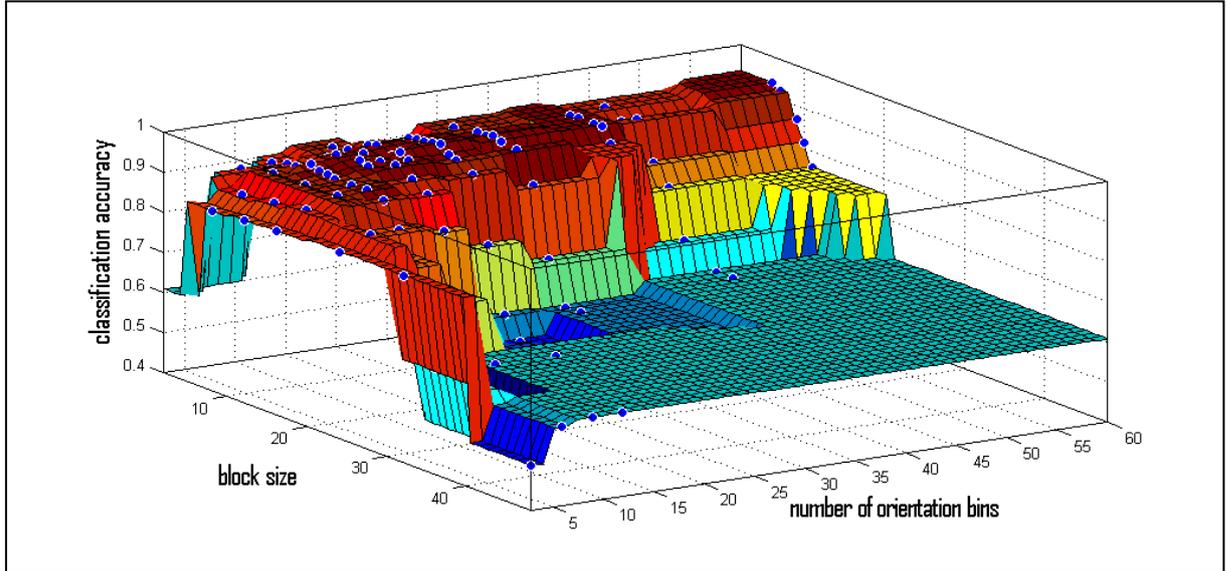
Neste trabalho, a superposição de blocos e a estrutura de varredura de janela não são usados, pois isso afeta o desempenho em tempo real. Principalmente, porque o método foi desenvolvido para trabalhar com múltiplos robôs na busca e operação de resgate, que atua em tempo real com as rotinas de localização e mapeamento.

A superposição de blocos aumenta não somente a quantidade de cálculos, mas também requer espaço de memória extra. O conjunto de imagens a ser analisado é variado em poses e condições diversas de iluminação, e o ângulo de abordagem do robô para a vítima cria variações. Se a vítima não se encaixar na janela de detecção, provavelmente, não será detectada pelo robô.

A classificação das imagens é feita pelo algoritmo SVM, com núcleo polinomial de segunda ordem. Nas figuras 28 e 29, em seguida, são apresentadas as exatidões de classificação para o conjunto de 100 diferentes representações do HOG usado nos testes experimentais. Na figura 28, a superfície de exatidão de classificação é apresentada. Cada ponto em azul mostra uma apresentação do HOG. O eixo X mostra o tamanho do sub-bloco, e o eixo Y, apresenta o número de orientações da divisão, e o eixo Z, o teste de exatidão de classificação.

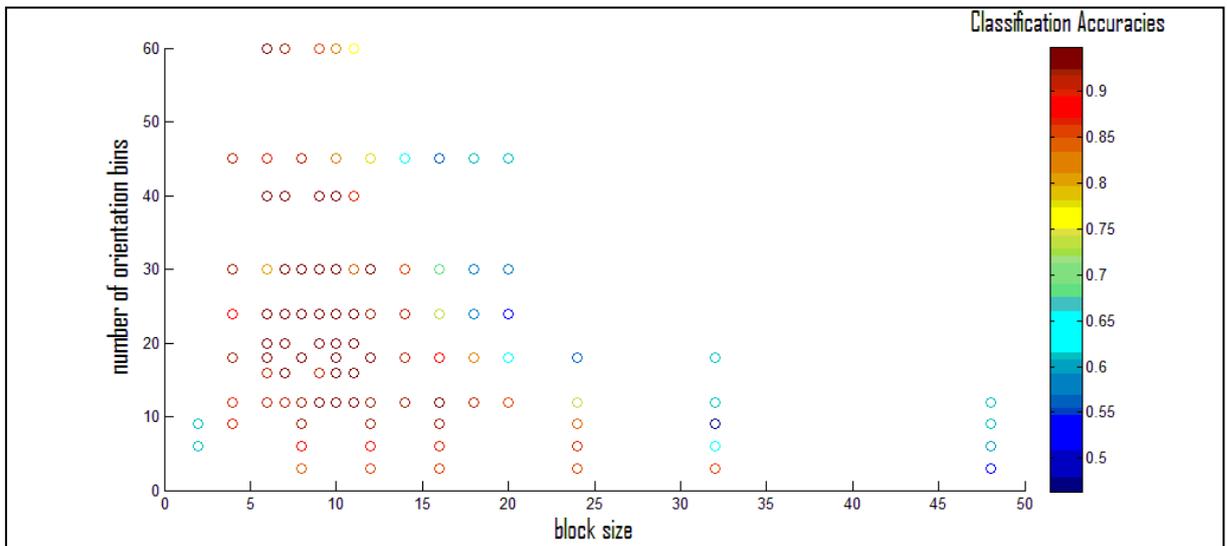
Na figura 29, cada círculo apresenta uma representação diferente do HOG. Os círculos coloridos mostram a exatidão da classificação.

Figura 28- Superfície de exatidão de classificação.



Fonte: Uzun et al., 2013.

Figura 29- Exatidão de classificação de 100 diferentes HOGs.



Fonte: Uzun et al., 2013.

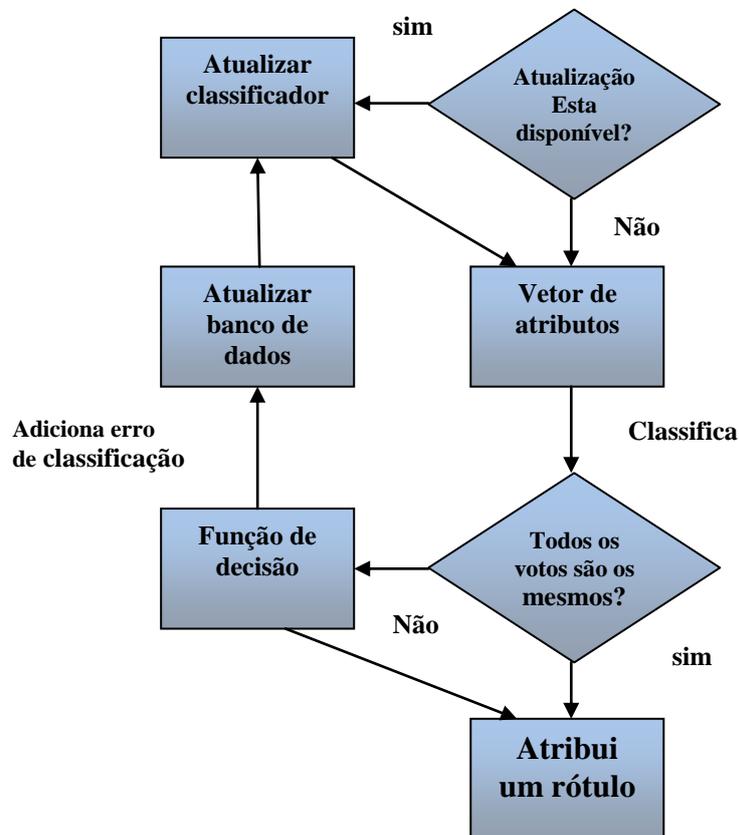
Os autores apresentam um resultado de exatidão de 95% . Todavia, foi testado em ambiente de simulação (USARSim), nada foi mencionado sobre as oclusões parciais ou quase totais; ou seja, não foi demonstrada a total eficiência desse método.

3.3 CONJUNTO DE CLASSIFICADORES COM APRENDIZADO INCREMENTAL PARA DETECÇÃO DE VÍTIMAS DE DESASTRES

O artigo “*Classifier ensemble with incremental learning for disaster victim detection*” (SONI; SOWMYA, 2012) aborda a detecção de vítimas de desastres em regiões urbanas pelo uso dos classificadores *AdaBoost*, *k*-vizinhos mais próximos (*k-nearest neighbours*) (*K-NN*) (FIX; HODGES, 1951) e SVM, em conjunto com a utilização de descritores de atributos como o HOG e o Atributo Robusto acelerado (Speeded Up Robust Features) (SURF) (BAY et al., 2006). O conjunto de classificadores usa tanto a votação por maioria (*majority voting*) quanto uma regra de decisão com base no histórico de classificação para determinar o resultado.

Uma novidade aplicada neste trabalho, a aprendizagem incremental, adquire um domínio de conhecimento enquanto realiza treinamento em paralelo, sem haver interrupção no processo de classificação. Segundo os autores, essa abordagem alcançou uma exatidão acima de 69%. Na figura 30, a seguir, observa-se um fluxograma de funcionamento do sistema proposto:

Figura 30- O sistema proposto por Soni.



Fonte: Soni; Sowmya, 2012.

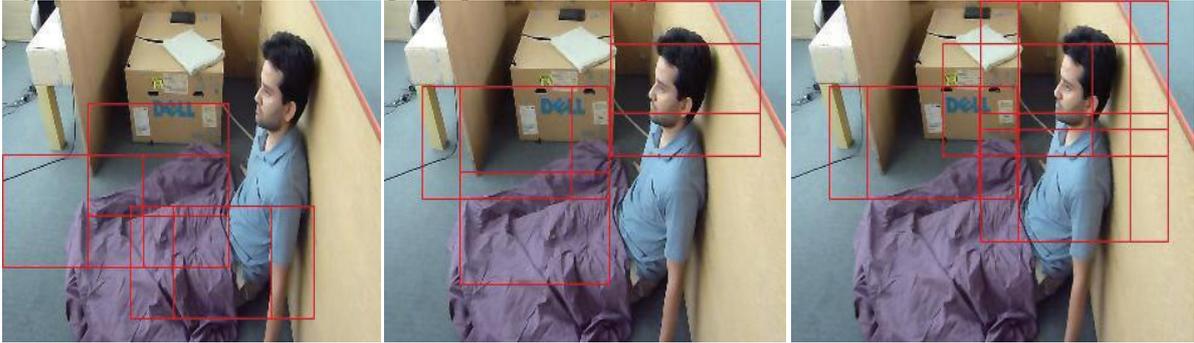
O funcionamento global da proposta figura 30 é apresentado nos passos a seguir:

- a) durante a classificação, cada classificador no conjunto vota em um rótulo de classe para um dado vetor de atributos;
- b) se todos os classificadores têm o mesmo voto, então, aquele rótulo de classe é usado como resultado;
- c) caso contrário, o histórico da maioria ponderada é usado por uma função de decisão para determinar o resultado;
- d) os classificadores que votaram contra o resultado da classificação o fizeram em um rótulo de classe errado, e são marcados para uma atualização;
- e) o vetor de atributos com classificação errada tem o nome do classificador e o rótulo da classe correta adicionados para uma atualização do banco de dados;
- f) a atualização do banco de dados continua a coletar os vetores com classificações erradas, até alcançar um específico limiar, *threshold*, sobre o qual uma atualização do classificador marcado é iniciada;
- g) antes de iniciar uma atualização, o sistema cria uma cópia do classificador, tal que o atual classificador pode continuar seu processo de classificação, enquanto sua cópia é atualizada;
- h) uma vez finalizada a atualização, o correspondente classificador no conjunto e sua cópia (atualizada), tanto quanto as amostras no banco de dados atualizado, são todos excluídos.

Além do classificador *K-NN*, atualizado com novos dados, o algoritmo SVM e o AdaBoost são atualizados por um novo treinamento, usando os dados originais do treinamento inicial, juntamente com os novos vetores de atributos. A qualquer momento, o número máximo de atualização no processo de classificação é o mesmo que o de classificadores no conjunto. Se uma atualização de um classificador está em curso e a atualização do banco de dados alcança o limiar, *threshold*, o dado atualizado é acumulado até o processo de atualização ser concluído.

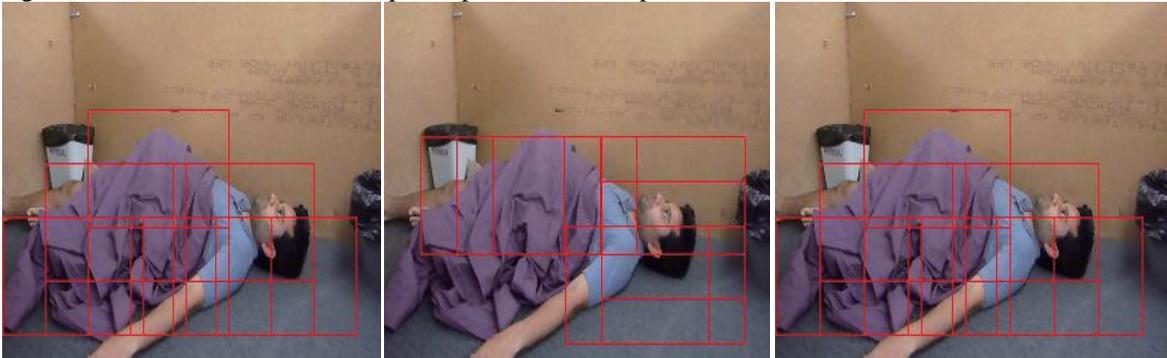
Os resultados obtidos dos experimentos são apresentados nas figuras 31 e 32, onde pode-se perceber que o número de janelas (quadro em vermelho) de detecções estão em maior quantidade sobre as figuras com aprendizagem incremental, IL, que foi a técnica que apresentou melhor resultado.

Figura 31- Resultados obtidos da esquerda para a direita respectivamente: HOG, SURF e HOG com IL.



Fonte: Soni; Sowmya, 2012.

Figura 32- Resultados obtidos da esquerda para a direita respectivamente: HOG, SURF e HOG com IL.



Fonte: Soni; Sowmya, 2012.

O melhor resultado final desse trabalho apresentou uma exatidão de 69% na configuração HOG-IL (aprendizado incremental), com taxa de FP igual a 18%.

3.4 DETECÇÃO DE VÍTIMA BASEADO EM VISÃO A PARTIR DE UM QUADRUTOR

O artigo “*Vision Based Victim Detection from Unmanned Aerial Vehicles*” (ANDRILUKA et al., 2010) trata diretamente do assunto abordado nesta dissertação, ou seja, não apenas aborda a detecção de pessoas em situação de desastres de forma geral, mas de pessoas deitadas sobre o chão em imagens tomadas em uma câmera integrada, *on-board*, de um veículo aéreo não tripulado, “*unmanned aerial vehicle*” (UAV).

Nesse artigo, os autores utilizaram várias abordagens de detecção de vítimas a partir de modelos de representação, baseados em partes flexíveis, e modelos esse fundamentados em treinamento de partes discriminativas (DPM), combinações de ambos e sensores de altitude e inclinação.

O descritor usado para extrair os atributos é o HOG, com as seguintes variantes:

- a) usado para treinamento do corpo humano completo;
- b) usado para treinamento na parte superior do corpo humano.

A detecção ocorre conforme modelos baseados em partes:

- a) detector de aparência de estrutura (PS) (FELZENSZWALB; HUTTENLOCHER; 2007). É o modelo de um objeto, que é determinado por um conjunto de partes com ligações entre certos pares de partes, desse modo, o objeto é representado como uma configuração de partes flexíveis e denotado por $L = \{I_0, \dots, I_N\}$, com I_i indicando a localização da parte i . Nessa formulação generativa, a posterior L (configurações), dada a evidência da imagem E , é obtida via regra de Bayes, $p(L|E) \propto p(L) p(E|L)$. Logo, a configuração posterior é:

$$p(L|E) \propto p(I_0) \cdot \prod_{i=0}^N p(E|I_i) \cdot \prod_{(i,j) \in G} p(l_i|l_j), \quad (48)$$

A propagação da crença da soma do produto é aplicada de modo a computar a posterior marginal do tronco humano, $P(I_0|E)$, que é, então, usado para delimitar a caixa limite de detecção. O modelo de aparência de estrutura, PS, utiliza partes superiores e inferiores do corpo humano, correspondentes a pernas, braços, tronco e cabeça, além de exemplos rotulados para treinamento. Os termos $p(E|I_i)$ são representados como parte discriminativa do classificador treinado com AdaBoost, e os termos $p(l_i|l_j)$ são estimados com a máxima probabilidade, usando a parte rotulada fornecida;

- b) o modelo de parte deformável, multi-escala, treinado discriminativamente (DPM) trata o formato do corpo humano como uma estrela, em que todas as partes são diretamente conectadas por uma raiz, ou seja, é o corpo completo. A diferença entre o modelo PS e o DPM é que, este, descobre automaticamente as partes do corpo, enquanto que PS depende de anotações manuais de rotulagem;
- c) o detector baseado em pose ou poselet-based (PBD) (BOURDEV; MALIK, 2009), depende de um largo número de detectores de partes para diversas regiões denotadas como “*poselet*” e que tem consistente aparência e correspondência similar para a configuração 3D do corpo humano. Para a detecção de diferentes *poselet*, são usados os procedimentos de votação probabilística, que remontam ao implícito modelo de forma com aprendizado ponderado, através da forma de máxima margem.

Com vistas a minimizar os inúmeros efeitos de oclusões, uma combinação dos modelos baseados em partes foi implementada. O DPM foi usado porque detecta a parte superior do corpo e o corpo completo. Também, o detector PS atua de forma similar e complementar (tronco, cabeça, braços e pernas). A fusão desses diferentes modelos é computada como a probabilidade posterior de cada hipótese k , dada a pontuação da detecção d_k do modo M como:

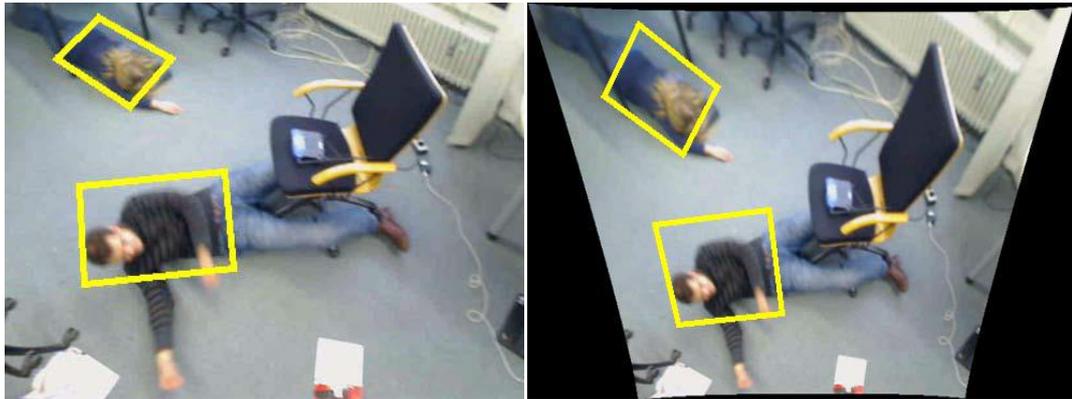
$$p(h_k | d_k, M) = \frac{p(d_k | h_k, M)}{p(d_k | h_k, M) + p(d_k | \neg h_k, M)}, \quad (49)$$

Em que h_k é uma variável Booleana correspondente para k -th hipótese, indicando se ela é correta ou incorreta. O termo $p(h_k | M)$ cancela, quando é assumido ser uniforme. A distribuição condicional $p(d_k | h_k, M)$ e $p(d_k | \neg h_k, M)$ é assumida ser Gaussiana e ajustada sobre um conjunto de detecções positivas e negativas.

O método de detecção de pessoas discutido é operado sobre a suposição de que a posição da câmera e a profundidade para cada pixel da imagem são desconhecidas. Isso implica que não está disponível a informação a priori sobre a escala das pessoas na imagem, e cada modelo precisa ser exaustivamente avaliado sobre todas as possíveis escalas. No entanto, o autor utilizou um *quadrotor* na aquisição das imagens, equipado com câmera calibrada, e sensores capazes de medir a altura e o ângulo do veículo. Isso permite, pela combinação desses sensores, determinar a distância do plano do solo para cada pixel da imagem. Desse modo, sabendo-se a posição da câmera com respeito ao plano do solo, pode-se projetar, de

volta, uma imagem sobre o plano do solo, considerando as transformações homográficas e as distorções da imagem. A figura 33 apresenta a projeção; nota-se que enquanto a escala das pessoas difere em imagem original, depois da projeção de volta, ela se torna aproximadamente a mesma.

Figura 33- Imagem original à esquerda e a imagem após a projeção de volta à direita.



Fonte: (ANDRILUKA et al., 2010).

Em adição à calibração da câmera e à projeção de volta, é habilitada a relação de altura da janela limite de detecção e a medida em pixels para a altura da pessoa medida em metros, o que permite definir a distribuição a priori sobre a altura da janela limitadora. O resultado comparativo entre os vários detectores é visto na figura 34. Na figura 35, o gráfico de desempenho dos modelos com e sem escala a priori, onde observa-se que o modelo PS, a priori, obteve uma melhoria de 16% EER, do que para o modelo DPM. E na figura 36, o gráfico de avaliação entre os diferentes modelos e combinações, observa-se que o melhor desempenho foi obtido pela combinação do corpo inteiro usado PS e parte superior do corpo também usado PS e DPM, o qual à combinação alcançou 66% ERR.

Figura 34 - Resultados de várias detecções obtidas dos modelos apresentados (amarelo *true positive* (+) e vermelho falso positivo (-)).

HOG



Poselets'



DPM



PS



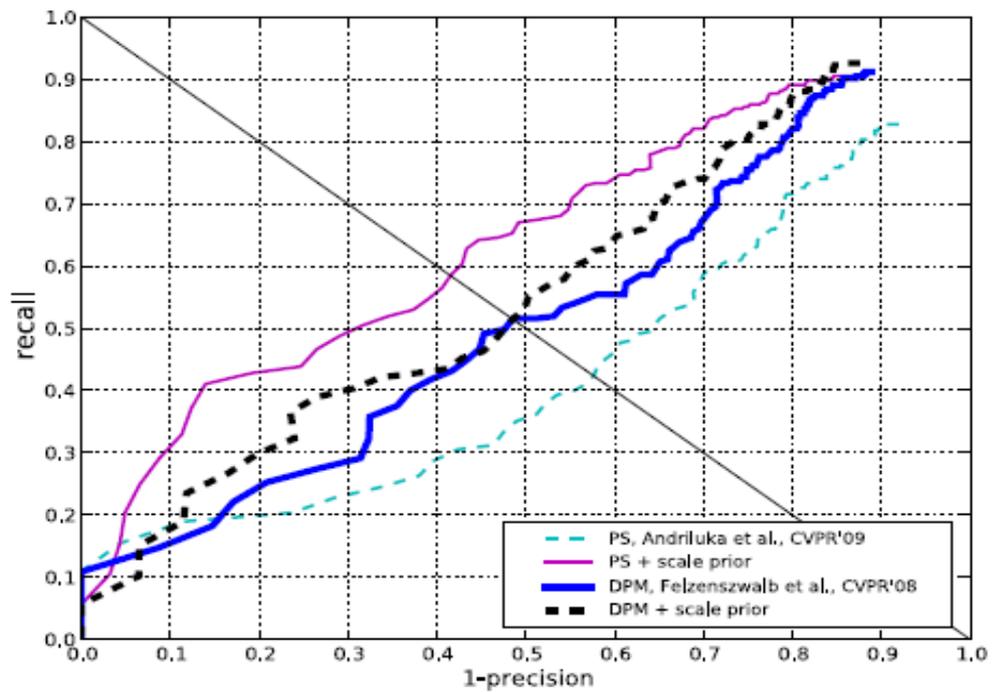
PS + DPM



Fonte: AndrilukAet at., 2010.

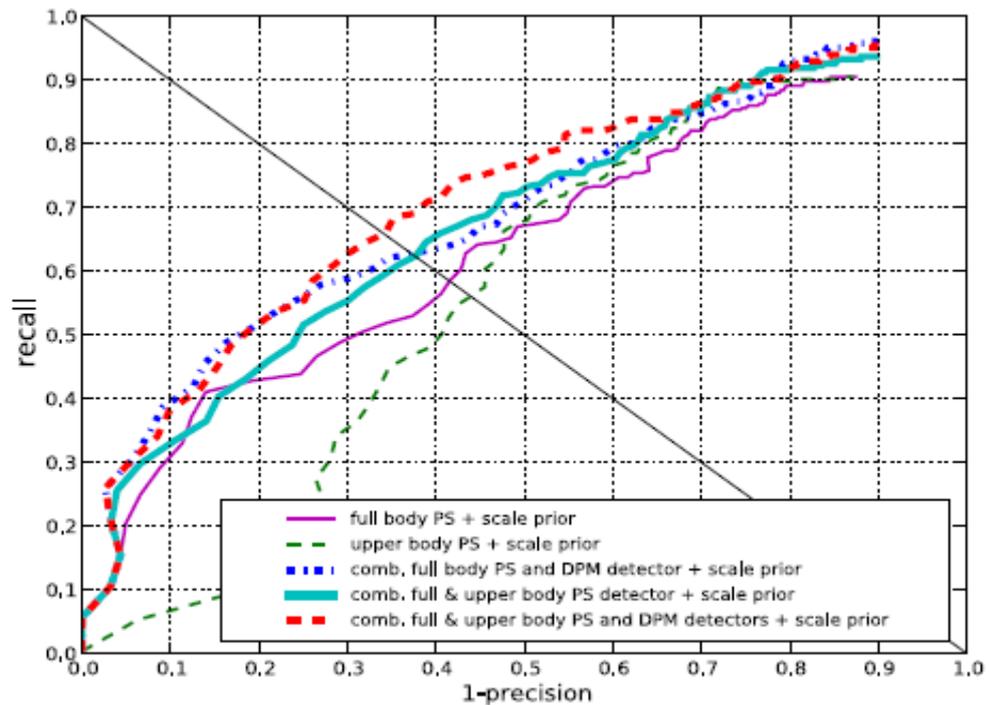
A conclusão obtida foi que os modelos baseados em partes têm um melhor desempenho para a detecção de vítimas em cenários de desastres, ou seja, PS+DPM, ver figura 34, do que modelos monolíticos (forma completa do corpo humano), ver figura 34 HOG, porque são adequados para representarem as variações em articulações e robustos para oclusões parciais. Neste artigo, nada foi mencionado sobre o funcionamento do sistema como um todo, em tempo real.

Figura 35- Gráfico de desempenho dos modelos com e sem escala a priori.



Fonte: Andriluka et al., 2010.

Figura 36- Avaliação entre os diferentes modelos e combinações.



Fonte: Andriluka et al., 2010.

3.5 CONSIDERAÇÕES FINAIS

A detecção de vítimas em um cenário de desastre é uma área relativamente nova de estudo, principalmente em cenas do ponto de vista aéreo, ou seja, de cima para baixo. Neste capítulo, foram detalhados alguns dos trabalhos que abordaram os vários desafios que transcendem o já bastante discutido problema de detecção de pedestres. Mesmo neste caso, ainda há muito a se fazer em termos de oclusões e funcionamento em tempo real.

O desafio de detecção de pessoas em cenários de desastres envolve a flexibilidade das articulações humanas em toda a sua plenitude, o vestuário diversificado, as condições variadas de iluminação e sombreamento, sem mencionar as oclusões, devido aos escombros e superposições de corpos. A abordagem com os tradicionais algoritmos usados na representação monolítica de pessoas é limitada diante de tais cenários complexos, o que pode ser visto nos pontos principais dos artigos apresentados. Assim, são recriadas combinações diversas de algoritmos e, principalmente, métodos baseados em modelos de partes do corpo humano, visando a obter um melhor resultado de desempenho em detecção e funcionamento em tempo real. Não há, portanto, uma solução única e definitiva. Esse conjunto de informações apresentadas forma a base teórica da proposta desta dissertação. No próximo capítulo, o modelo utilizado no desenvolvimento deste trabalho será apresentado em detalhes.

4 DETECÇÃO DE FORMAS HUMANAS EM CENAS AÉREAS DE DESASTRES PELO USO DE COMITÊ DE ESPECIALISTAS E REDE BAYESIANA

A abordagem proposta nesta dissertação é voltada para a detecção da forma humana em cenas de desastres com o uso de comitê de especialistas (Máquina de vetores de Suporte (SVM), árvore de decisão binária (Dtree) e rede neural artificial (NNT)), explicado no item 2.2 (Classificadores), o qual é responsável pela detecção de partes do corpo humano (cabeça, pernas e braços), e uma rede bayesiana para decisão se tais partes representam a existência de um corpo em uma imagem. Como já mencionado no item 2.1 (Descritores de atributos), o algoritmo CENTRIST foi escolhido como o descritor deste trabalho, pois, captura o contorno humano e elimina a textura, tem fácil implementação e trabalha em tempo linear $O(1)$. Esse atributo é a única informação fornecida aos vários classificadores, seja para captura da envoltória completa da forma humana, ou das partes específicas (braços, cabeça e pernas).

Há dois conceitos de classificação empregados neste trabalho: o primeiro é o de classificação em cascata, tal qual o algoritmo Adaboost, aqui utilizado para detecção do corpo completo. O segundo é de comitê de especialistas que atua em paralelo na classificação de uma mesma parte do corpo; por exemplo: a cabeça. Objetivo desse comitê é aumentar a confiabilidade de que uma amostra “observada” em um dado instante, janela de detecção, faz parte do conjunto de treinamento (cabeça, pernas e braços). A combinação dos resultados desses classificadores em paralelo é avaliada por uma função de corte, que determina um limiar de permissão de análise pela rede Bayesianapara tomada de decisão final.

4.1 AQUISIÇÃO DA IMAGEM E TRANSFORMAÇÕES

A figura 37 apresenta o fluxograma do algoritmo proposto neste trabalho. No início a imagem I na entrada (E) é tratada com uma rotina que suaviza a intensidade dos pixels com uma função Gaussiana e extrai o gradiente com um filtro Sobel (Sobel; Feldman, 1968), onde um importante parâmetro é a ponderação da magnitude dos gradientes nas direções G_x e G_y (bidimensional) da imagem I , pois, acrescenta ou reduz a quantidade de informação a ser analisada, afetando diretamente a velocidade de processamento das janelas. Este parâmetro também atua na sensibilidade do limiar de trabalho dos classificadores.

Pode-se observar pelas figuras 38, 39 e 40 que, ao decrementar o valor de α (valor adimensional), a informação também é reduzida, ou seja, a image tem menos detalhes, logo:

$$abs(G_x) * \alpha E \quad e \quad abs(G_y) * \alpha E, \quad (50)$$

O algoritmo CENTRIST é então aplicado sobre a imagem I , a qual foi usado o filtro Sobel. Uma janela de detecção de 80 x 80 pixels é então deslocada por toda imagem, gerando um vetor de atributos de dimensionalidade de 6400 pixels. A escolha do tamanho e forma da janela de detecção, 80 x 80 pixels, tem haver com a disposição do corpo em uma cena de desastre, pois, diferente da detecção de pedestre onde a posição vertical é predominante e a forma retangular da janela é adequada, em um desastre o corpo não tem disposição definida, e uma janela de formato equilibrado é o mais adequado, ou seja, um quadrado. A saída (O), do módulo de transformação CENTRIST, é então aplicada ao módulo de detecção de corpo completo o qual é constituído de uma cascata de classificadores, SVM-linear e HIK, tendo como alternativa o real adaboost no lugar do SVM-HIK.

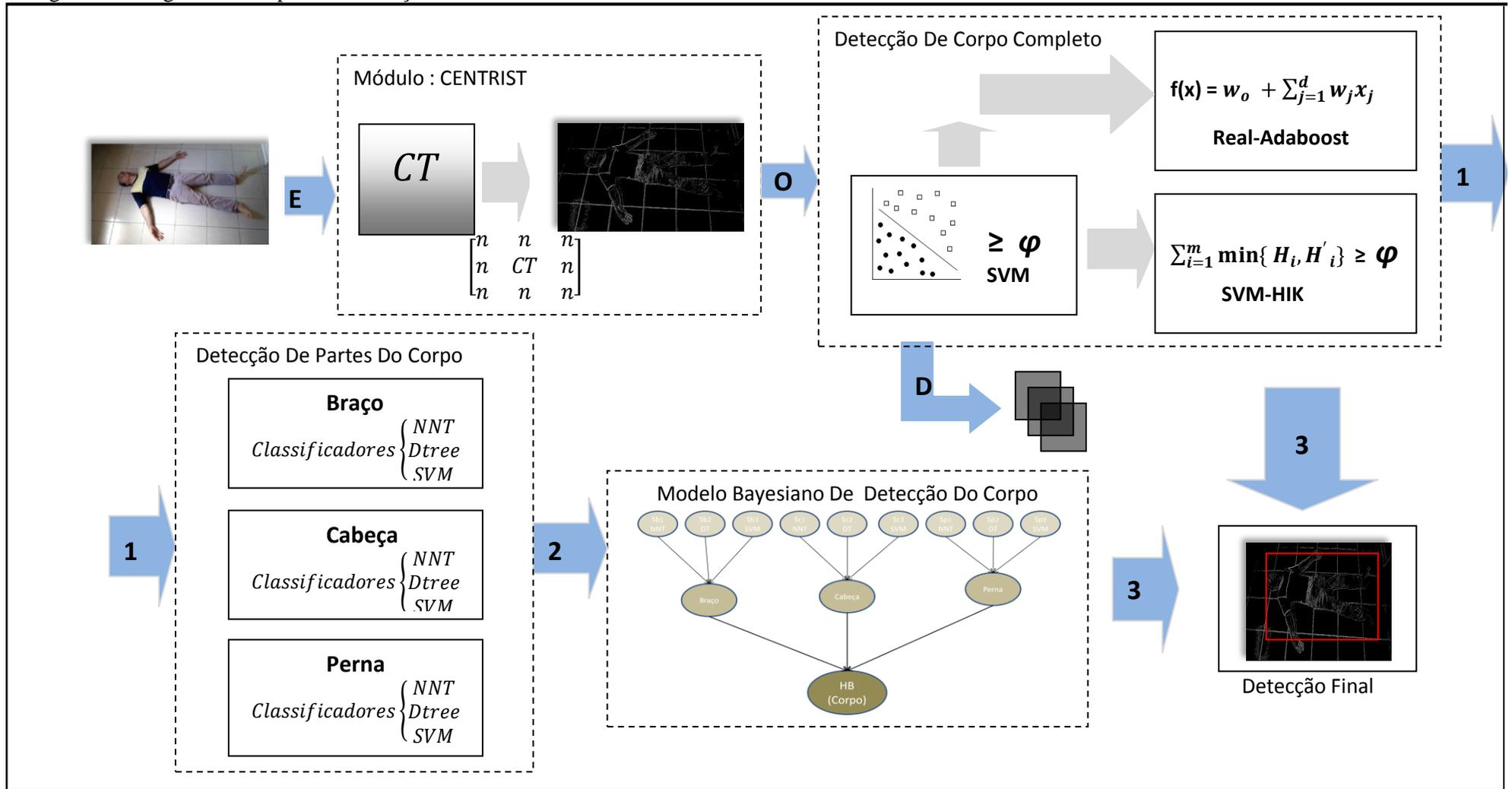
Uma relevante mudança na forma tradicional de predição de cada janela detectada, durante o processo de classificação, é que, na verdade, não é necessária nem a criação do próprio vetor de atributos. Ao supor que há uma imagem I , é então extraído o vetor de atributos $v \in \mathbb{R}^{256}$ e que o treinamento do classificador linear, obviamente já realizado, resulta em $\lambda \in \mathbb{R}^{256}$, onde λ representa os coeficientes do modelo de treinamento, a imagem I de interesse é, então, classificada corretamente, se e somente se $\lambda^T v \geq \varphi$ (limiar). Dado que a imagem I (CENTRIST) da janela de detecção tem N pixels com seus respectivos valores i , então o histograma do CENTRIST é satisfeito com $v = N$. Seguramente, todos os pixels contribuem, tal que $\lambda_i v_i = N\lambda_i$ para $\lambda^T v$, desde que $\lambda^T v = \sum_{i=1}^{256} \lambda_i v_i$. Pelo exposto, pode-se distribuir para cada pixel o valor respectivo de λ_i na imagem I (CENTRIST) e multiplicá-los obtendo uma soma com convergência para $N\lambda_i$. Logo constrói-se uma imagem auxiliar definida como:

$$I'(x, y) = \lambda CT_{(x,y)}, \quad (51)$$

onde I' é a imagem após a operação de distribuição de λ_i , coeficientes do modelo de treinamento do SVM-linear ou HIK, do vetor v na imagem após a transformação CT (CENTRIST).

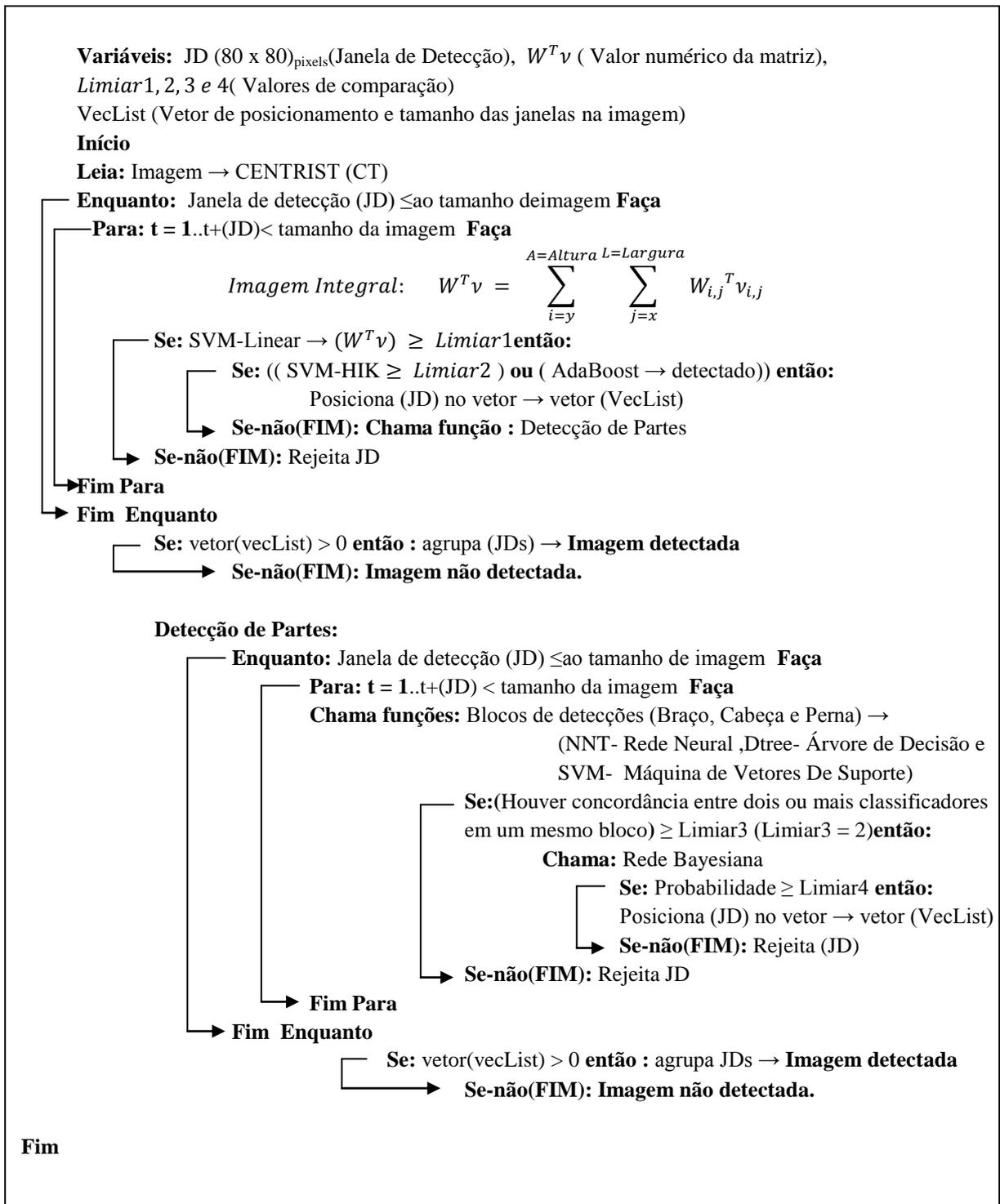
Essa simples operação matricial tem uma grande vantagem, visto que não é mais necessária a operação de predição dos valores de classificação e nem a criação do próprio vetor de atributos, basta a soma dos valores desta matriz, figura 42, e a comparação com um limiar.

Figura 37 - Algoritmo completo de detecção de formas humanas em cenas de desastres.



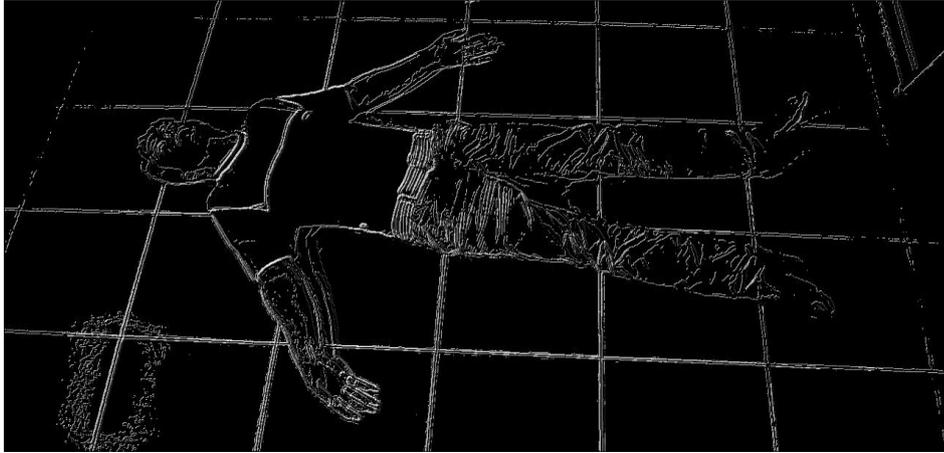
Fonte: Autor

Algoritmo 3 –Pseudo-código do algoritmo de detecção de formas humanas em cenas de desastres.



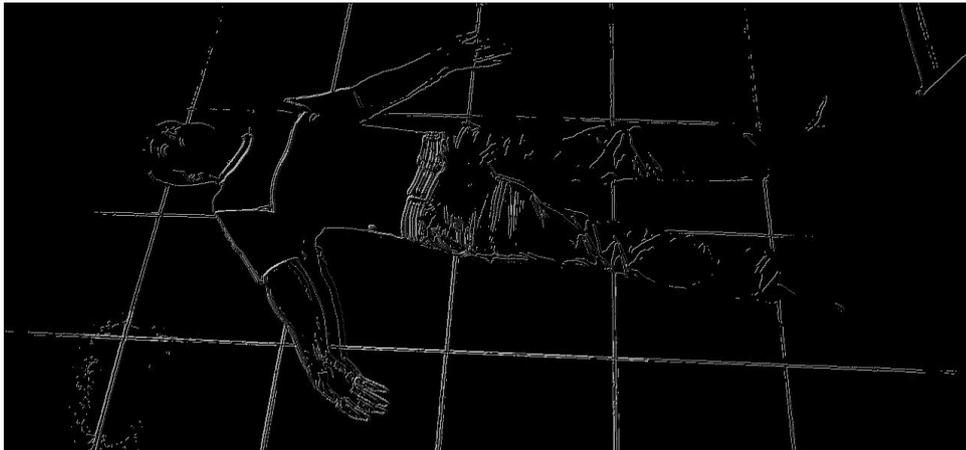
Fonte: Autor.

Figura 38– Ilustração da ponderação dos gradientes (G_x e G_y) no valor absoluto de 0.012.



Fonte: Autor.

Figura 39– Ilustração da ponderação dos gradientes (G_x e G_y) no valor absoluto de 0.009.



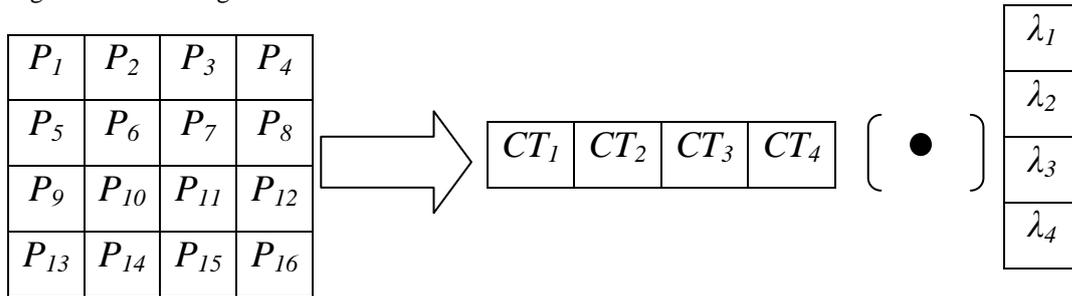
Fonte: Autor.

Figura 40– Ilustração da ponderação dos gradientes (G_x e G_y) no valor absoluto de 0.003.



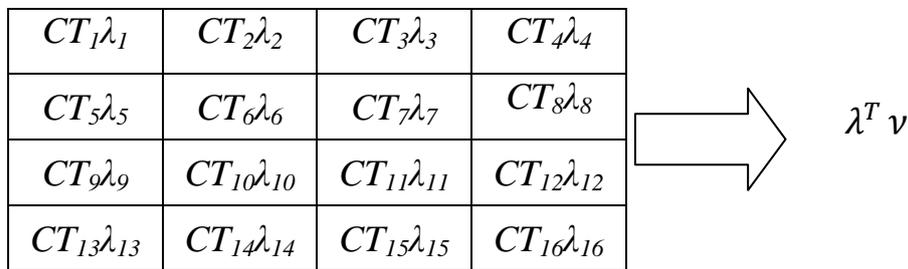
Fonte: Autor.

Figura 41 - Forma generalizada de calcular o vetor de atributos.



Fonte: Autor “adaptado de” CENTRIST Wu; J.Rehg M., 2011.

Figura 42 - Forma otimizada para obter o limiar de trabalho do classificador.



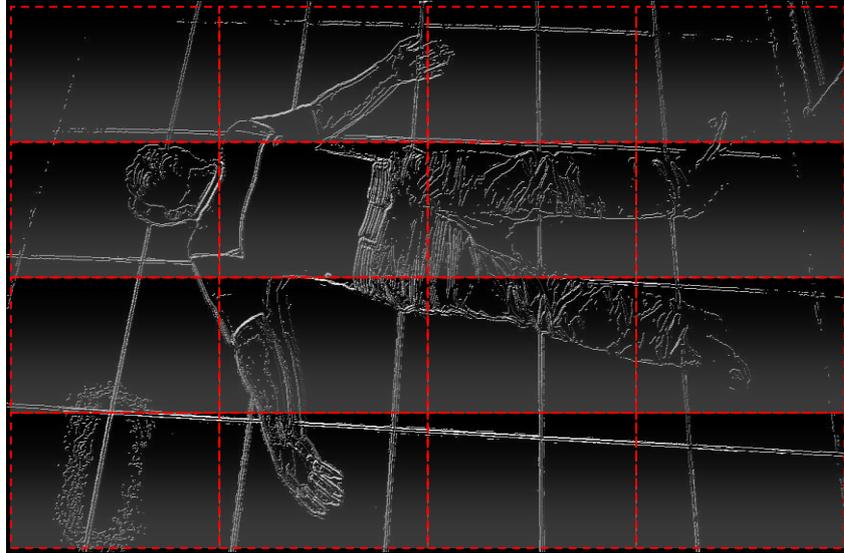
Fonte: Autor “adaptado de” CENTRIST Wu; J. Rehg M., 2011.

A figura 41 apresenta a classificação de uma imagem da forma tradicional, ou seja, cria-se o vetor de atributos e aplica-se a predição, onde P_i (matriz de pixels), CT_i (vetor centrist) e λ_i (coeficientes do modelo). Na figura 42 tem-se a matriz de valores obtidos da multiplicação de cada pixel obtido da imagem CENTRIST pelos coeficientes (pesos) do modelo SVM-linear ou HIK, tendo como resultado uma classificação mais rápida, pois, envolve apenas somas (valores da matriz), com pouco uso de memória e complexidade $O(1)$. Como a predição foi substituída por um limiar ($\lambda^T v \geq \varphi$), que corresponde à soma de todos os valores de $I'(x, y)$, é natural a utilização da técnica da Imagem Integral, o que leva a uma redução também no tempo de processamento.

A janela de detecção, usada para capturar uma imagem, gerar a matriz da figura 42 e posterior classificação, é subdividida em blocos (não sobrepostos) de pixels e é descrita a partir da estratégia ilustrada na figura 43, que apresenta a imagem I de tamanho A (altura) \times L (largura), dividida em blocos de tamanho $h_b = A/div_y$ e $w_b = L/div_x$, gerando a imagem $I'(x, y)$ (CENTRIST), em que A e L correspondem, respectivamente, à altura e à largura da imagem na janela de detecção. Os valores h_b e w_b são dimensões de cada bloco, definidas por div_x e div_y .

Durante a criação da imagem CENTRIST, não são considerados os pixels das bordas, pois, esse algoritmo requer uma matriz 3 x 3.

Figura 43 - Divisão da janela de detecção em blocos.



Fonte: Autor.

Durante a operação de varredura da imagem, cada janela de detecção tem os blocos ($h_b \times w_b$) percorridos, seguindo o movimento da esquerda para a direita, e de cima para baixo, a partir da posição (t, l):

$$\sum_{i=hb}^A \sum_{j=wb}^L \sum_{x=2}^{hb-1} \sum_{y=2}^{wb-1} \lambda_{i,j}^{C(t+(i-1)+x, l+(j-1)+y)} \geq \varphi, \quad (52)$$

onde $\lambda_{i,j}^K$ é a K -ésima componente de $\lambda_{i,j}$ na matriz de valores, e $C(x, y)$ é um pixel na imagem CENTRIST C . Observa-se que, para não contar com os valores de bordas, o processo se inicia em $x=2$ e $y=2$, terminando em $hb-1$ e $wb-1$.

A inequação 52 pode, então, ser agrupada conforme a inequação 53, na qual se observa, em vermelho, a própria definição do método da Imagem Integral.

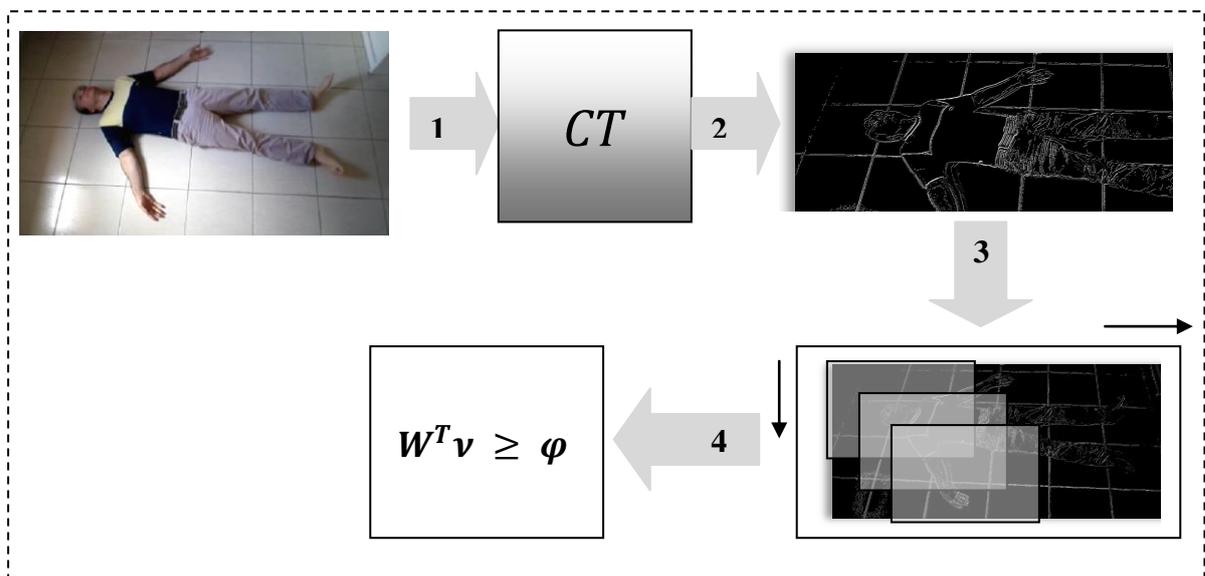
$$\sum_{i=hb}^A \sum_{j=wb}^L \left(\sum_{x=2}^{hb-1} \sum_{y=2}^{wb-1} \lambda_{i,j}^{C(t+(i-1)+x, l+(j-1)+y)} \right) \geq \varphi, \quad (53)$$

$$W^T \nu = \sum_{i=hb}^A \sum_{j=wb}^L \lambda_{i,j}^T \nu_{i,j} \geq \varphi, \quad (54)$$

A inequação 54 demonstra que, a partir da Imagem Integral, calculada em complexidade linear $O(1)$, essa inequação é avaliada segundo o limiar (φ), em que é definido o ponto de avaliação dos classificadores SVM-Linear e HIK.

O diagrama de blocos do exposto acima: fase inicial de aquisição de imagem (1), transformação CENTRIST (2), a varredura da janela de detecção sobre a imagem (3) e a obtenção do valor a ser usado como limiar de avaliação dos classificadores de corpo completo (SVM-Linear e HIK) (4) é apresentado na figura 44.

Figura 44 - Aquisição de imagem, *CENTRIST* e determinação do limiar



Fonte: Autor

Após a apresentação inicial de como é feita a detecção da imagem, a transformação *CENTRIST* e a determinação do valor a ser usado como limiar para avaliação dos classificadores SVM-linear e HIK, é apresentada então a estratégia de detecção da forma humana completa, a configuração dos classificadores e as decisões tomadas a partir do limiar no item 4.2

4.2 DETECÇÃO DE CORPO COMPLETO

No fluxograma da figura 37 ainda na entrada (O) do módulo de detecção de corpo completo, as janelas de detecções na imagem de entrada usam a estratégia de força-bruta, a qual examina todas as possibilidades existentes na dimensão (A, L) da imagem. Essas inúmeras janelas são, então, classificadas como sendo de interesse ou não. Aqui, dois parâmetros são de real importância no desempenho da exatidão versus tempo de processamento total da imagem completa:

- a) o valor do passo utilizado entre as janelas de detecções, ou seja, o deslocamento na direção L , pois influencia na quantidade de janelas a serem avaliadas e na exatidão de detalhes que podem ser desprezados. Neste trabalho, os valores de 3 e 5 pixels de deslocamento foram avaliados em função da resolução máxima de 640 x 480 pixels;
- b) o valor da escala de redimensionamento da imagem de entrada, já que pode haver imagens maiores que a janela de detecção (A_d, L_d) . Foi aplicada ainda a força bruta, o que afeta diretamente o desempenho, em razão do número de redimensionamento seguido do mesmo procedimento do item a acima. O passo de cada redimensionamento neste trabalho segue a relação $(A_d, L_d) \geq I_{input}$ (0.8) e (0.9), as duas condições são avaliadas em função da resolução máxima 640 x 480 pixels.

A avaliação do limiar fica a cargo dos classificadores SVM-Linear e SVM-HIK. Como já demonstrado nos parágrafos anteriores, não há necessidade de uma função de predição, e tais classificadores são mais rápidos na tomada de decisão. Esses classificadores estão estruturados em uma configuração de cascata, em que a entrada é constituída pelo SVM-Linear, que tem o limiar ajustado para que rejeite a maioria das janelas que contenham o plano de fundo, observar figura 37 a saída (D) do módulo de detecção de corpo completo. Porém, deixa passar grande parte das janelas, para o estágio seguinte. Pode-se considerar que o SVM-linear é um filtro grosso de rejeição, mas com menor exatidão na tomada de decisão.

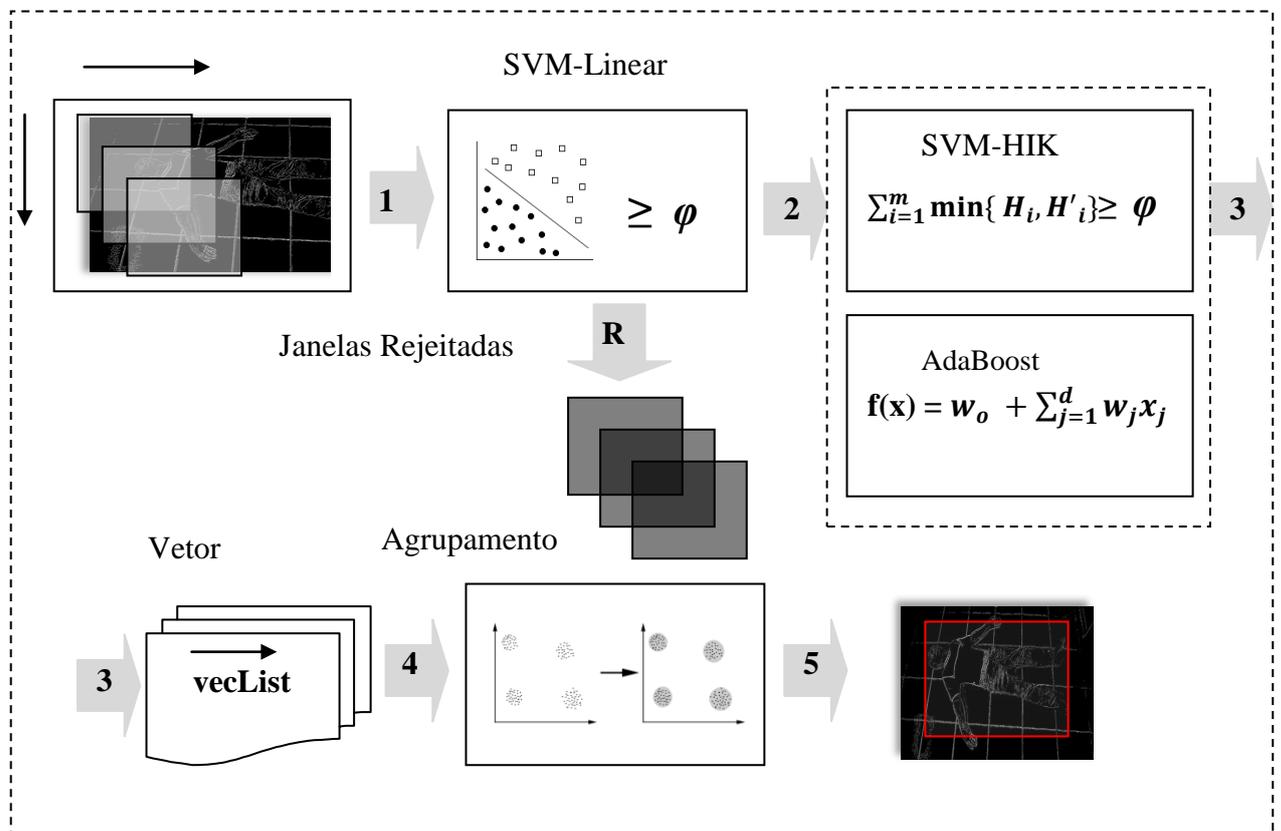
O estágio seguinte nessa estrutura em cascata é o SVM-HIK, que também trabalha sob um limiar de decisão. Tem complexidade $O(1)$, e recebe as janelas de detecções, aprovadas pelo estágio anterior, e toma a decisão final, se há a presença de uma forma humana completa ou não.

Uma alternativa ao segundo estágio, constituída pelo classificador SVM-HIK, é o uso do classificador *real AdaBoost* (*Adaptive Boosting*), cujo objetivo é o teste de desempenho com relação ao HIK.

As janelas de detecções aprovadas nos dois estágios, incluindo o *AdaBoost*, são encaminhadas a um vetor (*vecList*), que mantém todas essas em uma ordem de posição (P_{Lx} , P_{Ay}), referente à imagem de entrada e também ao tamanho do redimensionamento (*Largura*, *Altura*). Ao final de cada redimensionamento, é então aplicado um algoritmo de agrupamento (*Clustering*) no vetor (*vecList*). O objetivo dessa etapa é minimizar o efeito de multi-deteções, devido ao redimensionamento da imagem.

O diagrama de blocos da fase de classificação, com os dois estágios, e o algoritmo de agrupamento, é apresentado na figura 45: Na saída (1) tem-se as janelas de detecções, após a aplicação da estratégia de força bruta, as saídas (R) e (2) são respectivamente as janelas rejeitadas pelo classificador SVM-linear e as que foram aprovadas, em (3) é a etapa de saída do classificador em cascata (SVM-HIK ou Adaboost), em (4) é a saída do vetor *vecList* e ao final em (5) a saída das janelas de detecções já agrupadas.

Figura 45 - Estrutura de classificação em cascata.



Fonte: Autor.

4.3 DETECÇÃO DE PARTES DO CORPO

Na seção anterior, a estratégia para detecção do corpo completo foi apresentada, através dos classificadores SVM-Linear e HIK; porém, como já abordado, não é a solução única ou definitiva para a detecção da forma humana, principalmente em situações adversas em que são submetidas durante eventos de desastres. As janelas de detecções rejeitadas pelo classificador SVM-Linear, no primeiro estágio da estrutura em cascata, são todas descartadas, pois o objetivo aqui é a eliminação do plano de fundo. No caso do segundo estágio, SVM-HIK ou SVM-Adaboost, as amostras rejeitadas não são definitivamente descartadas, e sim, encaminhadas para posterior análise, chamada de detecção de partes responsável por detectar (braço, cabeça e perna).

A detecção de partes é formada por três blocos de três classificadores (comitê de especialistas), ver figura 37 saída (1) de detecção de corpo completo, que trabalham de forma paralela, com o objetivo de aumentar a confiabilidade da amostra observada, janela de detecção, baseada na exatidão de cada classificador e do número de classificadores concordantes na observação. Todavia, não é aplicada aqui a análise de concordância estatística *Kappa* (COHEN; 1960), e sim seleção (ver apêndice A), com o uso de uma simples função de corte, usada na avaliação inicial das amostras. A decisão final é feita por uma rede Bayesiana, cujos valores de exatidões obtidas dos classificadores atuará como nós pais, avaliará as amostras concordantes no comitê de especialistas e calculará a probabilidade de existência de uma forma humana ou não em uma determinada imagem.

A tabela 1 e a inequação 55, a seguir, apresentam as exatidões usadas na rede Bayesiana, além da função de avaliação de corte das amostras observadas:

Tabela 1 - Exatidões dos classificadores usadas na Rede Bayesiana.

Classifiers	NNT-1 Rede Neural	Dtree-2 Árvore de Decisão	SVM-3 Máquina de Vetores de Suporte
Atributos	Exatidão	Exatidão	Exatidão
Braço Sb	ACC_{B1}	ACC_{B2}	ACC_{B3}
Cabeça Sc	ACC_{C1}	ACC_{C2}	ACC_{C3}
Perna Sp	ACC_{P1}	ACC_{P2}	ACC_{P3}

Fonte: Autor.

$$F_C \leq ((Sb_1) + (Sb_2) + (Sb_3)) + ((Sc_1) + (Sc_2) + (Sc_3)) + ((Sp_1) + (Sp_2) + (Sp_3)), \quad (55)$$

onde Sb_i , Sc_i e Sp_i são valores binários, correspondentes à detecção de um classificador i , indicando ser positiva (1) ou a não detecção (0). O valor de cada bloco $((Sb_1) + (Sb_2) + (Sb_3))$ ou $((Sc_1) + (Sc_2) + (Sc_3))$ ou $((Sp_1) + (Sp_2) + (Sp_3))$ é assumido ser maior ou igual a dois (Bloco ≥ 2), em razão do exposto sobre a concordância entre classificadores; ou seja, no mínimo, há concordância entre dois classificadores dentro do mesmo bloco. A função de corte é assumida ser maior ou igual a um ($F_C \geq 1$), pois precisa de pelo menos um bloco concordante para encaminhar à análise pela rede Bayesiana.

As imagens rejeitadas pela etapa anterior, figura 37 saída(1) detecção de corpo completo, são então encaminhadas para uma nova operação de varredura. Contudo, essa tem o passo reduzido para 3 pixels, com vistas a aumentar a exatidão de detecção, e também tem reduzido o número de redimensionamentos da janela, para não afetar o desempenho do tempo de processamento final.

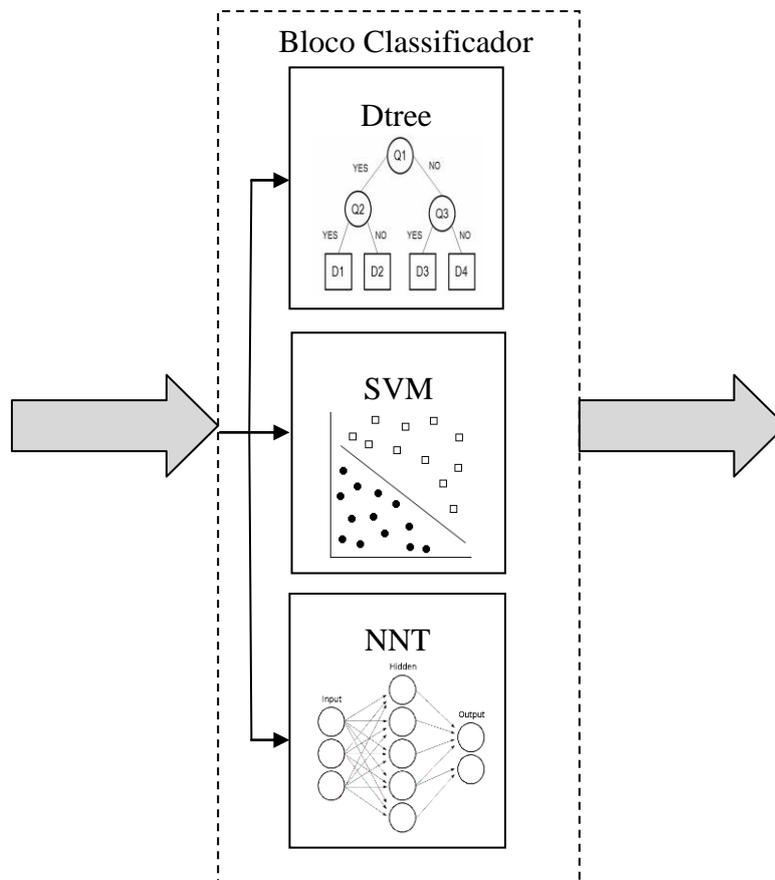
A estrutura de formação dos três blocos de classificadores é apresentada na figura 46. Cada bloco é formado por três classificadores (NNT-Rede Neural, Dtree-Árvore de Decisão e SVM-Máquina de Vetores de Suporte) que têm abordagens distintas, treinados para detectar partes específicas do corpo humano (Braço, Cabeça e Perna). A meta aqui foi escolher classificadores distintos, para classificar partes iguais do corpo humano como por exemplo: uma cabeça. Cujo objetivo é evitar vícios de amostras e classificações. Os tamanhos das amostras utilizadas nos treinamentos dos blocos são apresentados como segue:

$$\text{a) bloco-1 (Classificador de Braço com 278 amostras)} \begin{cases} NNT & Sb_1 \\ Dtree & Sb_2 \\ SVM & Sb_3 \end{cases}$$

$$\text{b) bloco-2 (Classificador de Cabeça com 162 amostras)} \begin{cases} NNT & Sc_1 \\ Dtree & Sc_2 \\ SVM & Sc_3 \end{cases}$$

$$\text{c) bloco-3 (Classificador de Perna com 282 amostras)} \begin{cases} NNT & Sp_1 \\ Dtree & Sp_2 \\ SVM & Sp_3 \end{cases}$$

Figura 46 - Estrutura de formação dos blocos classificadores



Fonte: Autor.

A figura 46 mostra que cada bloco é formado por classificadores como árvore de decisão (Dtree), máquina de vetores de suporte (SVM) e rede neural artificial (NNT). Esses classificadores trabalham em estruturas padronizadas do openCV e apresentam funções de predições, diferentemente da versão de corpo completo com uso de limiar.

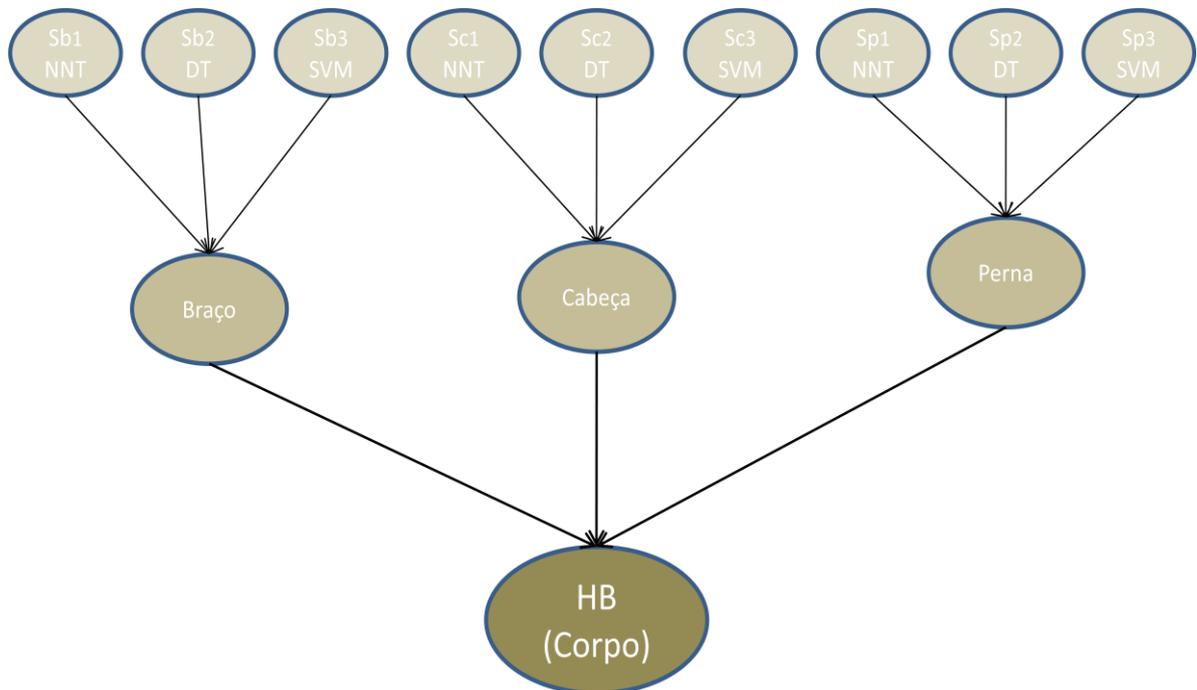
As configurações desses classificadores são apresentadas adiante:

- dtree→É uma árvore binária, em que cada folha tem dois filhos e é construída recursivamente; inicia-se no nó raiz, e todas as amostras (treinamento) são usadas para dividir esse nó. A profundidade máxima adotada neste trabalho é de 25 ramos;
- SVM→ É a classificação binária com um núcleo linear (*Linear Kernel*);
- NNT→É uma rede neural de alimentação direta, com três camadas; com número de entradas dependente do vetor de atributos, 2 nós escondidos e duas saídas (classes) do tipo *Perceptrons*. A função de ativação é uma Sigmóide, e o treinamento usa o *Back-propagation*.

4.4 MODELO BAYESIANO DE DETECÇÃO DO CORPO

O módulo final do algoritmo consiste na estratégia de análise das amostras detectadas dos classificadores de partes, da seção anterior. Essa estratégia é baseada nas redes Bayesianas, pois a presença de um corpo (HB) ou não (\neg HB) é dependente da estrutura das variáveis (Braço, Cabeça e Perna). Essa, por consequência, apresenta uma dependência das variáveis de entrada (Sb_1, Sb_2, Sb_3), (Sc_1, Sc_2, Sc_3) e (Sp_1, Sp_2, Sp_3), seus pais, identificados por S (Sensor), b , c , e p (braço, cabeça e perna), e o índice i (1, 2 e 3), por (1) NNT-Rede Neural, (2) DTree-Árvore de Decisão e (3) SVM-Máquina de Vetores de Suporte. E toda etapa dessa rede, em cada nó (variável), apresenta uma tabela de distribuição de probabilidade condicional (TPC), que quantifica o efeito dos pais sobre o nó filho, tal que $P(X_i|Pais(X_i))$. A figura 47, a seguir, apresenta o grafo direcionado e acíclico, com as variáveis da rede e suas dependências:

Figura 47 - Grafo direcionado e acíclico de dependências das variáveis no modelo do corpo humano.



Fonte: Autor.

A topologia da rede Bayesiana observada na figura 47 também representa relações de independência condicional entre os nós (Braço), (Cabeça) e (Perna) e as variáveis de entrada (Sb_1, Sb_2, Sb_3), (Sc_1, Sc_2, Sc_3) e (Sp_1, Sp_2, Sp_3), causas diretas da presença de um Braço, Cabeça ou Perna.

Como essa rede reflete o conhecimento, pode-se, então, montar a Tabela de Probabilidade Condicional (TPC) para avaliação das distribuições condicionais. Nessa, cada linha em uma (TPC) contém a probabilidade condicional de cada valor de nó para um caso de condicionamento (combinação possível de valores para os nós superiores).

Na montagem de cada Tabela de Probabilidade Condicional (TPC), usada neste trabalho, algumas observações são essenciais:

- a) os nós sem pais ((Sb_1, Sb_2, Sb_3) , (Sc_1, Sc_2, Sc_3) e (Sp_1, Sp_2, Sp_3)) têm como probabilidade a priori a exatidão obtida no treinamento de cada classificador de partes, ou seja, ao classificador de Rede Neural (NNT), Árvore de Decisão (DTree) e Máquina de vetores de suporte (SVM);
- b) na tabela, cada variável booleana com K pais (também booleanos), possui 2^K probabilidades; assim a rede completa é igual a $O(n2^K)$, em que n é o número de nós;
- c) quando não houver o valor da probabilidade condicional (relacionamentos entre nós pais e filhos são arbitrários), é utilizada a distribuição canônica, a partir dos nós determinísticos (item a) e do método do OU-ruidoso.

As tabelas 2 a 5 apresentam as (TPCs) aplicadas nesta dissertação. Os nós a priori são apresentados nas probabilidades de (56) a (64), e na cor vermelha, o valor da exatidão de cada classificador (Sb_i, Sc_i e Sp_i) índice i (1, 2 e 3) (1) NNT-Rede Neural, (2) DTree-Árvore de Decisão e (3) SVM-Máquina de Vetores de Suporte, treinados nos atributos Sb (braço), Sc (cabeça) e Sp (pernas). Os demais valores nas tabelas são obtidos pelo método do OU-Ruidoso.

$$P(\neg B | \neg Sb_1, \neg Sb_2, Sb_3) = \text{Exatidão de } Sb_3 = \mathbf{(1-AccSb_3)}, \quad (56)$$

$$P(\neg B | \neg Sb_1, Sb_2, \neg Sb_3) = \text{Exatidão de } Sb_2 = \mathbf{(1-AccSb_2)}, \quad (57)$$

$$P(\neg B | Sb_1, \neg Sb_2, \neg Sb_3) = \text{Exatidão de } Sb_1 = \mathbf{(1-AccSb_1)}, \quad (58)$$

Tabela 2 - Tabela de Probabilidade Condicional (TPC) para a variável Braço.

Sb_1	Sb_2	Sb_3	$P(B Sb_1, Sb_2, Sb_3)$	$P(\neg B Sb_1, Sb_2, Sb_3)$
0	0	0	0	1
0	0	1	$((AccSb_3))$	$(1-(AccSb_3))$
0	1	0	$((AccSb_2))$	$(1-(AccSb_2))$
0	1	1	$1-((1-(AccSb_2)) \times (1-(AccSb_3)))$	$(1-(AccSb_2)) \times (1-(AccSb_3))$
1	0	0	$((AccSb_1))$	$(1-(AccSb_1))$
1	0	1	$1-((1-(AccSb_1)) \times (1-(AccSb_3)))$	$(1-(AccSb_1)) \times (1-(AccSb_3))$
1	1	0	$1-((1-(AccSb_1)) \times (1-(AccSb_2)))$	$(1-(AccSb_1)) \times (1-(AccSb_2))$
1	1	1	$1-((1-(AccSb_1)) \times (1-(AccSb_2)) \times (1-(AccSb_3)))$	$(1-(AccSb_1)) \times (1-(AccSb_2)) \times (1-(AccSb_3))$

Fonte: Autor.

$$P(\neg C | \neg Sc_1, \neg Sc_2, Sc_3) = \text{Exatidão de } Sc_3 = (1-AccSc_3), \quad (59)$$

$$P(\neg C | \neg Sc_1, Sc_2, \neg Sc_3) = \text{Exatidão de } Sc_2 = (1-AccSc_2), \quad (60)$$

$$P(\neg C | Sc_1, \neg Sc_2, \neg Sc_3) = \text{Exatidão de } Sc_1 = (1-AccSc_1), \quad (61)$$

Tabela 3 - Tabela de Probabilidade Condicional (TPC) para a variável Cabeça.

Sc_1	Sc_2	Sc_3	$P(C Sc_1, Sc_2, Sc_3)$	$P(\neg C Sc_1, Sc_2, Sc_3)$
0	0	0	0	1
0	0	1	$(AccSc_3)$	$(1-(AccSc_3))$
0	1	0	$(AccSc_2)$	$(1-(AccSc_2))$
0	1	1	$1-((1-(AccSc_2)) \times (1-(AccSc_3)))$	$(1-(AccSc_2)) \times (1-(AccSc_3))$
1	0	0	$(AccSc_1)$	$(1-(AccSc_1))$
1	0	1	$1-((1-(AccSc_1)) \times (1-(AccSc_3)))$	$(1-(AccSc_1)) \times (1-(AccSc_3))$
1	1	0	$1-((1-(AccSc_1)) \times (1-(AccSc_2)))$	$(1-(AccSc_1)) \times (1-(AccSc_2))$
1	1	1	$1-((1-(AccSc_1)) \times (1-(AccSc_2)) \times (1-(AccSc_3)))$	$(1-(AccSc_1)) \times (1-(AccSc_2)) \times (1-(AccSc_3))$

Fonte: Autor.

$$P(\neg P | \neg Sp_1, \neg Sp_2, Sp_3) = \text{Exatidão de } Sp_3 = (1-AccSp_3), \quad (62)$$

$$P(\neg P | \neg Sp_1, Sp_2, \neg Sp_3) = \text{Exatidão de } Sp_2 = (1-AccSp_2), \quad (63)$$

$$P(\neg P | Sp_1, \neg Sp_2, \neg Sp_3) = \text{Exatidão de } Sp_1 = (1-AccSp_1), \quad (64)$$

Tabela 4 - Tabela de Probabilidade Condicional (TPC) para a variável Perna.

Sp_1	Sp_2	Sp_3	$P(P Sp_1, Sp_2, Sp_3)$	$P(\neg P Sp_1, Sp_2, Sp_3)$
0	0	0	0	1
0	0	1	(AccSp ₃)	(1-(AccSp₃))
0	1	0	(AccSp ₂)	(1-(AccSp₂))
0	1	1	1-((1-(AccSp ₂)) x (1-(AccSp ₃)))	(1-(AccSp ₂)) x (1-(AccSp ₃))
1	0	0	(AccSp ₁)	(1-(AccSp₁))
1	0	1	1-((1-(AccSp ₁)) x (1-(AccSp ₃)))	(1-(AccSp ₁)) x (1-(AccSp ₃))
1	1	0	1-((1-(AccSp ₁)) x (1-(AccSp ₂)))	(1-(AccSp ₁)) x (1-(AccSp ₂))
1	1	1	1-((1-(AccSp ₁)) x (1-(AccSp ₂)) x (1-(AccSp ₃)))	(1-(AccSp ₁)) x (1-(AccSp ₂)) x (1-(AccSp ₃))

Fonte: Autor.

Na tabela 5 (TPC), relativa às variáveis B, C e P (Braço, Cabeça e Perna), figura 47, a determinação das probabilidades condicionais da presença do corpo (HB) ou não ($\neg HB$), são consideradas todas as probabilidades resultantes das tabelas 2, 3 e 4. Logo cada expressão abaixo corresponde a uma linha a priori, circundada em vermelho, da tabela 5:

para a condição; $P(\neg HB | \neg B, \neg C, P)$ tem-se as seguintes probabilidades na tabela 5:

$$1, \mathbf{(1-(AccSp_3))}, \mathbf{(1-(AccSp_2))}, (1-(AccSp_2)) \times (1-(AccSp_3)), \mathbf{(1-(AccSp_1))}, (1-(AccSp_1)) \times (1-(AccSp_3)), (1-(AccSp_1)) \times (1-(AccSp_2)), (1-(AccSp_1)) \times (1-(AccSp_2)) \times (1-(AccSp_3)), \quad (65)$$

para a condição; $P(\neg HB | \neg B, C, \neg P)$ tem-se as seguintes probabilidades na tabela 5:

$$1, \mathbf{(1-(AccSc_3))}, \mathbf{(1-(AccSc_2))}, (1-(AccSc_2)) \times (1-(AccSc_3)), \mathbf{(1-(AccSc_1))}, (1-(AccSc_1)) \times (1-(AccSc_3)), (1-(AccSc_1)) \times (1-(AccSc_2)), (1-(AccSc_1)) \times (1-(AccSc_2)) \times (1-(AccSc_3)), \quad (66)$$

para a condição; $P(\neg HB | B, \neg C, \neg P)$ tem-se as seguintes probabilidades na tabela 5:

$$1, \mathbf{(1-(AccSb_3))}, \mathbf{(1-(AccSb_2))}, (1-(AccSb_2)) \times (1-(AccSb_3)), \mathbf{(1-(AccSb_1))}, (1-(AccSb_1)) \times (1-(AccSb_3)), (1-(AccSb_1)) \times (1-(AccSb_2)), (1-(AccSb_1)) \times (1-(AccSb_2)) \times (1-(AccSb_3)), \quad (67)$$

Para facilitar a leitura da tabela 5 de $P(HB)$, foram criadas as seguintes legendas (A, B, C e D) e substituídas nas expressões 65, 66 e 67. Desse modo, foram obtidas as expressões 68, 69 e 70, que correspondem às probabilidades dos nós apriori (P, C e B) da tabela 5, e auxiliam na determinação das outras relações probabilísticas com o uso do OU-Ruidoso.

$((1-(AccSp_2)) \times (1-(AccSp_3))) = p_1$ $((1-(AccSp_1)) \times (1-(AccSp_3))) = p_2$ $((1-(AccSp_1)) \times (1-(AccSp_2))) = p_3$ $((1-(AccSp_1)) \times (1-(AccSp_2)) \times (1-(AccSp_3))) = p_4$	A	$((1-(AccSc_2)) \times (1-(AccSc_3))) = c_1$ $((1-(AccSc_1)) \times (1-(AccSc_3))) = c_2$ $((1-(AccSc_1)) \times (1-(AccSc_2))) = c_3$ $((1-(AccSc_1)) \times (1-(AccSc_2)) \times (1-(AccSc_3))) = c_4$	B
$((1-(AccSb_2)) \times (1-(AccSb_3))) = b_1$ $((1-(AccSb_1)) \times (1-(AccSb_3))) = b_2$ $((1-(AccSb_1)) \times (1-(AccSb_2))) = b_3$ $((1-(AccSb_1)) \times (1-(AccSb_2)) \times (1-(AccSb_3))) = b_4$	C	$AccSp_1 = Ap_1 \quad AccSp_2 = Ap_2 \quad AccSp_3 = Ap_3$ $AccSc_1 = Ac_1 \quad AccSc_2 = Ac_2 \quad AccSc_3 = Ac_3$ $AccSb_1 = Ab_1 \quad AccSb_2 = Ab_2 \quad AccSb_3 = Ab_3$	D

Legendas A, B, C, e D.

substituindo as legendas A em (65): $(1, 1-Ap_3, 1-Ap_2, p_1, 1-Ap_1, p_2, p_3, p_4)$, (68)

substituindo as legendas B em (66): $(1, 1-Ac_3, 1-Ac_2, c_1, 1-Ac_1, c_2, c_3, c_4)$, (69)

substituindo as legendas C em (67): $(1, 1-Ab_3, 1-Ab_2, b_1, 1-Ab_1, b_2, b_3, b_4)$, (70)

A tabela de probabilidade condicional (TPC) 5 mostra que todas as probabilidades de $P(HB)$ não são simples relações diretas das variáveis pais (B, C e P). Isso porque estes pais variam conforme uma distribuição condicional dos conjuntos $((Sb_1, Sb_2, Sb_3)$, (Sc_1, Sc_2, Sc_3) e (Sp_1, Sp_2, Sp_3)), ou seja, para uma dada consulta $P(HB | \neg B, \neg C, P)$, há uma combinação de oito probabilidades, e o conjunto total de $2^9 = 512$ possíveis probabilidades de existência de um corpo em uma dada imagem. Entretanto, este trabalho considera apenas os resultados concordantes de dois ou três classificadores para cada parte do corpo detectada. Assim, há 64 possíveis probabilidades apresentadas pelas saídas S3, S5, S6 e S7.

Tabela 5 - Tabela de Probabilidade Condicional (TPC) da presença do corpo $P(HB)$ ou não $P(\neg HB)$, dadas as probabilidades determinísticas Braço (B), Cabeça(C) e Perna(P) (em vermelho) + uso do OU-Ruidoso.

NNT			0	0	0	1	1	1	1	1	1	1	1	1	1	
DTree			0	1	1	0	0	1	0	0	1	1	1	1	1	
SVM			1	0	1	0	1	0	1	0	1	0	1	1	1	
variáveis			Saída: S1		Saída: S2		Saída: S3		Saída: S4		Saída: S5		Saída: S6		Saída: S7	
B	C	P	$P(HB)$	$P(\neg HB)$	$P(HB)$	$P(\neg HB)$	$P(HB)$	$P(\neg HB)$	$P(HB)$	$P(\neg HB)$	$P(HB)$	$P(\neg HB)$	$P(HB)$	$P(\neg HB)$	$P(HB)$	$P(\neg HB)$
0	0	0	0	1	0	1	0	1	0	1	0	1	0	1	0	1
0	0	1	Ap_3	$1-Ap_3$	Ap_2	$\neg Ap_2$	$1-p_1$	p_1	Ap_1	$1-Ap_1$	$1-p_2$	p_2	$1-p_3$	p_3	$1-p_4$	p_4
0	1	0	Ac_3	$1-Ac_3$	Ac_2	$\neg Ac_2$	$1-c_1$	c_1	Ac_1	$1-Ac_1$	$1-c_2$	c_2	$1-c_3$	c_3	$1-c_4$	c_4
0	1	1	$1-((1-Ap_3) \times (1-Ac_3))$	$(1-Ap_3) \times (1-Ac_3)$	$1-((1-Ap_2) \times (1-Ac_2))$	$(1-Ap_2) \times (1-Ac_2)$	$1-((1-c_1) \times (1-p_1))$	$(1-c_1) \times (1-p_1)$	$1-((1-Ap_1) \times (1-Ac_1))$	$(1-Ap_1) \times (1-Ac_1)$	$1-((1-c_2) \times (1-p_2))$	$(1-c_2) \times (1-p_2)$	$1-((1-c_3) \times (1-p_3))$	$(1-c_3) \times (1-p_3)$	$1-((1-c_4) \times (1-p_4))$	$(1-c_4) \times (1-p_4)$
1	0	0	Ab_3	$1-Ab_3$	Ab_2	$1-Ab_2$	$1-b_1$	b_1	Ab_1	$1-Ab_1$	$1-b_2$	b_2	$1-b_3$	b_3	$1-b_4$	b_4
1	0	1	$1-((1-Ab_3) \times (1-Ap_3))$	$(1-Ab_3) \times (1-Ap_3)$	$1-((1-Ab_2) \times (1-Ap_2))$	$(1-Ab_2) \times (1-Ap_2)$	$1-((1-b_1) \times (1-p_1))$	$(1-b_1) \times (1-p_1)$	$1-((1-Ab_1) \times (1-Ap_1))$	$(1-Ab_1) \times (1-Ap_1)$	$1-((1-b_2) \times (1-p_2))$	$(1-b_2) \times (1-p_2)$	$1-((1-b_3) \times (1-p_3))$	$(1-b_3) \times (1-p_3)$	$1-((1-b_4) \times (1-p_4))$	$(1-b_4) \times (1-p_4)$
1	1	0	$1-((1-Ab_3) \times (1-Ac_3))$	$(1-Ab_3) \times (1-Ac_3)$	$1-((1-Ab_2) \times (1-Ac_2))$	$(1-Ab_2) \times (1-Ac_2)$	$1-((1-b_1) \times (1-c_1))$	$(1-b_1) \times (1-c_1)$	$1-((1-Ab_1) \times (1-Ac_1))$	$(1-Ab_1) \times (1-Ac_1)$	$1-((1-b_2) \times (1-c_2))$	$(1-b_2) \times (1-c_2)$	$1-((1-b_3) \times (1-c_3))$	$(1-b_3) \times (1-c_3)$	$1-((1-b_4) \times (1-c_4))$	$(1-b_4) \times (1-c_4)$
1	1	1	$1-((1-Ap_3) \times (1-Ac_3) \times (1-Ab_3))$	$(1-Ap_3) \times (1-Ac_3) \times (1-Ab_3)$	$1-((1-Ap_2) \times (1-Ac_2) \times (1-Ab_2))$	$(1-Ap_2) \times (1-Ac_2) \times (1-Ab_2)$	$1-((1-p_1) \times (1-c_1) \times (1-b_1))$	$(1-p_1) \times (1-c_1) \times (1-b_1)$	$1-((1-Ap_1) \times (1-Ac_1) \times (1-Ab_1))$	$(1-Ap_1) \times (1-Ac_1) \times (1-Ab_1)$	$1-((1-p_2) \times (1-c_2) \times (1-b_2))$	$(1-p_2) \times (1-c_2) \times (1-b_2)$	$1-((1-p_3) \times (1-c_3) \times (1-b_3))$	$(1-p_3) \times (1-c_3) \times (1-b_3)$	$1-((1-p_4) \times (1-c_4) \times (1-b_4))$	$(1-p_4) \times (1-c_4) \times (1-b_4)$

Fonte: Autor

4.5 CONSIDERAÇÕES FINAIS

Neste capítulo, foi descrito, em detalhes, como a abordagem de detecção de corpo completa, o comitê de especialistas e a rede bayesiana trabalha na tentativa de detecção da forma humana, utilizando apenas o CENTRIST como atributo. Também foram apresentados todos os parâmetros possíveis de atuação no algoritmo, que possibilitam afetar diretamente o desempenho (exatidão vs. tempo de processamento). No próximo capítulo, serão apresentados os resultados dos experimentos, por meio de alguns conjuntos de imagens criados para esta dissertação e outros adquiridos pela internet, em situações de ambientes diversos.

5 EXPERIMENTOS E RESULTADOS

Neste capítulo, serão discutidos os resultados obtidos durante os experimentos realizados com os conjuntos de imagens, formados por três grupos distintos: conjunto A (imagens feitas em ambiente interno), conjunto B (imagens feitas por um quadrotor em ambiente externo) e conjunto C (imagens de cenas de desastres retiradas da internet). Os resultados também serão comparados com os de outras propostas. Inicialmente, serão descritos os bancos de imagens criados para uso no treinamento dos classificadores e os conjuntos de imagens usados para testes.

5.1 BANCO DE IMAGENS E CONJUNTOS DE TESTES

A realização deste trabalho ocorreu após a criação de um banco de imagens voltado especificamente para a aplicação do treinamento dos classificadores do corpo humano completo e parcial (Braços, Cabeças e Pernas). Essas imagens foram capturadas em ambiente interno, com iluminação artificial fluorescente simples; foram ainda capturadas outras com iluminação natural. Todas são coloridas e com resolução de 1280 x 720 pixels, com o plano de fundo retirado por edição de imagens. Elas tentam representar um variado número de poses e orientações de direções (90°, 180° e 270°), em um total de 1000 imagens, figuras 48 e 49. No caso dos classificadores de partes do corpo, foram feitas edições no banco de 1000 imagens, extraíndo-se 278 amostras de braços, 162 de cabeças e 282 de pernas; novamente, foram exploradas diferentes orientações (90°, 180° e 270°), figuras 50 e 51. Com relação aos conjuntos de imagens usados para os testes de desempenho do algoritmo proposto são apresentados a seguir:

- a) conjunto A: usado para teste, é formado por imagens coloridas com resolução 1280 x 720 pixels, feitas em ambiente interno, simulando cenas de pessoas caídas e objetos diversos, figuras 52 e 53. Esse conjunto apresenta 69 imagens (48 planos de fundo e 21 corpos humanos) de uma única pessoa, simulando possíveis cenas de desastres em ambiente iluminado artificialmente por lâmpadas fluorescentes, capturadas manualmente, a uma média de 2,5 a 3 metros de altura da cena focada. As cenas tentam abranger situações que vão de nenhuma oclusão até quase total oclusão, e é o conjunto que apresenta melhores condições para a tarefa de detecção;
- b) conjunto B: também usado nos testes, é formado por imagens coloridas, com resolução de 240 x 320 pixels, capturadas de vídeos a partir de um

quadrotor AR.Drone da Parrot SA., figura 54, disponibilizado pelo Centro Universitário da Fundação Educacional Inaciana Pe. Sabóia de Medeiros (FEI). Esse conjunto contém 71 imagens (37 planos de fundo e 34 corpos humanos), de três a quatro pessoas, simulando cenas de desastre em campo aberto, FEI, figuras 55 e 56, e capturadas a alturas que variavam de cinco a 20 metros. O conjunto não apresenta grandes complicadores de oclusões; porém, tem baixa resolução, saturação de iluminação, distorções nas imagens (devido a interferências), imagens duplas (câmera frontal) retiradas por edição e uma diversidade de alturas na captura de cenas, que levam a falsas detecções.

- c) conjunto C: constituído por 49 imagens (29 planos de fundo e 20 imagens contendo corpos diversos), extraídas da internet, com resolução de 240 x 160 pixels. Representam cenas reais do desastre ocorrido nos Alpes Franceses, da empresa aérea alemã Germanwings (2015), e também a destruição da aeronave MH17, da Malaysia Airlines, no leste da Ucrânia, figuras 57 e 58. Esse conjunto reúne as condições mais desafiadoras de avaliação do algoritmo proposto, visto que apresenta situações diversas de oclusões em cenas que, em alguns casos, fogem à situação desejável de captura das imagens proposta neste trabalho, ou seja, de cima para baixo (imagens aéreas).

Figura 48 - Imagem de corpo completo com Iluminação artificial.



Fonte: Autor.

Figura 49 - Imagem de corpo completo com Iluminação natural



Fonte: Autor.

Figura 50 - Imagem de parte do corpo: Braço.



Fonte: Autor.

Figura 51 - Imagem de parte do corpo: Perna.



Fonte: Autor.

Figura 52 - Imagem de ambiente interno1.



Fonte: Autor.

Figura 53 - Imagem de ambiente interno2.



Fonte: Autor.

Figura 54 - QuadrotorAR.Drone da Parrot SA



Fonte: FEI, laboratório mestrado 2013.

Figura 55 - Imagem feita pelo quadrotor (FEI).



Fonte: Autor.

Figura 56 - Imagem feita pelo quadrotor (FEI).



Fonte: Autor.

Figura 57 - Imagem real de desastre.



Fonte: Google Imagens.

Figura 58 - Imagem real de desastre



Fonte: Google Imagens.

5.2 INDICADORES DE AVALIAÇÃO DE DESEMPENHO

A forma mais eficaz de avaliar o desempenho de classificadores supervisionados é o uso da matriz de confusão, ao mostrar o número de classificações corretas versus as classificações preditas, para cada classe, sobre um conjunto K de exemplos. O número de acertos para cada classe está posicionado na diagonal principal $MF(d_i, d_i)$ da matriz, enquanto as demais posições (elementos) $MF(d_i, d_j)$, para $i \neq j$, representam erros na classificação. Na figura 59, adiante, é apresentada a matriz de confusão, com os principais erros de classificação usados neste trabalho, orbitando fora da diagonal principal:

Figura 59 - Matriz de confusão.

		MATRIZ DE CONFUSÃO				
		População total (TPP) = $TP + TN + FP + FN$	Condição Positiva (CP) = $TP + FN$	Condição Negativa (CN) = $FP + TN$		
RESULTADO DO TESTE (RT)	Resultado de Teste Positivo (RTP) = $TP + FP$	Verdadeiro Positivo (TP)	Falso Positivo (FP)	Preditivo positivo (PPV) ou Precisão = $\frac{\sum TP}{\sum RTP}$	Taxa de detecção falsa (FDR) = $\frac{\sum FP}{\sum RTP}$	
	Resultado de Teste Negativo (RTN) = $TN + FN$	Falso Negativo (FN)	Verdadeiro Negativo (TN)	Taxa de Omissão Falsa (FOR) = $\frac{\sum FN}{\sum RTN}$	Preditivo negativo (NPV) = $\frac{\sum TN}{\sum RTN}$	
	Exatidão (ACC) = $\frac{\sum TP + \sum TN}{\sum TPP}$	Taxa de Positivo Verdadeiro (TPR) ou <i>Sensitividade</i> , Recall = $\frac{\sum TP}{\sum CP}$	Taxa de falso positivo (FPR), Fall-out = $\frac{\sum FP}{\sum CN}$	Taxa de falso negativo Miss Rate (FNR) = $\frac{\sum FN}{\sum CP}$	Taxa de verdadeiro negativo (TNR) ou <i>Especificidade</i> (SPC) = $\frac{\sum TN}{\sum CN}$	

Fonte: "Adaptado de", Confusionmatrix; WIKIPEDIA.

Os erros de classificação são definidos como:

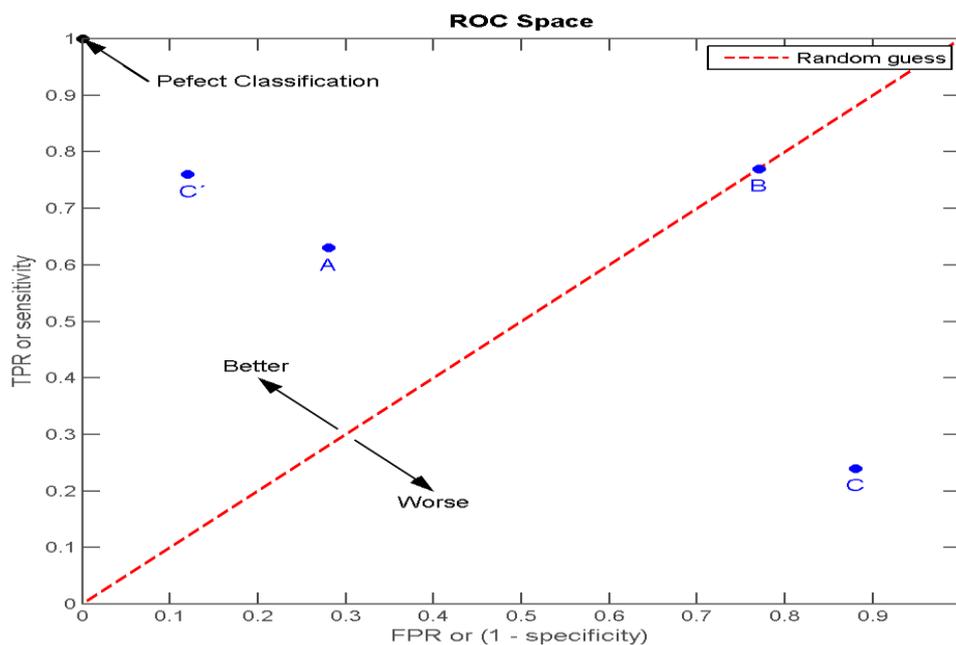
- a) sensibilidade (TPR) ou Recall = porcentagem de amostras corretamente classificadas como positivas, sobre o total de amostras positivas (CP);
- b) especificidade (SPC) = porcentagem de amostras corretamente classificadas como negativas, sobre o total de amostras negativas (CN);
- c) taxa de falso positivo (FPR) ou Fall-out ($1-SPC$) = porcentagem de amostras incorretamente classificadas como positivas, sobre o total de amostras negativas (CN);
- d) taxa de falso negativo (FNR) ou Miss Rate = porcentagem de amostras incorretamente classificadas como negativas, sobre o total de amostras positivas (CP);
- e) acurácia (ACC) = porcentagem de amostras positivas e negativas corretamente classificadas, sobre o total de amostras positivas e negativas (TPP);
- f) valor preditivo positivo (PPV) ou precisão = porcentagem de amostras corretamente classificadas como positivas, sobre o total de amostras classificadas como positivas (RTP);
- g) valor preditivo negativo (NPV) = porcentagem de amostras classificadas corretamente como negativas, sobre o total de amostras classificadas como negativas (RTN);
- h) taxa de detecção falsa (1-PPV) ou (FDR) = porcentagem de amostras classificadas incorretamente como positivas, sobre o total de amostras classificadas como positivas (RTP);
- i) taxa de omissão falsa (FOR) = porcentagem de amostras classificadas incorretamente como negativas, sobre o total de amostras classificadas como negativas (RTN).

A curva característica de operação do receptor, ou “*Receiver Operating Characteristic*” (ROC), figura 60, é outra forma de apresentação dos dados; porém, gráfica, mostrando a relação de compromisso entre as taxas de TP (TPR) e FP (FPR) de um classificador. Em uma curva ROC ideal com o valor de $TPR = 1$ e $FPR = 0$, todas as amostras positivas e negativas são classificadas corretamente. No caso de $TPR = FPR = 0$ ou $TPR = FPR = 1$, são classificadores que, respectivamente, classificam todas as amostras como negativas ou positivas. No caso em que $TPR = 0$ e $FPR = 1$, o classificador classifica incorretamente todas as amostras. Um eficiente e rápido indicador do desempenho

de um classificador é calcular a área AUC, abaixo da curva ROC, em que, quanto maior esse valor, melhor seu desempenho global.

Neste trabalho com uso de comitê de especialistas, existem parâmetros a serem ajustados para melhorar a relação de compromisso entre as taxas de TP (TPR) e FP (FPR), de modo que cada ajuste forneça um par (FP,TP) para construção da curva ROC, conforme figura a seguir:

Figura 60 - Espaço da curva ROC , com pontos A e C com boa classificação, ponto B com classificação aleatória e ponto C com péssima classificação.



Fonte: ROC curve; WIKIPEDIA.

5.3 RESULTADO DOS INDICADORES DE DESEMPENHO INDIVIDUAL DE CADA CLASSIFICADOR

Nesta seção, será apresentado o método de treinamento dos classificadores de corpo completo, das partes e os valores dos indicadores de desempenho obtidos. O tamanho total das amostras, para uso nos testes, está na tabela 6, e nas demais tabelas, 7, 8, 9 e 10, os valores obtidos dos erros dos classificadores SVM (linear e HIK) e real adaboost, o treinamento esta descrito na figura 61 para obtenção dos modelos.

Uma importante ação para obter resultados confiáveis durante a classificação é o treinamento dos classificadores. Nessa fase, um banco de imagens próprio foi criado, contendo 1000 imagens positivas P , de corpo completo, em diversos posicionamentos,

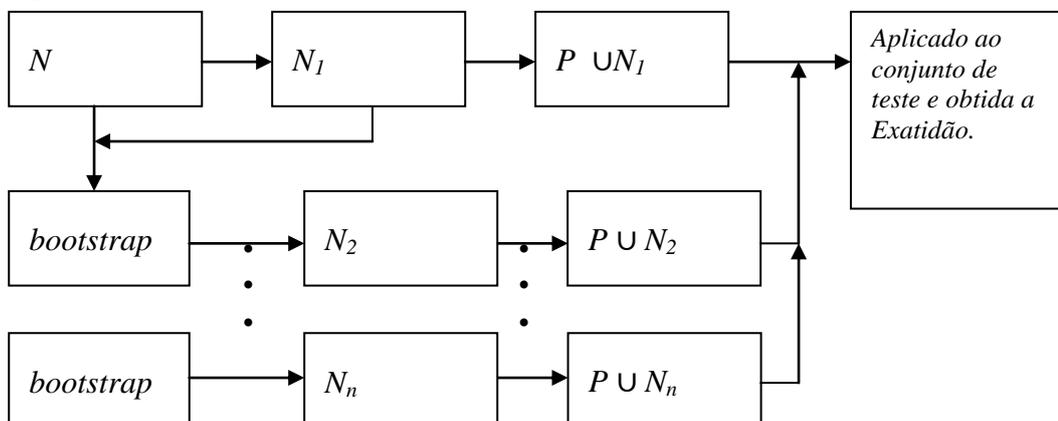
redimensionadas para a janela de detecção (w_d , h_d). O conjunto de imagens negativas foi criado a partir de um conjunto de 1500 imagens diversas, que não continham formas humanas, e eram exploradas de forma aleatória, gerando um total de 6000 imagens negativas N .

No início, figura 61, escolheu-se um conjunto aleatório correspondente a 2/3 das amostras positivas P , e 2/3 do conjunto N para o treinamento, enquanto 1/3 de P e N (P_T, N_T) era usado para teste. No conjunto formado por $P \cup N$, usado para o treinamento, foi escolhido um pequeno conjunto aleatório $N_1 \in N$, formando um novo conjunto $P \cup N_1$ e treinando o classificador SVM-Linear. O resultado desse treinamento, $R_{1(\text{modelo})}$, foi aplicado ao conjunto de teste (P_T, N_T) e sua exatidão avaliada. O processo de *bootstrap* (EFRON; 1979), com reposição de amostras, é aplicado a N e é gerado um novo conjunto N_2 . Ao novo conjunto $P \cup N_2$, é realizado um novo treinamento no classificador SVM-Linear, resultando $R_{2(\text{modelo})}$ e uma nova avaliação da exatidão. Este processo é repetido inúmeras vezes, até obter um valor que quase não tenha mudanças significativas no valor final da exatidão.

Após o treinamento final do classificador SVM-Linear, obtido o modelo, o processo é repetido para o treinamento do classificador SVM-HIK, porém, partindo do último conjunto obtido da combinação de $P \cup N_{\text{final}}$, no caso do Adaboost o processo é idêntico ao SVM-Linear. Quanto ao procedimento de treinamento dos classificadores de partes, é idêntico àquele utilizado para os classificadores de corpo completo, usando as mesmas 6000 amostras negativas. Todo esses processos de treinamento (SVM-Linear e HIK) para obtenção dos modelos são realizados utilizando a biblioteca libHIK-2.7 (WU; 2010).

A figura 61 segue o diagrama de blocos do processo de treinamento aplicado neste trabalho.

Figura 61 - Processo de treinamento dos classificadores.



Fonte: Autor.

Tabela 6 - Tamanho total das amostras.

AMOSTRAS	QUANTIDADE	TAMANHO DAS AMOSTRAS TREINAMENTO	TAMANHO DAS AMOSTRAS TESTE
NEGATIVAS	6000	4000	2000
CORPO COMPLETO	1000	660	340
BRAÇO	278	180	98
CABEÇA	162	93	69
PERNA	282	185	97

Fonte: Autor.

Tabela 7 - Indicadores para a amostra de corpo completo.

CORPO COMPLETO			
	SVM-Linear	SVM-HIK	Adaboost
FN	90	98	90
FP	147	122	467
TP	250	242	250
TN	1853	1878	1533
TPR	0,735294118	0,711764706	0,735294118
SPC	0,9265	0,939	0,7665
PPV	0,629722922	0,664835165	0,348675035
NPV	0,953679876	0,950404858	0,944547135
FPR	0,0735	0,061	0,2335
FNR	0,264705882	0,288235294	0,264705882
ACC	0,898717949	0,905982906	0,761965812
FDR	0,370277078	0,335164835	0,651324965
FOR	0,046320124	0,049595142	0,055452865

Fonte: Autor.

Tabela 8 - Indicadores para a amostra Braço.

	BRAÇOS		
	Dtree	NNT	SVM
FN	13	15	20
FP	240	230	301
TP	85	83	78
TN	1761	1771	1700
TPR	0,867347	0,846939	0,795918
SPC	0,88006	0,885057	0,849575
PPV	0,261538	0,265176	0,205805
NPV	0,992672	0,991601	0,988372
FPR	0,11994	0,114943	0,150425
FNR	0,132653	0,153061	0,204082
ACC	0,879466	0,883278	0,84707
FDR	0,738462	0,734824	0,794195
FOR	0,007328	0,008399	0,011628

Fonte: Autor.

Tabela 9 - Indicadores para a amostra Cabeça.

	CABEÇAS		
	Dtree	NNT	SVM
FN	7	3	10
FP	51	116	187
TP	62	66	59
TN	1950	1885	1814
TPR	0,898551	0,956522	0,855072
SPC	0,974513	0,942029	0,906547
PPV	0,548673	0,362637	0,239837
NPV	0,996423	0,998411	0,994518
FPR	0,025487	0,057971	0,093453
FNR	0,101449	0,043478	0,144928
ACC	0,971981	0,942512	0,904831
FDR	0,451327	0,637363	0,760163
FOR	0,003577	0,001589	0,005482

Fonte: Autor.

Tabela 10 - Indicadores para a amostra Perna.

PERNAS			
	Dtree	NNT	SVM
FN	23	19	26
FP	181	276	338
TP	74	78	71
TN	1820	1725	1663
TPR	0,762887	0,804124	0,731959
SPC	0,909545	0,862069	0,831084
PPV	0,290196	0,220339	0,173594
NPV	0,98752	0,989106	0,984606
FPR	0,090455	0,137931	0,168916
FNR	0,237113	0,195876	0,268041
ACC	0,902765	0,85939	0,826501
FDR	0,709804	0,779661	0,826406
FOR	0,01248	0,010894	0,015394

Fonte: Autor.

A tabela 6 apresenta os tamanhos das amostras utilizadas durante o procedimento descrito de treinamento dos classificadores; corpo completo (SVM-Linear, SVM-HIK e AdaBoost), e dos classificadores de partes do corpo (Dtree-Árvore de Decisão, NNT-Rede Neural Artificial e SVM-Máquina de Vetores de Suporte-Linear) e das amostras usadas para os testes de desempenho. E a tabela 7 apresenta os valores dos indicadores de desempenho dos classificadores de corpo completo. Nota-se que a exatidão (ACC) do classificador SVM-HIK é a que mostra melhor desempenho.

As tabelas 8, 9 e 10 apontam os valores dos indicadores de desempenho, obtidos dos classificadores (Dtree, NNT e SVM), usados para o treinamento das partes do corpo humano (Braços, Cabeça e Pernas). Observa-se que as exatidões (em negrito) são os valores usados como nó a priori nas tabelas de probabilidades condicionais (TPCs), na formação da rede Bayesiana, apresentada na seção 4.4. Ademais, que os classificadores com melhor desempenho são aqueles relativos ao atributo (Cabeça), por serem mais discriminantes do que os atributos (Braços) e (Pernas), que são similares.

5.4 RESULTADO DO OU-RUIDOSO NA REDE BAYESIANA

A seguir, são apresentadas as Tabelas de Probabilidades Condicionais (TPCs) tabelas 11, 12, 13, e 14, usadas neste trabalho. Deve-se observar que os nós a priori (em vermelho) são obtidos a partir dos valores de exatidão referentes às tabelas 8, 9 e 10 da seção 5.3, e que as outras probabilidades são obtidas pela aplicação do OU-Ruidoso. Observa-se como exemplo que, para obter o valor de probabilidade da presença do atributo BRAÇO = 0.982148 (em verde), com os classificadores NNT (Rede Neural Artificial) e SVM-Linear detectando esse atributo e o classificador Dtree (Árvore de Decisão) não detectando, é necessária a operação $1 - (1 - ACC_{NNT} \times 1 - ACC_{SVM})$, ou seja, $1 - (0.11673 \times 0.15293) = 0.982148$:

Tabela 11-TPC aplicada para o nó Braço usando OU-Ruidoso.

Dtree	NNT	TPC - BRAÇO		
		SVM	BRAÇO	~ BRAÇO
0	0	0	0	1
0	0	1	0,84707	0,15293
0	1	0	0,88327	0,11673
0	1	1	0,982148	0,017852
1	0	0	0,87946	0,12054
1	0	1	0,981566	0,018434
1	1	0	0,985929	0,014071
1	1	1	0,997848	0,002152

Fonte: Autor.

Tabela 12-TPC aplicada para o nó Cabeça usando OU-Ruidoso.

Dtree	NNT	TPC-CABEÇA		
		SVM	CABEÇA	~ CABEÇA
0	0	0	0	1
0	0	1	0,9048	0,0952
0	1	0	0,9425	0,0575
0	1	1	0,994526	0,005474
1	0	0	0,97198	0,02802
1	0	1	0,997332	0,002668
1	1	0	0,998389	0,001611
1	1	1	0,999847	0,000153

Fonte: Autor.

Tabela 13 -TPC aplicada para o nó Perna usando OU-Ruidoso

Dtree	NNT	TPC-PERNA		
		SVM	PERNA	¬ PERNAS
0	0	0	0	1
0	0	1	0,8265	0,1735
0	1	0	0,85939	0,14061
0	1	1	0,975604	0,024396
1	0	0	0,90276	0,09724
1	0	1	0,983129	0,016871
1	1	0	0,986327	0,013673
1	1	1	0,997628	0,002372

Fonte: Autor.

A Tabela de Probabilidade Condicional (TPC), tabela 14 apresenta o resultado final da relação de probabilidades, entre as variáveis Braço, Cabeça e Perna (nós pais) e a variável nó filho, (Corpo Completo). Nota-se que, nas saídas S3, S5, S6 e S7 da tabela 14, há sete probabilidades de existência de qualquer uma das variáveis (Braço, Cabeça e Perna), mais a condição inicial 0, de não existência de qualquer uma delas (\neg Braço, \neg Cabeça e \neg Perna) levando em consideração a concordância dos classificadores, ou seja, quando estes assumem valor igual a 1 (colunas em amarelo). Os valores apresentados na cor vermelha são usados para a construção da TPC, pelo uso do OU-Ruidoso (nós a priori). A decisão da rede Bayesiana foi então ajustada a um valor maior ou igual 99.2% de probabilidade limiar (aceitação da existência de algum corpo humano na imagem analisada), pois, a presença de apenas uma parte do corpo não garante que haja um corpo completo, logo; o valor de probabilidade de 99.2%, obtido experimentalmente durante os testes de classificação, aumenta esta possibilidade pela união de um numero maior de partes.

Tabela 14- TPC aplicada para o nó (Corpo Completo HB) usando OU-Ruidoso

VARIÁVEL			SAÍDA: S1		SAÍDA: S2		SAÍDA: S3		SAÍDA: S4		SAÍDA: S5		SAÍDA: S6		SAÍDA: S7	
BRAÇO	CABEÇA	PERNA	CORPO	→ CORPO	CORPO	→ CORPO	CORPO	→ CORPO	CORPO	→ CORPO	CORPO	→ CORPO	CORPO	→ CORPO	CORPO	→ CORPO
0	0	0	0	1	0	1	0	1	0	1	0	1	0	1	0	1
0	0	1	0,8265	0,1735	0,8593	0,1407	0,9756	0,0244	0,9027	0,0973	0,9831	0,0169	0,9863	0,0137	0,9976	0,0024
0	1	0	0,9048	0,0952	0,9425	0,0575	0,9945	0,0055	0,9719	0,0281	0,9973	0,0027	0,9983	0,0017	0,9998	0,0002
0	1	1	0,98348	0,01652	0,99191	0,00809	0,99987	0,000134	0,99727	0,00273	0,99995	4,563E-05	0,99998	2,33E-05	1	4,8E-07
1	0	0	0,84707	0,15293	0,8832	0,1168	0,9821	0,0179	0,8794	0,1206	0,9815	0,0185	0,9859	0,0141	0,9978	0,0022
1	0	1	0,97347	0,02653	0,98357	0,016434	0,99956	0,000437	0,98827	0,01173	0,99969	0,0003127	0,99981	0,000193	0,99999	5,28E-06
1	1	0	0,98544	0,01456	0,99328	0,006716	0,9999	9,84E-05	0,99661	0,00339	0,99995	4,995E-05	0,99998	2,4E-05	1	4,4E-07
1	1	1	0,99747	0,00253	0,99906	0,000945	1	2,4E-06	0,99967	0,00033	1	8,442E-07	1	3,28E-07	1	1,06E-09

Fonte: Autor

5.5 EXPERIMENTOS

Os experimentos objetivaram a analisar os resultados do algoritmo proposto, aplicado sobre cada um dos conjuntos de imagens descritos na seção 5.1 (Banco de Imagens e Conjuntos de Testes), e comparar esses resultados a alguns trabalhos relevantes, já discutidos no início deste trabalho, no capítulo 3 (Trabalhos Relacionados). Os parâmetros de atuação no algoritmo principal foram explorados, de maneira a obter o melhor desempenho de exatidão e velocidade de processamento. A tabela 15 apresenta todos os parâmetros envolvidos diretamente nesse algoritmo; mas, o desempenho foi avaliado segundo a velocidade de deslocamento da janela de detecção Δ ($\Delta = 3$ ou $\Delta = 5$) e o redimensionamento dessa σ ($\sigma = 0.9$ ou $\sigma = 0.8$), valores respectivos para menor velocidade de processamento; porém, maior precisão de captura de detalhes, ou maior velocidade de processamento com menor precisão de captura. O valor de β (ajuste de ponderação dos gradientes) foi ajustado, de forma que o algoritmo possa se “adaptar” à iluminação do ambiente e obter a informação a ser processada.

A construção de cada curva ROC foi obtida pelo ajuste de β , nos conjuntos de testes (A, B e C), para as duas condições de velocidade de processamento, procurando minimizar a taxa de falso positivos (FPR) e falsos negativos (FNR), maximizar a taxa de verdadeiro positivo (TPR) e maximizar a exatidão (ACC). As tabelas que se seguem apresentam os valores dos indicadores FPR, FNR, TPR e ACC para as duas opções de classificadores em cascata: (SVM-Linear + Adaboost) ou (SVM-Linear + HIK). Os demais indicadores usados na matriz de confusão não serão usados para análise, porém estão disponíveis no apêndice B.

Tabela 15 - Parâmetros do algoritmo.

Símbolo	Descrição	Valor
β	Ajuste de ponderação dos gradientes (GX, GY)	(ajustável)
Δ	Passo de deslocamento da janela de detecção.	$\Delta = 3$ ou $\Delta = 5$
σ	Passo de redimensionamento da janela de detecção	$\sigma = 0.9$ ou $\sigma = 0.8$
Φ_1	Limiar de trabalho do SVM-Linear	$\Phi_1 \geq 3$
Φ_2	Limiar de trabalho do SVM-HIK	$\Phi_2 \geq 50$
P	Probabilidade limiar da rede Bayesiana	0.992

Fonte: Autor.

5.5.1 Experimentos com o Conjunto de Imagens Indoor (A)

Para o primeiro experimento, conjunto (A) foi utilizado, seção 5.1, com 69 (48 plano de fundo e 21 corpos) imagens feitas em ambiente interno, simulando cenas de indivíduo caído em situações diversas. Como a resolução das imagens é elevada (1280 x 720) e há iluminação local artificial adequada, observa-se que o nível de contraste é também elevado, ou seja, é a condição ideal para a perfeita discriminação entre as formas.

No experimento com o classificador SVM-Adaboost, o nível dos gradientes (β) foi ajustado, de forma gradual ($\beta = 0.0010$ a $\beta = 0.0020$), partindo do menor contraste detectável até próximo à saturação, de modo a alcançar o melhor desempenho. As tabelas 16 e 17 representam os indicadores FPR, FNR, TPR e ACC, onde os valores V1, V2, V3, V4 e V5 são os pontos de ajuste de β . Como a variação de β não corresponde de forma linear aos resultados obtidos (para $\beta = 0.00180$ tem-se TPR = 0.952 e $\beta = 0.001$ tem-se TPR = 0 na condição de velocidade de processamento menor com $\Delta = 3$ e $\sigma = 0.9$), qualquer valor fora dessa faixa provoca grandes alterações. Na condição de maior velocidade de processamento ($\Delta = 5$ e $\sigma = 0.8$), os resultados foram além das expectativas, pois era esperada uma queda no desempenho; porém, foi mantido e ainda reduzido o tempo de processamento em 50% (de $\cong 1.0$ seg. para 0.5 seg. por foto) observar tabela 18. A curva ROC de cada tabela 16 e 17 é apresentada na figura 62 e 63, e o resultado da detecção, imagens classificadas, na figura 64:

Tabela 16- Indicadores (SVM-Adaboost) para o conjunto (A) com $\Delta = 3$ e $\sigma = 0.9$.

AdaBoost Conj. A ($\Delta = 3$ $\sigma = 0.9$)					
	V1	V2	V3	V4	V5
TPR	0	0,38	0,762	0,81	0,952
FPR	0,021	0,04	0,167	0,23	0,438
FNR	1	0,62	0,238	0,19	0,048
ACC	0,681	0,78	0,812	0,78	0,681

Fonte: Autor.

Tabela 17 - Indicadores (SVM-Adaboost) para o conjunto (A) com $\Delta = 5$ e $\sigma = 0.8$.

AdaBoost Conj. A ($\Delta = 5$ $\sigma = 0.8$)					
	V1	V2	V3	V4	V5
TPR	0	0,19	0,524	0,81	0,905
FPR	0,021	0,06	0,063	0,19	0,313
FNR	1	0,81	0,476	0,19	0,095
ACC	0,681	0,71	0,812	0,81	0,754

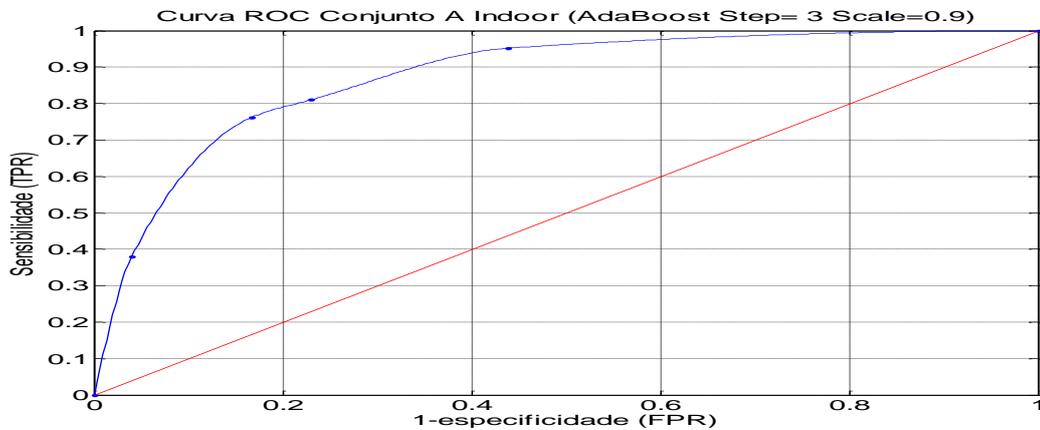
Fonte: Autor.

Tabela 18- Tempo de processamento para o conjunto (A) com (SVM-Adaboost)

$\Delta = 3$ e $\sigma = 0.9$.	1 imagen/seg.
$\Delta = 5$ e $\sigma = 0.8$.	2 imagens/seg.

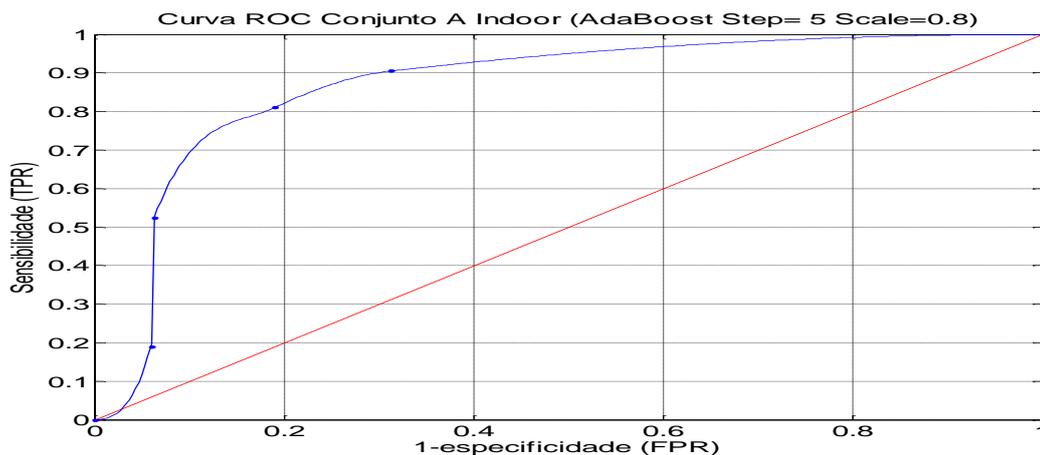
Fonte: Autor.

Figura 62-Curva ROC do classificador em cascata (SVM-Adaboost) na condição $\Delta = 3$ e $\sigma = 0.9$ para o conjunto A.



Fonte: Autor.

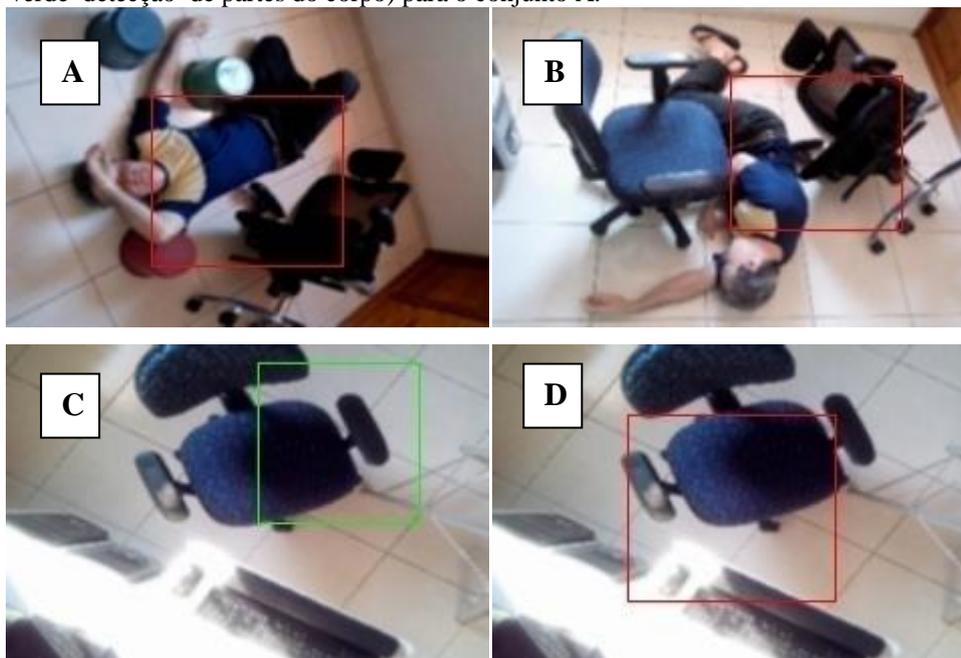
Figura 63 - Curva ROC do classificador em cascata (SVM-Adaboost) na condição $\Delta = 5$ e $\sigma = 0.8$ para o conjunto A.



Fonte: Autor.

Nas tabelas 16 e 17, são apresentados os parâmetros de desempenho. Observa-se que, nessas tabelas para o conjunto (A-Indoor), as exatidões são semelhantes, próximas a uma taxa de verdadeiro positivo (TPR = 0.81). Contudo, a tabela 17 apresenta uma taxa de falsos positivos menor (FPR), neste ponto e nos demais. A curva ROC, figura 63, da tabela 17 mostra, de forma clara, que há uma resposta mais eficiente na condição $\Delta = 5$ e $\sigma = 0.8$, já que a subida da curva está mais rápida no intervalo (0 – 0.2). Esse resultado positivo obtido na condição $\Delta = 5$ e $\sigma = 0.8$ (maior velocidade de processamento) não era esperado, pois os detalhes foram perdidos. Todavia, como o conjunto (A) apresenta as melhores condições de contraste e iluminação, isso faz a diferença na detecção.

Figura 64 - Saída do classificador SVM-Adaboost (Em vermelho detecção de corpo, em verde detecção de partes do corpo) para o conjunto A.



Fonte: Autor.

Na figura 64, nota-se a saída da classificação, SVM-Adaboost, com alguns resultados representados por um quadrado em vermelho, na detecção de corpo completo e em verde, a classificação de partes do corpo. As figuras 64-A e B são classificações corretas, a figura 64-C é detecção de membro incorreta, pois detectou o braço da cadeira como parte de um corpo (Braço ou Perna). A figura 64-D é a detecção incorreta de corpo completo, ocasionada, provavelmente, como resultado de um amplo conjunto de treinamento, incluindo posições semi-sentadas.

Os resultados obtidos com o classificador SVM-HIK são apresentados na tabela 19 e 20, com as respectivas curvas ROC nas figuras 65 e 66, e o resultado da detecção, imagens classificadas, na figura 67. Aqui, também o nível dos gradientes (β) foi ajustado de forma gradual ($\beta=0.001$ a $\beta=0.003$), partindo do menor contraste detectável até próximo à saturação, de modo a alcançar o melhor desempenho. Nota-se que, no caso do SVM-HIK, nas tabelas 19 e 20 e que representam os indicadores FPR, FNR, TPR e ACC, onde os valores V1, V2, V3, V4 e V5 são os pontos de ajuste de β e nas curvas ROC, figuras 65 e 66, houve uma queda no valor da exatidão (de 0.8 para $\cong 0.6$) e um aumento do FPR (0.19 para $\cong 0.4$), comparado ao SVM-Adaboost nas tabelas 16 e 17, próximo ao valor de TPR = 0.8. Todavia, há ganhos na velocidade de processamento, pois, enquanto o classificador SVM-Adaboost na condição ($\Delta = 3$ e $\sigma = 0.9$) trabalha em torno de uma imagem por segundo (1 Imagem /seg.), o classificador SVM-HIK trabalha com duas (2 Imagens/seg.), observar tabela 21. No caso da condição ($\Delta =$

5 e $\sigma = 0.8$), essa diferença sobe para cinco (5 imagens/seg.), observar tabela 21. Esta situação é verdadeira caso o algoritmo de corpo completo não recorra à detecção de partes do corpo:

Tabela 19 - Indicadores (SVM-HIK) para o conjunto (A) com $\Delta = 3$ e $\sigma = 0.9$.

SVM HIK Conj. A ($\Delta = 3$ $\sigma = 0.9$)					
	V1	V2	V3	V4	V5
TPR	0	0,571	0,714	0,86	0,95
FPR	0,042	0,125	0,146	0,52	0,69
FNR	1	0,429	0,286	0,14	0,05
ACC	0,667	0,783	0,812	0,59	0,51

Fonte: Autor.

Tabela 20 - Indicadores (SVM-HIK) para o conjunto (A) com $\Delta = 5$ e $\sigma = 0.8$

SVM HIK Conj. A ($\Delta = 5$ $\sigma = 0.8$)					
	V1	V2	V3	V4	V5
TPR	0	0,667	0,762	0,81	0,86
FPR	0,021	0,146	0,25	0,38	0,71
FNR	1	0,333	0,238	0,19	0,14
ACC	0,681	0,797	0,754	0,68	0,46

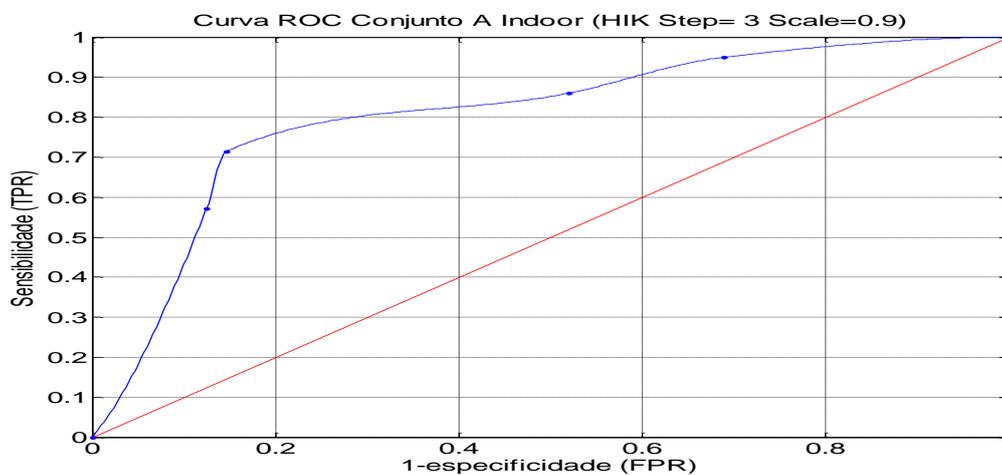
Fonte: Autor.

Tabela 21- Tempo de processamento para o conjunto (A) com (SVM-HIK)

$\Delta = 3$ e $\sigma = 0.9$.	2 imagens/seg.
$\Delta = 5$ e $\sigma = 0.8$.	5 imagens/seg.

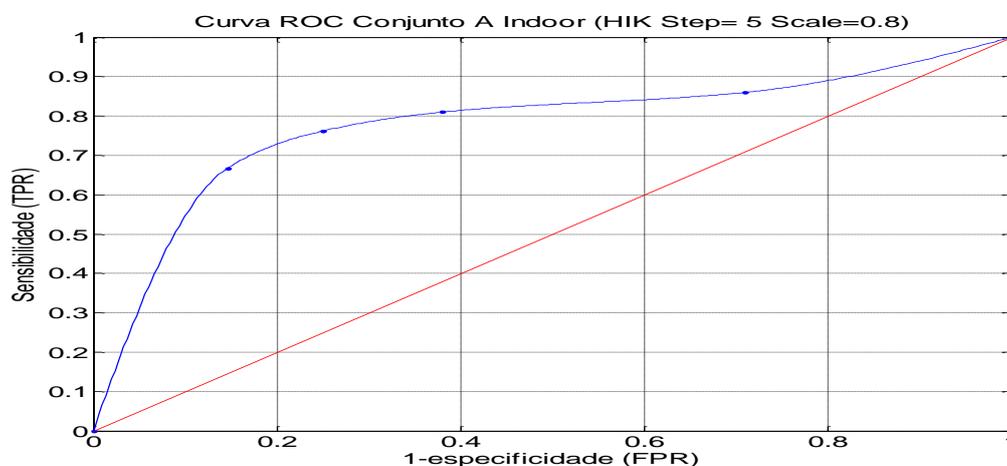
Fonte: Autor.

Figura 65 - Curva ROC do classificador em cascata (SVM-HIK) na condição $\Delta = 3$ e $\sigma = 0.9$ para o conjunto A.



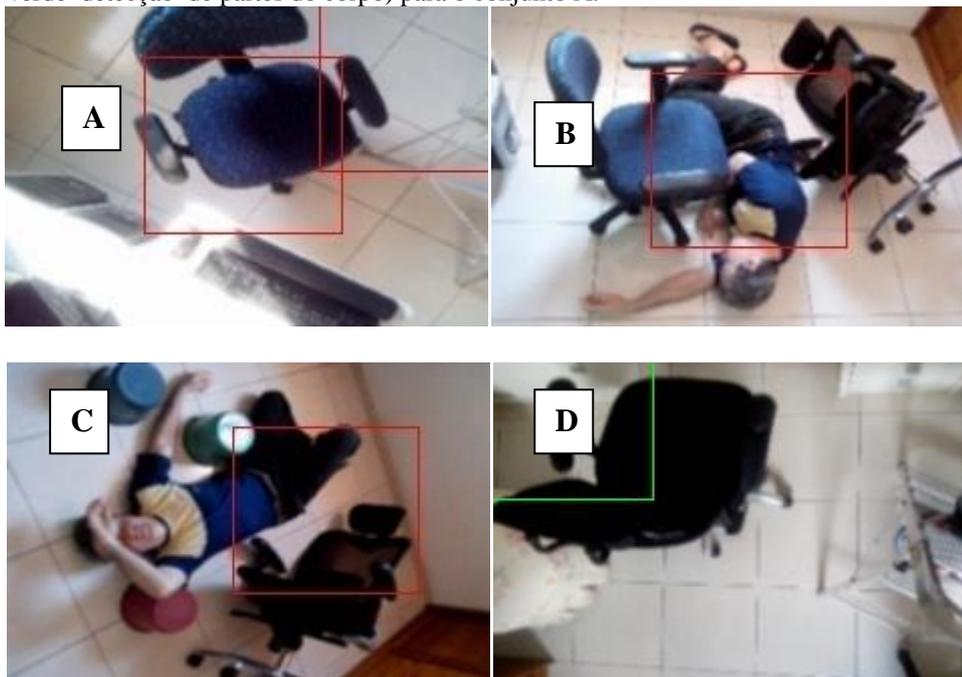
Fonte: Autor.

Figura 66 - Curva ROC do classificador em cascata (SVM-HIK) na condição $\Delta = 5$ e $\sigma = 0.8$ para o conjunto A.



Fonte: Autor.

Figura 67 - Saída do classificador SVM-HIK (Em vermelho detecção de corpo, em verde detecção de partes do corpo) para o conjunto A.



Fonte: Autor.

Na figura 67, observa-se a saída da classificação SVM-HIK, com alguns resultados representados por um quadrado em vermelho na detecção de corpo completo, e em verde, na classificação de partes do corpo. As figura 67-B e C mostram classificações corretas de corpo completo, enquanto a figura 67-A, uma detecção incorreta. A figura 67-D é a detecção de parte do corpo incorreta. As razões dessas detecções incorretas são as mesmas justificadas na figura 64.

5.5.2 Experimentos com o Conjunto de Imagens feitas pelo Drone (B)

No segundo experimento, o conjunto (B) foi utilizado, com imagens de resolução (240 x 320) feitas pelo quadrotor, formando 71 imagens (37 planos de fundo e 34 corpos humanos), em campo aberto, FEI, em diferentes altitudes.

Nesse experimento com o classificador SVM-AdaBoost, o nível dos gradientes (β) foi ajustado, de forma gradual ($\beta = 0.003$ a $\beta = 0.008$), até próximo à saturação, de modo a alcançar o melhor desempenho, tabelas 22 e 23. Aqui, a variação de β não corresponde de forma linear aos resultados obtidos, e diferentemente do conjunto (A), um aumento considerável desse valor foi necessário, já que as imagens fornecidas pelo quadrotor apresentam duas desvantagens para a detecção: baixa resolução e grande quantidade de iluminação, o que leva à saturação e ao baixo nível de contraste nas imagens. O reflexo dessa condição é visto nas tabelas 22 e 23 e que representam os indicadores FPR, FNR, TPR e ACC, onde os valores V1, V2, V3, V4 e V5 são os pontos de ajuste de β e na curva ROC, figuras 68 e 69, em que, para o classificador SVM-AdaBoost na condição ($\Delta = 3$ e $\sigma = 0.9$), o melhor desempenho é de TPR = 0.865, com FPR = 0.412. Na condição de maior velocidade de processamento ($\Delta = 5$ e $\sigma = 0.8$), foi obtido TPR = 0.622 com FPR = 0.235; ou seja, uma alta taxa de FPR é causa direta de ruídos, devido à saturação no contraste. Quanto ao tempo de processamento houve um aumento em relação ao conjunto (A), pois, com o reforço no nível de contraste, ou seja, uma maior quantidade de informação o tempo de processamento também aumenta, observar tabela 24. As figuras 70 e 71 apresentam o resultado da detecção do classificador SVM-AdaBoost:

Tabela 22- Indicadores (SVM-Adaboost) para o conjunto (B) com $\Delta = 3$ e $\sigma = 0.9$.

AdaBoost Conj. B ($\Delta = 3$ $\sigma = 0.9$)					
	V1	V2	V3	V4	V5
TPR	0	0,35	0,622	0,76	0,865
FPR	0,029	0,03	0,235	0,38	0,412
FNR	1	0,65	0,378	0,24	0,135
ACC	0,465	0,65	0,69	0,69	0,732

Fonte: Autor.

Tabela 23 - Indicadores (SVM-Adaboost) para o conjunto (B) com $\Delta = 5$ e $\sigma = 0.8$.

AdaBoost Conj. B ($\Delta = 5$ $\sigma = 0.8$)					
	V1	V2	V3	V4	V5
TPR	0	0,51	0,622	0,81	0,946
FPR	0,029	0,21	0,235	0,79	0,706
FNR	1	0,49	0,378	0,19	0,054
ACC	0,465	0,65	0,69	0,52	0,634

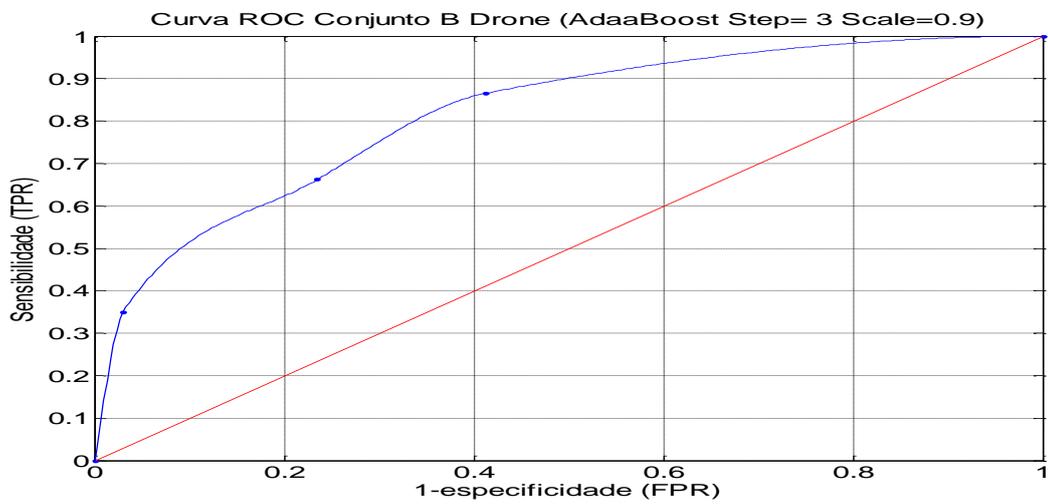
Fonte: Autor.

Tabela 24- Tempo de processamento para o conjunto (B) com (SVM-Adaboost)

$\Delta = 3$ e $\sigma = 0.9$.	0.5 imagens/seg.
$\Delta = 5$ e $\sigma = 0.8$.	1 imagens/seg.

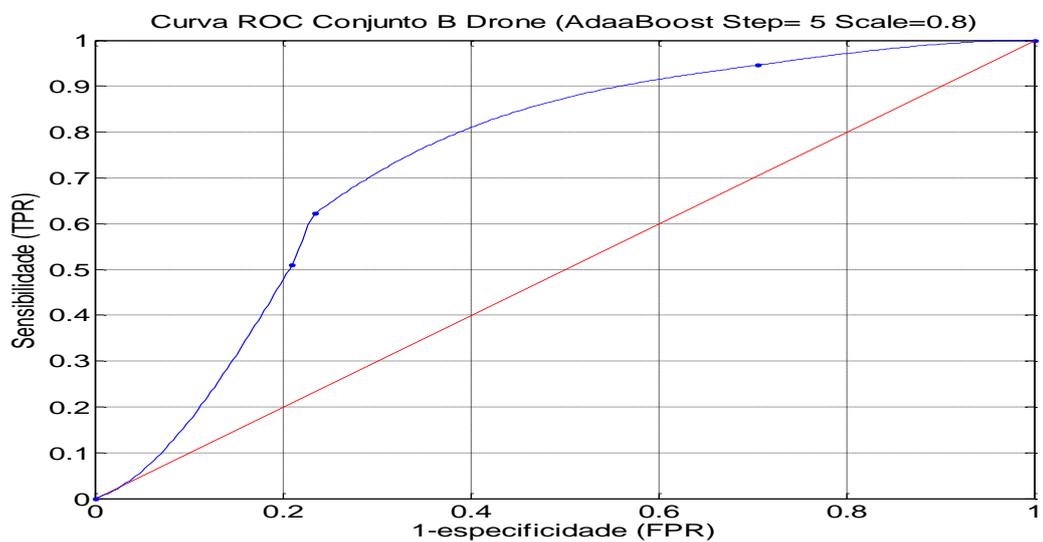
Fonte: Autor.

Figura 68 - Curva ROC do classificador em cascata (SVM-Adaboost) na condição $\Delta = 3$ e $\sigma = 0.9$ para o conjunto B.



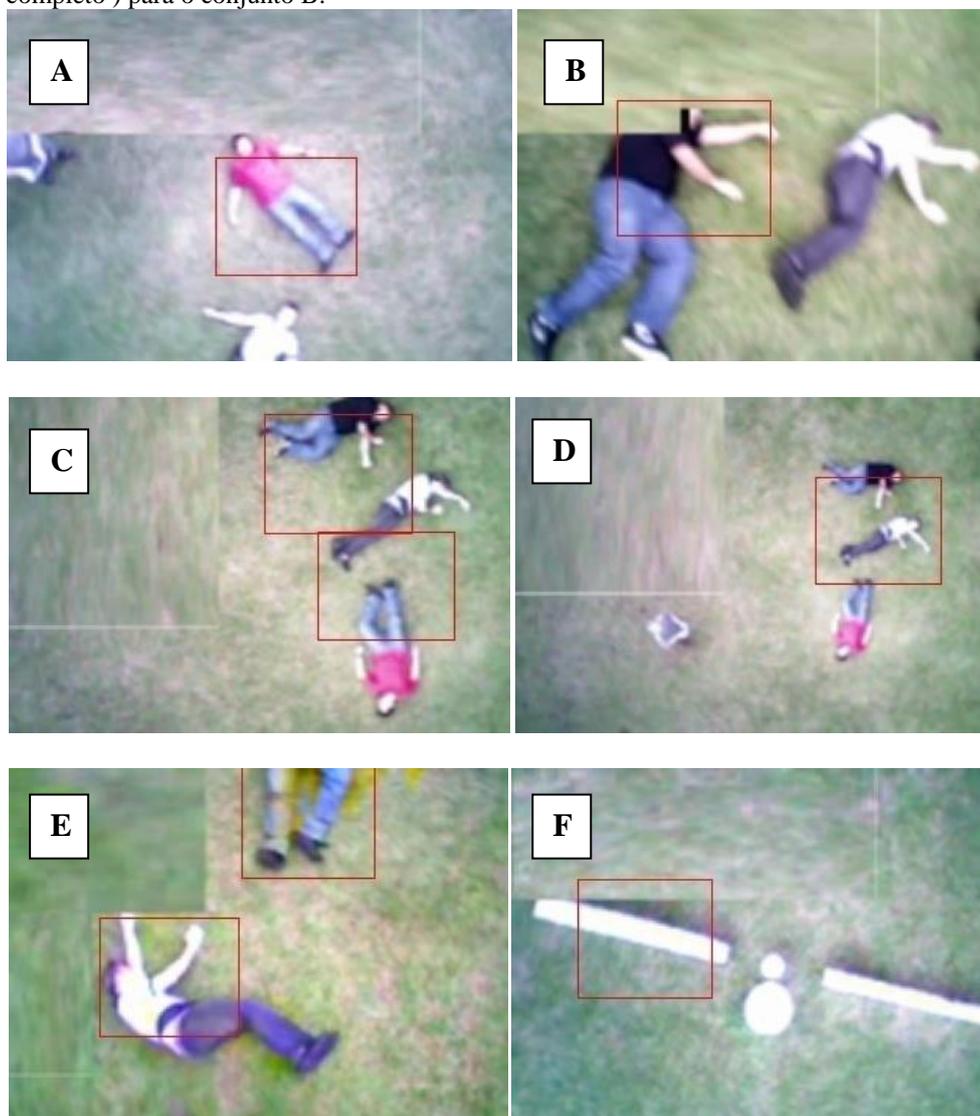
Fonte: Autor.

Figura 69 - Curva ROC do classificador em cascata (SVM-Adaboost) na condição $\Delta = 5$ e $\sigma = 0.8$ para o conjunto B.



Fonte: Autor.

Figura 70 - Saída do classificador SVM-Adaboost (Em vermelho detecção de corpo completo) para o conjunto B.



Fonte: Autor.

Figura 71 - Saída do classificador SVM-Adaboost (Em verde detecção de partes do corpo) para o conjunto B.



Fonte: Autor.

As figuras 70 e 71 apresentam a saída do classificador SVM-Adaboost, em que o quadrado em vermelho é a detecção de corpo completo, e em verde, a classificação de partes do corpo. As figura 70-A, B, C, D e E são classificações corretas de corpo completo, enquanto a figura 70-F é uma detecção incorreta. Nota-se que algumas imagens de corpos não foram detectadas, devido ao nível de contraste baixo e à criação de ruídos na tentativa de compensação. As figuras 71-A e B são saídas de detecções positivas de partes do corpo (Braços), e as figuras 71-C e D, de detecções incorretas. A figura 71-C é uma detecção incorreta de cabeça e a figura 71-D, de braço ou perna.

Os resultados obtidos com o classificador SVM-HIK são apresentados nas tabelas 25 e 26 e que representam os indicadores FPR, FNR, TPR e ACC, onde os valores V1, V2, V3, V4 e V5 são os pontos de ajuste de β nas curvas ROC, figuras 72 e 73. Observa-se que, para o conjunto B, esse classificador não é adequado, pois, devido ao elevado nível de (β) , para compensar a falta de contraste das imagens, foram gerados ruídos. Conseqüentemente, isso elevou a taxa de falsos positivos, ou seja, na condição ($\Delta = 3$ e $\sigma = 0.9$) com TPR = 0.757, tem-se FPR = 0.353, e na condição ($\Delta = 5$ e $\sigma = 0.8$) com TPR = 0.65, tem-se FPR = 0.41. O tempo de processamento apresentou o mesmo problema que a condição com o uso do SVM-Adaboost

para o conjunto (B), ou seja, o tempo de processamento também aumentou, observar tabela 27. O resultado da detecção de imagens classificadas é apresentado na figura 74:

Tabela 25 - Indicadores (SVM-HIK) para o conjunto (B) com $\Delta = 3$ e $\sigma = 0.9$.

	SVM HIK Conj. B ($\Delta = 3$ $\sigma = 0.9$)				
	V1	V2	V3	V4	V5
TPR	0,027	0,568	0,757	0,81	0,97
FPR	0	0,265	0,353	0,62	0,91
FNR	0,973	0,432	0,243	0,19	0,03
ACC	0,493	0,648	0,704	0,61	0,55

Fonte: Autor.

Tabela 26 - Indicadores (SVM-HIK) para o conjunto (B) com $\Delta = 5$ e $\sigma = 0.8$

	SVM HIK Conj. B ($\Delta = 5$ $\sigma = 0.8$)				
	V1	V2	V3	V4	V5
TPR	0,081	0,27	0,595	0,65	0,68
FPR	0	0,088	0,5	0,41	0,38
FNR	0,919	0,73	0,405	0,35	0,32
ACC	0,521	0,577	0,549	0,62	0,65

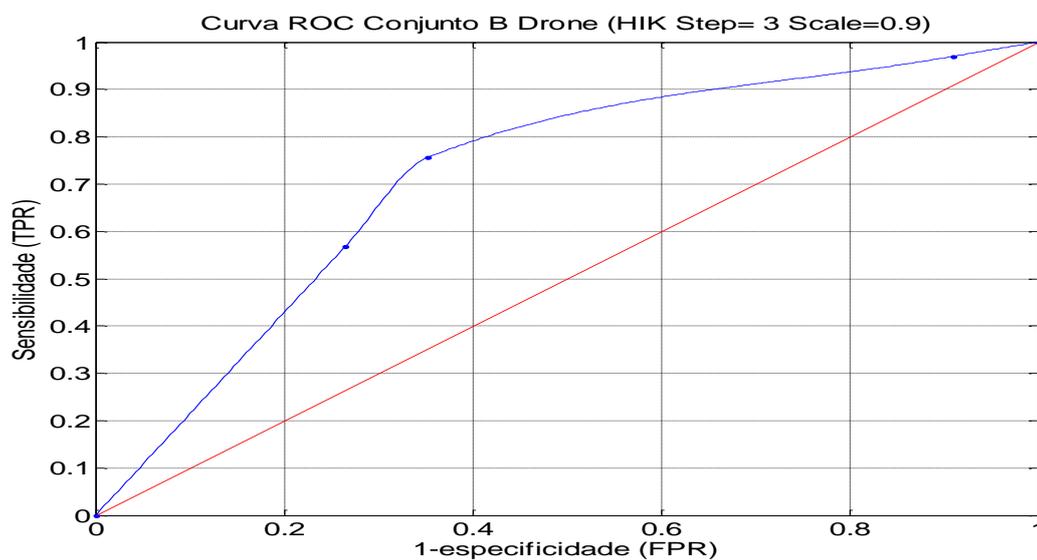
Fonte: Autor.

Tabela 27- Tempo de processamento para o conjunto (B) com (SVM-HIK)

$\Delta = 3$ e $\sigma = 0.9$.	1 imagens/seg.
$\Delta = 5$ e $\sigma = 0.8$.	2 imagens/seg.

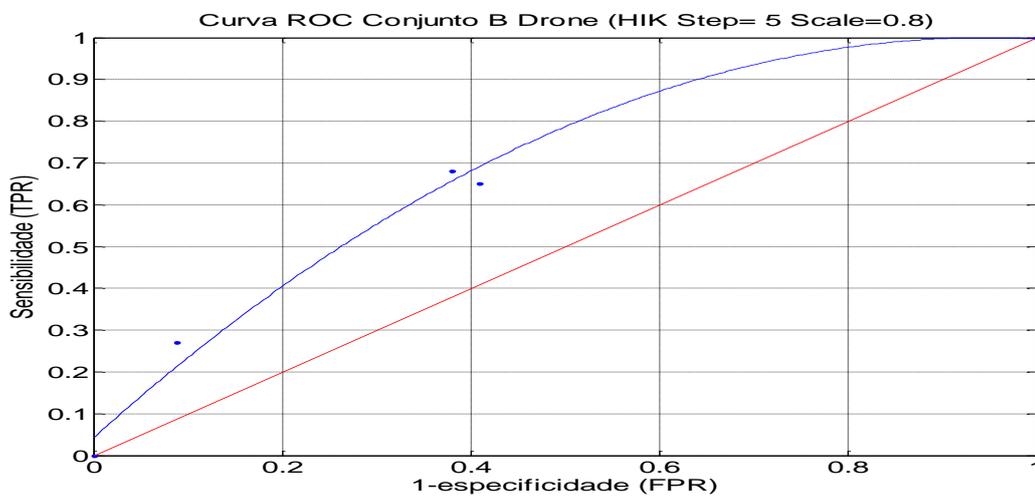
Fonte: Autor

Figura 72 -Curva ROC do classificador em cascata (SVM-HIK) na condição $\Delta = 3$ e $\sigma = 0.9$ para o conjunto B.



Fonte: Autor.

Figura 73 -Curva ROC do classificador em cascata (SVM-HIK) na condição $\Delta = 5$ e $\sigma = 0.8$ para o conjunto B.



Fonte: Autor.

Figura 74 - Saída do classificador SVM-HIK (Em vermelho detecção de corpo completo e em verde detecção de partes do corpo) para o conjunto B.



Fonte: Autor.

Na figura 74, é apresentada a saída da classificação, SVM-HIK, para o conjunto (B). O quadrado em vermelho é a detecção de corpo completo, e em verde, a classificação de partes do corpo. A figura 74-A e C são classificações corretas de corpo completo e partes do corpo, a

figura 74-B é uma detecção incorreta de corpo completo, devido ao ruído, e a figura 74-D, uma detecção incorreta de parte do corpo (cabeça).

5.5.3 Experimentos com o Conjunto de Imagens de Desastres (C)

O terceiro e último experimento com o conjunto (C) é formado por imagens reais do desastre ocorrido nos Alpes Franceses, da empresa aérea alemã Germanwings, e da destruição da aeronave MH17, da Malaysia Airlines, no leste da Ucrânia, extraídas da internet, com 29 imagens de plano de fundo e 20 imagens positivas de presença de corpo humano, porém nem todas as imagens são do ponto de vista aéreo. Nesse último experimento, o nível dos gradientes foi ajustado em $\beta = 0.00142$, pois, no caso desse conjunto de imagens, não há necessidade de reforço no contraste para melhorar a detecção. Os resultados são apresentados nas tabelas 28 e 29 e que representam os indicadores FPR, FNR, TPR e ACC, onde os valores V1, V2, V3, V4 e V5 são os pontos de ajuste de β e que corresponde às respectivas curvas ROC, figuras 75 e 76. Nesse, para o classificador SVM-Adaboost na condição ($\Delta = 3$ e $\sigma = 0.9$), o melhor desempenho é de TPR = 0.75, com FPR = 0.207, e na condição de maior velocidade de processamento ($\Delta = 5$ e $\sigma = 0.8$), foi obtido TPR = 0.85, com FPR = 0.28. Os tempos de processamento respectivos obtidos foram de (1.5 imagens/seg.) e (5 imagens/seg.), com elevado valor de TRF e baixo FPR, observar tabela 30. Aqui, os resultados são comparados aos obtidos no conjunto (A), com exatidão próxima de (ACC=0.8). Isso porque há uma boa condição de contraste nas imagens, sem a necessidade de reforço nos gradientes, garantindo, assim, tempos reduzidos no processamento. A figura 77 apresenta o resultado da detecção do classificador SVM-Adaboost.

Tabela 28 - Indicadores (SVM-Adaboost) para o conjunto (C) com $\Delta = 3$ e $\sigma = 0.9$.

AdaBoost Conj. C ($\Delta = 3$ $\sigma = 0.9$)					
	V1	V2	V3	V4	V5
TPR	0	0,3	0,75	0,8	0,9
FPR	0,103	0,17	0,207	0,45	0,586
FNR	1	0,7	0,25	0,2	0,1
ACC	0,531	0,61	0,776	0,65	0,612

Fonte: Autor.

Tabela 29 - Indicadores (SVM-Adaboost) para o conjunto (C) com $\Delta = 5$ e $\sigma = 0.8$.

AdaBoost Conj. C ($\Delta = 5$ $\sigma = 0.8$)					
	V1	V2	V3	V4	V5
TPR	0,05	0,3	0,55	0,85	0,95
FPR	0,069	0,1	0,172	0,28	0,655
FNR	0,95	0,7	0,45	0,15	0,05
ACC	0,571	0,65	0,714	0,78	0,592

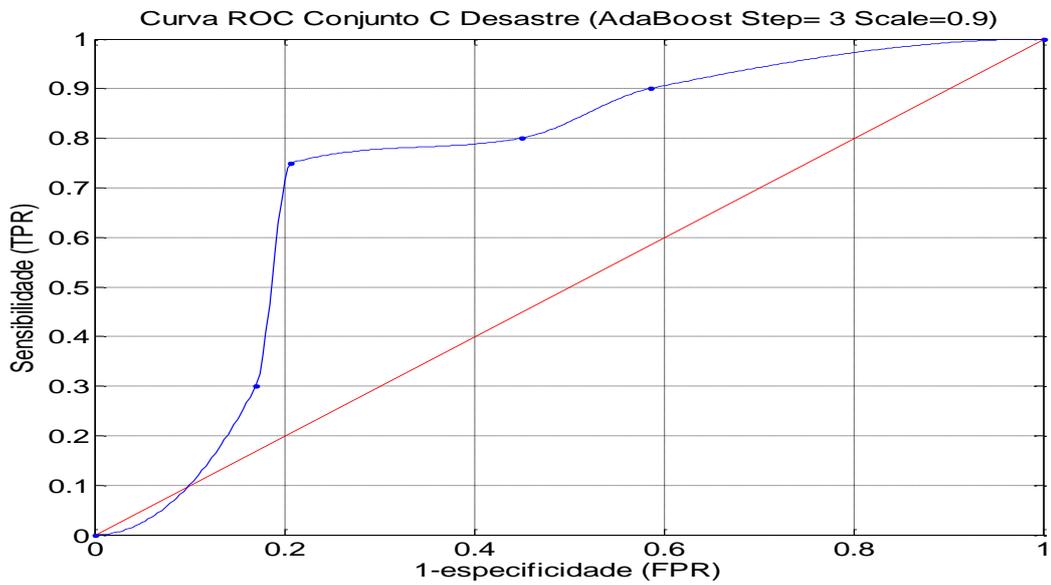
Fonte: Autor.

Tabela 30- Tempo de processamento para o conjunto (C) com (SVM-Adaboost)

$\Delta = 3$ e $\sigma = 0.9$.	1.5 imagens/seg.
$\Delta = 5$ e $\sigma = 0.8$.	5.0 imagens/seg.

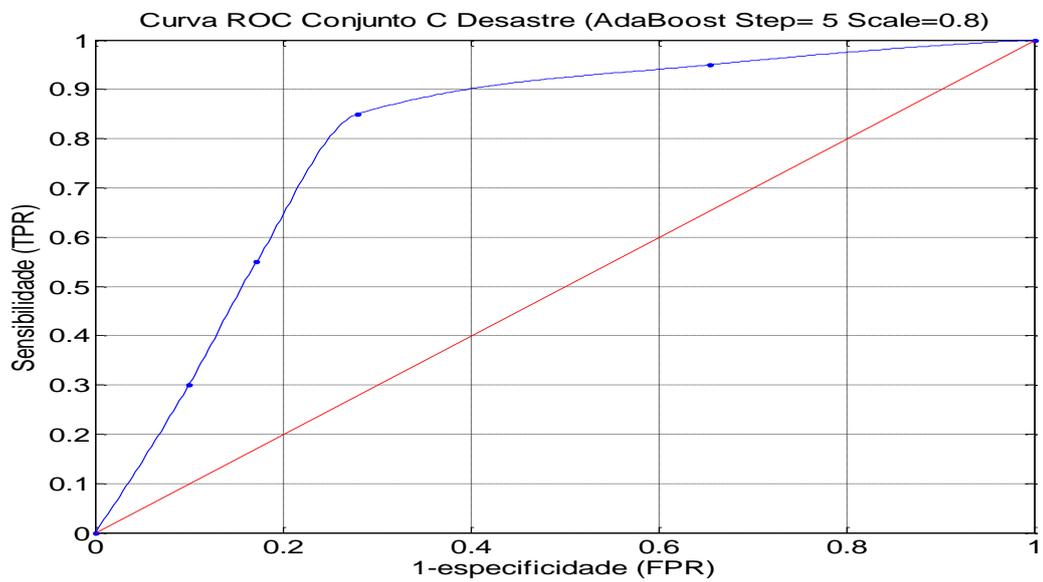
Fonte: Autor.

Figura 75 -Curva ROC do classificador em cascata (SVM-Adaboost) na condição $\Delta = 3$ e $\sigma = 0.9$ para o conjunto C.



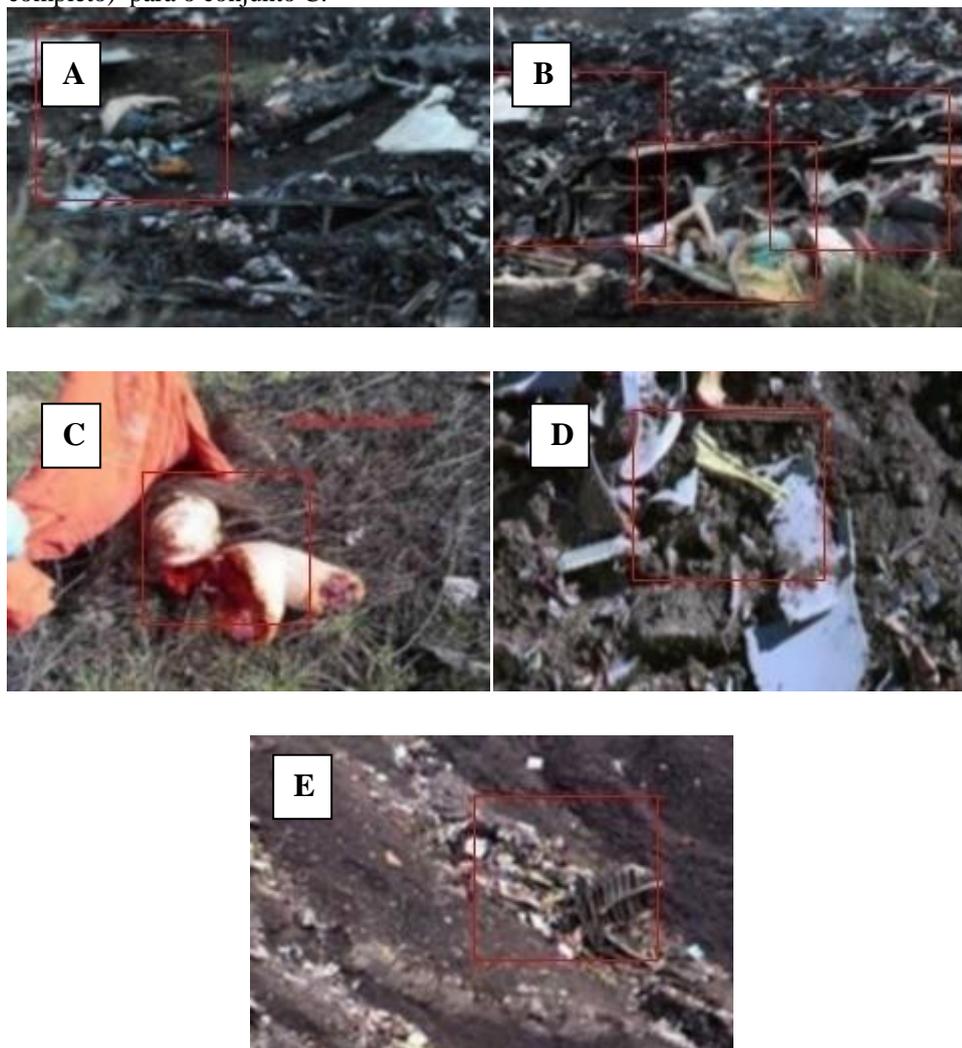
Fonte: Autor.

Figura 76-Curva ROC do classificador em cascata (SVM-Adaboost) na condição $\Delta = 5$ e $\sigma = 0.8$ para o conjunto C.



Fonte: Autor.

Figura 77 - Saída do classificador SVM-Adaboost (Em vermelho detecção de corpo completo) para o conjunto C.



Fonte: Autor.

A figura 77 apresenta a saída do classificador SVM-Adaboost para o conjunto (C), em que o quadrado em vermelho é a detecção de corpo completo. As figuras 77-A, B e C são classificações corretas, apesar de algumas não serem do ponto de vista aéreo. A figura 77-D é uma classificação incorreta, pois revela uma imagem feita à curta distância de um chinelo. A figura 77-E é também uma classificação incorreta, e neste caso, é uma imagem feita provavelmente à grande distância de destroços.

Os resultados obtidos com o classificador SVM-HIK são apresentados nas tabelas 31 e 32 e que representam os indicadores FPR, FNR, TPR e ACC, onde os valores V1, V2, V3, V4 e V5 são os pontos de ajuste de β das respectivas curvas ROC nas figuras 78 e 79. Nesse caso, o classificador SVM-HIK apresenta um resultado um pouco abaixo do obtido com o classificador SVM-AdaBoost; ou seja, na condição ($\Delta = 3$ e $\sigma = 0.9$), com TPR = 0.7 e FPR =

0.28, e na condição ($\Delta = 5$ e $\sigma = 0.8$), com $TPR = 0.75$ e $FPR = 0.31$; porém, com um tempo de processamento médio um pouco menor em ambas as condições de (2.5 imagens/seg.) e (6.25 imagens/seg.) respectivamente, observar tabela 33. A figura 80 apresenta o resultado da detecção do classificador SVM-HIK.

Tabela 31- Indicadores (SVM-HIK) para o conjunto (C) com $\Delta = 3$ e $\sigma = 0.9$.

	SVM HIK Conj. C ($\Delta = 3$ $\sigma = 0.9$)				
	V1	V2	V3	V4	V5
TPR	0	0,05	0,45	0,7	0,95
FPR	0,103	0,138	0,207	0,28	0,55
FNR	1	0,95	0,55	0,3	0,05
ACC	0,531	0,531	0,653	0,71	0,65

Fonte: Autor.

Tabela 32 - Indicadores (SVM-HIK) para o conjunto (C) com $\Delta = 5$ e $\sigma = 0.8$.

	SVM HIK Conj. C ($\Delta = 5$ $\sigma = 0.8$)				
	V1	V2	V3	V4	V5
TPR	0	0,25	0,55	0,75	0,95
FPR	0,138	0,172	0,241	0,31	0,93
FNR	1	0,75	0,45	0,25	0,05
ACC	0,51	0,592	0,673	0,71	0,43

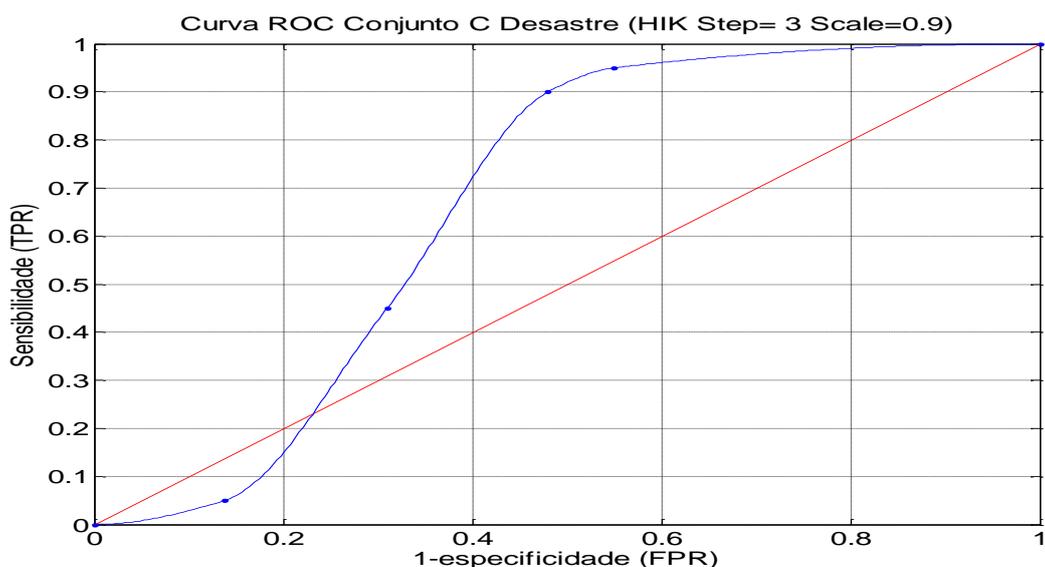
Fonte: Autor.

Tabela 33- Tempo de processamento para o conjunto (C) com (SVM-HIK)

$\Delta = 3$ e $\sigma = 0.9$.	2.5 imagens/seg.
$\Delta = 5$ e $\sigma = 0.8$.	6.25 imagens/seg.

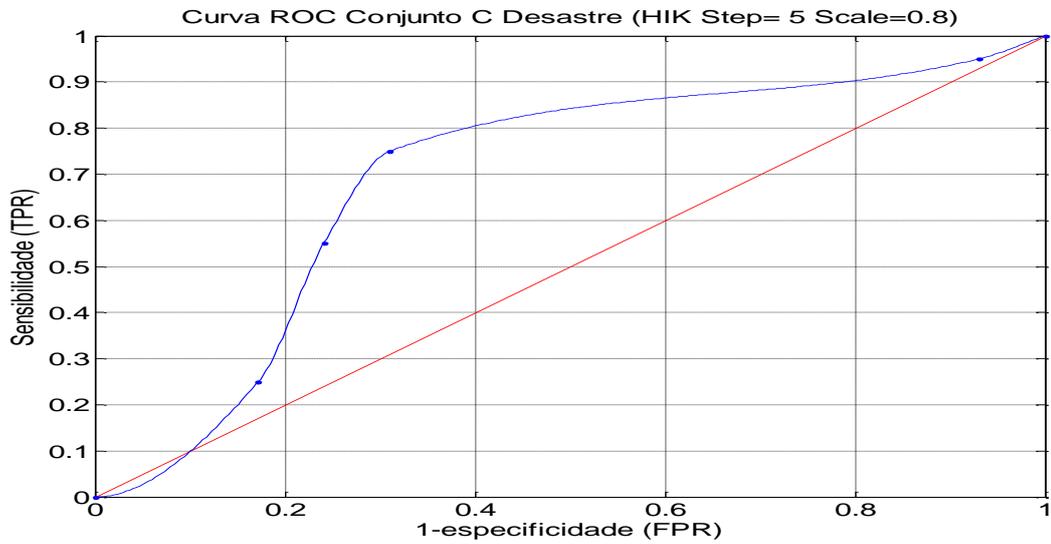
Fonte: Autor

Figura 78 -Curva ROC do classificador em cascata (SVM-HIK) na condição $\Delta = 3$ e $\sigma = 0.9$ para o conjunto C.



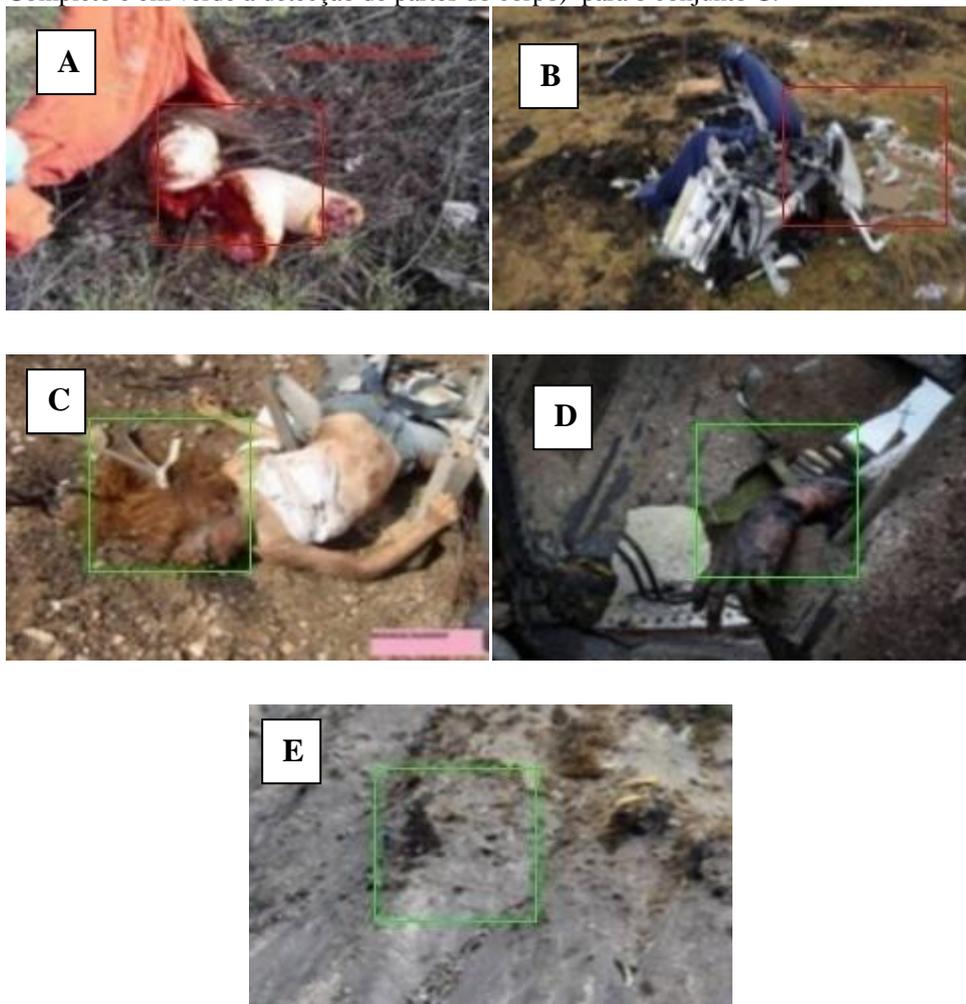
Fonte: Autor.

Figura 79 -Curva ROC do classificador em cascata (SVM-HIK) na condição $\Delta = 5$ e $\sigma = 0.8$ para o conjunto C



Fonte: Autor.

Figura 80 - Saída do classificador SVM-HIK (Em vermelho a detecção de corpo Completo e em verde a detecção de partes do corpo) para o conjunto C.



Fonte:Autor.

Na figura 80, a saída do classificador SVM-HIK para o conjunto C é demonstrada, observar que algumas imagens não são do ponto de vista aéreo, porém mesmo assim houve uma classificação. As figuras com quadrado em vermelho são a detecção de corpo completo, e em verde, a classificação de partes do corpo. A figura 80-A é um exemplo de classificação correta de corpo completo, enquanto a figura 80-B representa uma detecção incorreta. As figuras 80-C e D apresentam detecções corretas de partes do corpo (Cabeça e Braço), e a figura 80-E, uma detecção incorreta de imagem capturada, provavelmente, a uma grande distância.

5.5.4 Discussão dos Resultados

A apresentação final dos resultados está presente nas tabelas 34 e 35 onde o classificador em cascata SVM-AdaBoost na condição ($\Delta = 5$ e $\sigma = 0.8$) apresentou um desempenho melhor em relação ao SVM-HIK, em quase todos os conjuntos (A, B e C) analisados. Cabe notar também que a condição ($\Delta = 5$ e $\sigma = 0.8$), ou seja, de maior velocidade de processamento, apresentou a mais adequada solução para este trabalho, em termos gerais (TPR x ACC x Tempo de processamento), do que a condição ($\Delta = 3$ e $\sigma = 0.9$). O conjunto B (imagens feitas pelo drone) não apresentou um bom resultado, pelos motivos já expostos antes: baixa resolução de imagem e saturação de iluminação exterior.

Um comparativo entre trabalhos correlatos está presente na tabela 36, em que se verifica que não há um padrão comum de dados estatísticos a ser avaliado, visto que alguns avaliam em termos de acurácia (ACC), e outros, pelo *Equal Error Rate* (ERR). Outro ponto relevante é que não há também um banco de imagens disponível a ser usado de forma comum para a avaliação de desempenho, ou seja, cada trabalho cria seu próprio conjunto de imagens.

Tabela 34 - Comparação dos indicadores para os conjuntos (A, B e C) com $\Delta = 3$ e $\sigma = 0.9$.

$\Delta = 3$ e $\sigma = 0.9$	Conjunto A (Indoor)		Conjunto B (Drone)		Conjunto C (Desastre)	
	SVM-Adaboost	SVM-HIK	SVM-Adaboost	SVM-HIK	SVM-Adaboost	SVM-HIK
TPR	0.76	0.714	0.76	0.757	0.75	0.7
PPV	0.67	0.68	0.68	0.7	0.714	0.64
FPR	0.167	0.146	0.38	0.353	0.207	0.28
ACC	0.81	0.81	0.69	0.704	0.776	0.71

Fonte: Autor.

Tabela 35 - Comparação dos indicadores para os conjuntos (A, B e C) com $\Delta = 5$ e $\sigma = 0.8$.

$\Delta = 5$ e $\sigma = 0.8$	Conjunto A (Indoor)		Conjunto B (Drone)		Conjunto C (Desastre)	
	SVM-Adaboost	SVM-HIK	SVM-Adaboost	SVM-HIK	SVM-Adaboost	SVM-HIK
TPR	0.81	0.762	0.622	0.68	0.85	0.75
PPV	0.65	0.57	0.742	0.66	0.68	0.63
FPR	0.19	0.25	0.235	0.38	0.28	0.31
ACC	0.81	0.754	0.69	0.65	0.78	0.71

Fonte: Autor.

Tabela 36 - Comparação de trabalhos correlatos.

No	Trabalhos	Resultados
1	(Andriluka et al.,2010)	Equalerror rate (ERR) = 66%
2	(Kleiner et al.,2007)	Taxa de erro de 7.5%
3	(Castillo et al.,2005)	ACC= 78%
4	Cubber et al.,2009)	ACC = 65 %
5	(SONI;SOWMYA , 2012)	ACC = 69 %
6	Este Trabalho	Resultado do conjunto A(Indoor) ACC= 81% Resultado do conjunto B(Drone) ACC= 69% Resultado do conjunto C(Desastres Reais) ACC= 78%

Fonte: Autor.

5.6 CONSIDERAÇÕES FINAIS

Os experimentos conduzidos e relatados neste trabalho demonstram que é possível obter resultados favoráveis, com a associação de um algoritmo de corpo completo e comitês de especialistas na detecção de partes do corpo e utilizando-se apenas de um único atributo. Contudo, é evidente que há uma clara dependência do nível de contraste envolvido, já que isso afeta o desempenho na detecção e o tempo de processamento. Esse, por sua vez, tem seu desempenho afetado diretamente pelo número de janelas de detecções analisadas, principalmente pelo bloco de detecção de partes do corpo, que pode elevar o tempo em duas ou até cinco vezes, dependendo de quantas vezes esse bloco é requisitado pelo algoritmo principal, SVM-AdaBoost ou SVM-HIK. No próximo capítulo, conclui-se este trabalho, apresentando todos os pontos relevantes desenvolvidos.

6 CONCLUSÃO E TRABALHOS FUTUROS

Esse trabalho apresentou os resultados da associação de um algoritmo detector de corpo completo, comitês de especialistas na detecção de partes do corpo e rede bayesiana na tomada de decisão final da detecção de formas humanas, em cenas aéreas de desastres. Demonstrou-se que essa abordagem foi eficiente para o objetivo proposto, utilizando como atributo apenas o contorno da forma humana completa e parcial. Procurou-se demonstrar, pelos resultados experimentais, que alguns casos de oclusões parciais ou quase totais foram detectadas. Para essa comprovação, foram feitos três experimentos para aproximação das cenas reais, com diferentes situações em ambientes distintos, como dentro de uma residência com iluminação artificial, externo em campo aberto com iluminação natural e cenas reais de acidentes retiradas da internet. Por esses experimentos, evidenciou-se que o contraste entre as formas em uma cena é fundamental para perfeito reconhecimento e atuação do algoritmo como um todo. No entanto, esse mesmo contraste deve estar balanceado, para que não ocorra uma saturação de informação, levando a um aumento no tempo de processamento ou ao surgimento de ruídos, e elevando a taxa de falsos positivos (FPR). Uma observação importante a favor desse algoritmo (distinto de outros trabalhos) foi a classificação de imagens em diferentes altitudes, visto que as cenas analisadas podem levar ao erro de classificação, como no caso das cenas do quadrotor.

Concluiu-se ainda que, diferentemente do que se esperava para o classificador SVM-HIK, seu desempenho ficou abaixo do classificador SVM-Adaboost, quando analisada a taxa de verdadeiros positivos (TPR) versus a de falsos positivos (FPR), e não apenas o total de acertos com a exatidão; porém, a expectativa de tempo de processamento menor foi mantida, em razão do uso de um simples limiar de decisão. A razão para a queda no desempenho do classificador SVM-HIK se deve, no caso do conjunto (B), realizado em campo aberto na FEI, ao ajuste elevado no valor dos gradientes (β) para compensar a perda de contraste nas imagens. Outro fator relevante que pode ter levado à queda de desempenho é o fato de não ter sido utilizada a técnica de pirâmide espacial no algoritmo CENTRIST, já que impactava no aumento do tempo de processamento.

O bloco detector de partes do corpo humano também apresentou uma eficiência adequada a que se propôs, na alternativa de não detecção do corpo completo pelo algoritmo em cascata SVM-Adaboost ou do SVM-HIK; mas, a um custo de processamento de duas a cinco vezes o tempo utilizado pelo algoritmo detector de corpo completo. Em alguns casos de detecções incorretas pelo detector de partes, notou-se que erros ocorriam pela proximidade do

formato do braço e da perna, isto é, os classificadores se tornaram pouco discriminantes e não conseguiram uma distinção eficiente. O ideal seria a mudança de um dos atributos, perna ou braço, para, por exemplo, o tronco humano.

Ainda como contribuição do trabalho, a criação do banco de imagens mostrou-se adequada para o treinamento dos modelos dos classificadores de corpo completo e de detecção de partes. Todavia, como foi utilizado apenas um modelo de corpo humano, o ideal para a melhoria do desempenho seria a utilização de novas amostras e da variedade de posicionamentos.

Como trabalhos futuros, a implementação de classificadores multi-classes específicos (com algumas posições de cenas de desastres), operando em paralelo, seria uma alternativa para o classificador de corpo completo, que tenta “aprender” todas as posições possíveis, e acaba por falhar. O controle automático de contraste é outro avanço que poderia ocorrer ao nível físico, com o uso de sensores de intensidade luminosa, alinhados a uma avaliação da variância da intensidade dos pixels no quadro analisado. O tempo de processamento também, por ser otimizado pelo uso de sensores de altitude. Isso porque reduz o número de escalonamento dimensional das janelas de detecções, e o uso associativo da técnica de pirâmide espacial (que eleva o vetor de atributos, mas captura detalhes em uma imagem) com o algoritmo PCA, que diminui a dimensão desse mesmo vetor. Essas são algumas técnicas possíveis de ser aplicadas de forma prática.

REFERÊNCIAS

- ARIE, . M. et al. **Fast and stable human detection using multiple classifiers based on subtraction stereo with H.O.G. features.** In: IEEE INTERNATIONAL CONFERENCE ON ROBOTICS AND AUTOMATION, 2011, Shanghai, China, 2011, p. 868-873.
- BAGGIO, D. L. et al. **Mastering opencv with practical computer vision projects.** 1ª. ed. Birmingham: Packt Publishing, 2012.
- BARBER, D. **Bayesian reasoning and machine learning.** Cambridge: Cambridge University Press, 2010.
- BARLA, A.; ODOONE, F.; VERRI, A. **Histogram intersection kernel for image classification.** In: IMAGE PROCESSING, ICIP. PROCEEDINGS INTERNATIONAL CONFERENCE ON GENOVA, 2003, Genova, Italy, 2003 p. 513-516.
- BEN-GAL, I. **Bayesian Networks.** Israel: Tel Aviv University, Wiley e Sons Press, p. 1-6, 2007. Disponível em: <http://www.eng.tau.ac.il/~bengal/BN.pdf>. Acesso em: 04 dez. 2015.
- BHUMAN, S.; ARCOT, S. **Classifier ensemble with incremental learning for disaster victim detection.** In: PROCEEDINGS OF THE IEEE, 2012, Guangzhou, China, 2012, p. 1-6.
- BHUMAN, S.; ARCOT, S. **Victim Detection and Localization in an Urban Disaster Site.** In: PROCEEDING OF THE IEEE INTERNATIONAL CONFERENCE ON ROBOTICS AND BIOMIMETICS (ROBIO), Shenzhen, China, 2013, p. 2142-2147.
- BISHOP, C. M. **Pattern Recognition and Machine Learning.** Califórnia: Springer Press, 2006.
- BOUGHORBEL, S.; TAREL, J.-P.; BOUJEMAA, N. **Generalized histogram intersection kernel for image recognition.** In: INTERNATIONAL CONFERENCE ON IMAGE PROCESSING ICIP, 2005, Paris, France, 2005, p. 1-4.
- BRADSKI, G.; KAEHLER, A. **Learning OpenCV.** California: O'Reilly Press, 2008.
- BRAHMBHATT, S. **Practical OpenCv.** 1ª. ed. New York: Apress, 2013.
- BRAHNAM, S. et al. **Local Binary Patterns: New Variants and Applications.** 1ª. ed. Berlin: Springer-Verlag, 2014.
- CASTILLO, C.; CHANG, C. **A Method to Detect Victims in Search and Rescue Operations using Template Matching.** IEEE INTERNATIONAL WORKSHOP ON SAFETY, SECURITY AND RESCUE ROBOTICS, 2005, Kobe, Japan, 2005, p. 201-206.
- CAVALCANTI, G. D. D. C. **Curva ROC.** Universidade Federal De Pernambuco. Pernambuco, Brasil, p. 1-34. 2014.
- CHANG, C.-C. C.; CHIH-JEN, L. **LIBSVM: A library for support vector machines,** Taipei, Taiwan, 2013. p. 1-39.

CHENG, J.; GREINER, R. **Comparing bayesian network classifiers**, Alberta, Canada, 2013, p. 101-108.

CORVEE, E.; BREMOND, F. **Body parts detection for people tracking using trees of Histogram of Oriented Gradient descriptors**. In: SEVENTH IEEE INTERNATIONAL CONFERENCE ON ADVANCED VIDEO AND SIGNAL BASED SURVEILLANCE, 2010 Boston, 2010, p. 469-475.

CUBBER, G. D.; MARTON, G. **Human victim detection**. Department of mechanical engineering, Brussels, Belgium, 2015, p. 1-9.

D., M.; D., S. J.; C., T. C. **Machine learning, neural and statistical classification**. Cambridge, Cambridge, 1994.

DALAL, N.; TRIGGS, B. **Histograms of oriented gradients for human detection**. In: COMPUTER VISION AND PATTERN RECOGNITION CVPR, 2005, San Diego, 2005, p. 1-8.

DERPANIS, K. G. **Integral image-based representations**, Toronto, Canada, July 2007.

ENG, H.-L. et al. **A Bayesian framework for robust human detection and occlusion handling using human shape model**. In: IEEE INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION, Washington, 2004, p. 1-4.

FELZENSZWALB, P. F. et al. **Object detection with discriminatively trained part based models**. In: PATTERN ANALYSIS AND MACHINE INTELLIGENCE, IEEE, 2009, Chicago, 2009, p. 1627-1645.

FOROUGHI, H.; REZVANIAN, A.; PAZIRAEI, A. **Robust Fall Detection using Human Shape and Multi-class Support Vector Machine**. In: SIXTH INDIAN CONFERENCE ON COMPUTER VISION, GRAPHICS AND IMAGE PROCESSING, 2008, Bhubaneswar, India, 2008, p. 413-420.

FRIEDMAN, N.; GOLDSZMIDT, M. **Building classifiers using bayesian networks**. In: AAAI-96, 1996, Palo Alto, 1996, p. 1277-1284.

GRAUMAN, K.; DARRELL, T. **The pyramid match kernel: discriminative classification with sets of image features**. In: IEEE INTERNATIONAL CONFERENCE ON COMPUTER VISION, 2005, Beijing, China, 2005 p. 1-8.

GUALDI, G.; PRATI, A.; CUCCHIARA, R. **Covariance descriptors on moving regions for human detection in very complex outdoor scenes**. In: THIRD ACM/IEEE INTERNATIONAL CONFERENCE, 2009, p. 1-8.

GUO, H.; HSU, W. **A survey of algorithms for real-time bayesian network inference**. In: AAAI TECHNICAL REPORT WS-02-15, 2002, Kansas, 2002, p. 1-12.

HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. **The elements of statistical learning data mining, inference, and prediction**. 2 nd. ed. Santa Clara: Springer Press, 2008.

HAYKIN, S. **Neural networks and learning machines**. 3 rd. ed. Ontario, Canada: Pearson Education, 2009.

HENRLON, M. **Practical issues in constructing a bayes' belief network**, Pittsburgh, p. 132-139,2013.

HOU, Y.-L.; PANG, G. K. H. **Human detection in a challenging situation**.In: Image processing (ICIP), 2009 16th IEEE International, 2009, Cairo, Egypt, 2009, p. 2561-2564.

HSU, C.-W.; CHANG, C.-C.; LIN, C.-J. **Practical Guide to Support Vector Classification**, Taipei, Taiwan, p. 1-16, 2010.

JIA, H.-X.; ZHANG, Y.-J. **Fast human detection by boosting histograms of oriented gradients**. In:FOURTH INTERNATIONAL CONFERENCE ON IMAGE AND GRAPHICS, 2007, Beijing, China, 2007, p. 683-688.

KLEINER, A.; KUMMERLE, R. **Genetic MRF model optimization for real-time victim detection in Search and Rescue**. In:IEEE/RSJ INTERNATIONAL CONFERENCE ON INTELLIGENT ROBOTS AND SYSTEMS, 2007, San Diego, 2007, p. 3015-3030.

KUN-CHE, L.; DON-LIN, Y. **Image Processing and Image Mining using Decision Trees**. Journal of information science and engineering, Taiwan, China, p. 989-1003,2009.

LAGANIÈRE, R. **Opencv 2 computer vision application programming cookbook**. BIRMINGHAM - MUMBAI: Packt Publishing, 2011.

LAMPERT, C. H.; BLASCHKO, M. B.; HOFMANN, T. **Efficient Subwindow Search: A Branch and Bound Framework for Object Localization**. In:IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, 2008, Anchorage, 2008, p. 1-13.

LAW, M. **SVM support vector machines**. In: CSE 802 - PATTERN RECOGNITION AND ANALYSIS. Michigan: [s.n.]. p. 2-30,2005.

LAZEBNIK, S.; SCHMID, C.; PONCE, J. **Spatial pyramid matching**. INRIA, Alpes, France, p. 1-19,2009.

LI, M. et al. **Rapid and robust human detection and tracking based on omega-shape features**. In:IMAGE PROCESSING (ICIP), 16th IEEE INTERNATIONAL, 2009, Beijing, China, 2009, p. 2545-254.

LIN, Z. et al. **Hierarchical part-template matching for human detection and segmentation**. In: ICCV 2007. IEEE 11th INTERNATIONAL, 2007, Rio de Janeiro, Brasil, 2007, p. 1-8.

LIN, Z.; DAVIS, L. S. **Shape-based human detection and segmentation via hierarchical part-template matching**. In:IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, [S.l.], p. 604-618,2010.

MAJI, S.; BERG, A. C.; MALIK, J. **Classification using intersection kernel support vector machines is efficient**. In: IEEE COMPUTER VISION AND PATTERN RECOGNITION, ANCHORAGE, Alasca, 2008, p. 1-8,2008.

MARGOTTO, P. R. **Curva ROC como fazer e interpretar no SPSS**. Escola Superior de Ciências da Saúde. Brasília, Brasil, p. 1-10. 2010.

MUHAMMET, B. et al. **Usage of HoG (Histograms of Oriented Gradients) features for victim detection at disaster areas**. In: ELECTRICAL AND ELECTRONICS ENGINEERING (ELECO), INTERNATIONAL CONFERENCE 2013, Istanbul, Turkey, 2013, p. 535-538.

MYKHAYLO, A. et al. **Vision based victim detection from unmanned aerial vehicles**. IEEE/RSJ INTERNATIONAL CONFERENCE ON INTELLIGENT ROBOTS AND SYSTEMS, 2010, Taipei, Taiwan, 2010, p. 1740-1747.

MYRTUE, N. G. **Visual detection of humans in a disaster scenario**, Copenhagen, Denmark, p. 1-117, 2013.

NARAYANAN, V.; DESAI, R.; CHOUDHURY, S. **Bootstrapping Image Classification with Sample Evaluation**, Pittsburgh, p. 1-8, [s.d.].

NEUBECK, A.; GOOL, V. L. **Efficient non-maximum suppression**. In: Pattern Recognition, 2006. ICPR 2006. 18th INTERNATIONAL CONFERENCE, 2006, Hong Kong, China, 2006, p- 1-6.

NGUYEN, D. T.; LI, W.; OGUNBONA, P. **Human detection using local shape and non-redundant binary patterns**. In: CONF. CONTROL, AUTOMATION, ROBOTICS AND VISION, 2010, Singapore, Taiwan, 2010, p. 1145-1150.

NING, H. et al. **Research on human body detection methods based on the head features on the disaster scenes**. In: IEEE, 2010, China, 2010, p. 380-385.

NIXON, M. S.; AGUADO, A. S. **Feature extraction and image processing**. 2nd. ed. Oxford : Elsevier, 2008.

POURRET, O.; NAIM, P.; MARCOT, B. **Bayesian networks a practical guide to applications**. Chichester, England: John Wiley, 2008.

PRATI, R. C.; BATISTA, G. E. A. P. A.; MONARD, M. C. **Curvas ROC para avaliação de classificadores**, Brasil, p. 1-8, 2008.

RABUÑAL, J. R.; DORADO, J. **Artificial neural networks in real-Life applications**. [S.l.]: Idea Group Publishing, 2006.

RUSSELL, S.; NORVING, P. **Artificial intelligence a modern approach**. 3 rd. ed. [S.l.]: Pearson, 2014.

KUNCHEVA, L. **Combining pattern classifiers methods and algorithms**. ed. [S.l.] Wiley, 2004.

RUWEN, H. et al. **Heat mapping for improved victim detection**. In: Proceedings of the 2011 IEEE International Symposium on Safety, Security and Rescue Robotics, 2011, Kyoto, Japan, 2011, p. 116-121.

SAFAVIAN, S. R.; LANDGREBE, D. **A Survey of decision tree classifier methodology**. [S.l.]: IEEE, Indiana, P. 1-58, 1991.

SAID, Y.; ATRI, M.; TOURKI, R. **Human detection based on integral histograms of oriented gradients and SVM**. In: COMMUNICATIONS, COMPUTING AND CONTROL APPLICATIONS (CCCA), INTERNATIONAL CONFERENCE, Hammamet, 2011, Tunisia, 2011, p. 1-5.

SCHAPIRE, R. E. **Explaining adaBoost**, New Jersey, Princeton University Press, p. 1-16, 2013.

SMOLA, A.; S., V. V. N. **Introduction to machine learning**, 1. ed. Cambridge, United Kingdom: Cambridge, 2008.

SUDDERTH, E. B.; WAINWRIGHT, M. J.; WILLSKY, A. S. **Embedded trees: estimation of gaussian processes on graphs with cycles**. In: IEEE TRANSACTIONS ON SIGNAL PROCESSING, [S.l.], p. 3136-3150, 2004.

TEAM, O. **The opencv tutorials release 2.4.9.0**. [S.l.]: Intel, 2014.

VAPNIK, V. **Statistical learning theory**. New York: Wiley Press, 1998.

VIOLA, P.; JONES, M. **Rapid object detection using a boosted cascade of simple features**. In: ACCEPTED CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, 2001, Hawaii, 2001 p. 1-9.

WANG, L.; YUNG, N. H. C. **Three-dimensional model-based human detection in crowded scenes**. In: IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS, [S.l.], p. 691-703, 2012.

WANG, S.; ZABIR, S.; LEIBE, B. **Lying pose recognition for elderly fall detection**, Los Angeles, p. 1-8, 2011.

WANG, X.; HAN, X. T.; YAN, S. **An HOG-LBP human detector with partial occlusion handling**. In: COMPUTER VISION, 2009 IEEE 12th INTERNATIONAL CONFERENCE, 2009, Kyoto, Japan, 2009, p. 32 - 39.

WIEGERINCK, W.; KAPPEN, B.; BURGERS, W. **Bayesian networks for expert systems, theory and practical Applications**. In: INTERACTIVE COLLABORATIVE INFORMATION SYSTEM. [S.l.]: Springer Berlin Heidelberg, v. 281, p. 547-578, 2010.

WU, B.; NEVATIA, R. **Detection and tracking of multiple, partially occluded humans by bayesian combination of edgelet based part detectors**. In: INTERNATIONAL JOURNAL OF COMPUTER VISION, 2006, Los Angeles, 2006, p. 1-20.

WU, J. **A simple guide to libHIK**. Nanyang Technological University. Singapura, China, p. p. 1-9, 2010.

WU, J. **Histogram intersection kernel learning for multimedia applications**. In: The 2010 PACIFIC-RIM CONFERENCE ON MULTIMEDIA, 2010, Shanghai, China, 2010, p. 1-150.

WU, J.; GEYER, C.; REHG, J. M. **Real-time human detection using contour cues**. In:INTERNATIONAL CONFERENCE ON ROBOTICS AND AUTOMATION, Shanghai, China, 2011, p. 860-867,2011.

WU, J.; REHG, J. M. **CENTRIST: A visual descriptor for scene categorization**. In:PATTERN ANALYSIS AND MACHINE INTELLIGENCE, IEEE, Singapore, 2010, Taiwan, 2010,p. 1-14.

XU, Z.; CHEN, M.; WEINBERGER, K. Q. **Classifier cascade for minimizing feature evaluation cost**. In:FIFTEENTH INTERNATIONAL CONFERENCE ON ARTIFICIAL INTELLIGENCE AND STATISTICS (AISTATS-12), 2012, La Palma, Canary Island, 2012, p. 218-226.

YANG, M. et al. **Human action detection by boosting efficient motion features**. In:IEEE 12th INTERNATIONAL CONFERENCE ON COMPUTER VISION WORKSHOPS, ICCV, [S.l.], p. 522-529, 2009.

ZHANG, S.; BAUCKHAG, C.; CREMER, A. B. **Informed Haar-like Features Improve Pedestrian Detection**. In:CVPR, [S.l.], p. 1-8,2014.

ZHIHUI, L.; CHUNYAN, S.; DI, S. **Real-time Human Detection Based on Cascade Frame**. In:INTERNATIONAL CONFERENCE ON MECHATRONICS AND AUTOMATION, 2010, Beijing, China, 2010, p. 514-518.

B.,L. Singley. **Cena de desatre, 1900**. Disponível em:<http://en.wikipedia.org/wiki/1900_Galveston_hurricane>, acesso em agosto de 2013.

Fawcet, Tom; Powers David. **Confusion Matrix**. Disponível em:<http://en.wikipedia.org/wiki/Confusion_matrix>, acesso em Junho 2013.

Pedro, Silva. **Pedestrian Detection-My master's thesis**. Aveiro, Portugal, 2013. Disponível em:<<http://peddetect.blogspot.com.br/2013/04/randomness-generating-large-pool-of.html>>, acessado em 04/04/2015.

Students. Project 5: **ObjectDetection**, 2012. Disponível em:<<http://www.cs.cornell.edu/courses/cs4670/2012fa/projects/p5/index.html>>, acessado em 06/04/2015.

Intel Group. **CascadeClassification**, 2015. Disponível em:<http://docs.opencv.org/modules/objdetect/doc/cascade_classification.html?highlight=grouprectangles>, acessado em 15/04/2015.

Howard, Hamilton. **Confusion Matrix**, University of Regina 2012. Disponível em: <http://www2.cs.uregina.ca/~dbd/cs831/notes/confusion_matrix/confusion_matrix.html>, acessado em 20/05/2015.

Ball, Thorsten. **Train your own opencv Haar classifier**. Disponível em: <<http://coding-robin.de/2013/07/22/train-your-own-opencv-haar-classifier.html>>, acessado em 20/05/2015.

Breckon, Toby. **Comuter Vision and Image Processing**. Disponível em: <http://breckon.eu/toby/teaching/ml/examples/c++/opticaldigits_ex/boosttree.cpp>, acessado em 09/06/2015.

Ludwing, O., et al. **Video Content Description**. Disponível em: <<http://www.upb.ro/en/>>, acessado em 05/08/2015.

Li, Jia. **Data Mining**. Disponível em: <<http://www.stat.psu.edu/~jjali/course/stat557>>, acessado em 12/07/2015.

Buch, Peter. **Maximum-margin hyperplane and margins for an SVM trained**. Disponível em: <https://en.wikipedia.org/wiki/Support_vector_machine#/media/File:Svm_max_sep_hyperplane_with_margin.png>, acessado em 22/07/2015.

Google Images. **Cenas de desastres da Germanwings, 2015**. Disponível em: <<https://www.google.com.br/search?q=acidente+a%C3%A9reo+da+germanwings&biw=1440&bih=799&source=lnms&tbn=isch&sa=X&ved=0CAcQAUoAmoVChMIxfSN4dbOyAIViX6QCh3-oA2L>>, acesso em agosto de 2015.

Google Imagens. **Cenas de desastres da Malaysiaairlines, 2015**. Disponível em: <<https://www.google.com.br/search?q=acidente+a%C3%A9reo+da+malaysia+airlines&biw=1440&bih=799&source=lnms&tbn=isch&sa=X&ved=0CAGQAUoA2oVChMIz9WOkdfOyAIVwSOQCh3BBafE>>, acesso em agosto de 2015.

Barnum, Pete. **Passos para a computação do HOG, 2006**, disponível em: <<http://pt.slideshare.net/zukun/histogram-of-oriented-gradients-for-human-detection>>, acesso em dezembro de 2015.

APÊNDICE A – ESTRATÉGIA DE CONCORDÂNCIA ENTRE CLASSIFICADORES

Em sistemas com múltiplos classificadores podem-se combinar descritores para melhorar o desempenho de classificadores, tão quão, combinar estes próprios classificadores. Para tanto três abordagens estão presentes: fusão, seleção e híbridos.

- fusão: assume que cada classificador foi treinado sobre todo o espaço de atributos. A classificação, portanto, resulta da combinação de todos os classificadores;
- seleção: assume que cada classificador é um especialista em um atributo. Ao detectar que um determinado atributo caiu no espaço de classificação o especialista nesta característica atuará;
- híbridos: há uma ponderação entre os classificadores especialistas com base na proximidade do atributo com sua região de domínio.

Neste trabalho foi estudada duas estratégias de concordância entre classificadores a abordagem por fusão (acima descrita) por voto majoritario e o índice estatístico de kappa.

No caso de fusão por voto majoritário um vetor $[d_{i1}, d_{i2}, \dots, d_{ic}]^t$ onde $d_{ij} \in \{0,1\}$ indica a saída do classificador D_i , $i = 1, 2, \dots, L$, com relação á amostra x pertencer ou não à classe C_j , $j = 1, 2, \dots, c$. O voto majoritário escolhe a classe C_k onde:

$$\sum_{i=1}^L D_{i,k} = \max_{j=1}^c \sum_{i=1}^L d_{i,j} , \quad (71)$$

Fusão por voto majoritário é um método bastante utilizado, pois, sob determinadas condições, ele realmente aumenta a probabilidade de acerto na classificação onde a escolha pode ser por unanimidade, maioria simples (50% + 1) ou mais votado.

O índice de Kappa avalia o grau de concordância quando existem dois classificadores com uma única amostra ou um classificador com duas amostras. As interpretações dos valores do Kappa estão compreendidas entre $[0,1]$ onde:

- Kappa = 1, implica uma concordância perfeita.
- Kappa = 0, a concordância é um mero acaso.
- Kappa = < 0, a concordância é mais fraca que o esperado pelo acaso.

APÊNDICE B – ERROS DE CLASSIFICAÇÃO DOS CONJUNTOS (A), (B) E (C) PARA AS CASCATAS SVM-ADABOOST E SVM-HIK.

Neste apêndice são exibidos todos os erros de classificação obtidos da matriz de confusão referentes aos conjuntos de imagens A, B e C usados nas cascatas de classificadores SVM-Adaboost e SVM-HIK:

Tabela 37- Erros de classif. (SVM-Adaboost) para o conjunto (A) com $\Delta = 3$ e $\sigma = 0.9$.

AdaBoost Conj. A ($\Delta = 3$ $\sigma = 0.9$)					
	V1	V2	V3	V4	V5
FN	21	13	5	4	1
FP	1	2	8	11	21
TP	0	8	16	17	20
TN	47	46	40	37	27
TPR	0	0,38	0,762	0,81	0,952
SPC	0,979	0,96	0,833	0,77	0,563
PPV	0	0,8	0,667	0,61	0,488
NPV	0,691	0,78	0,889	0,9	0,964
FPR	0,021	0,04	0,167	0,23	0,438
FNR	1	0,62	0,238	0,19	0,048
ACC	0,681	0,78	0,812	0,78	0,681
FDR	1	0,2	0,333	0,39	0,512
FOR	0,309	0,22	0,111	0,1	0,036

Fonte: Autor.

Tabela 38 – Erros de classif. (SVM-Adaboost) para o conjunto (A) com $\Delta = 5$ e $\sigma = 0.8$.

AdaBoost Conj. A ($\Delta = 5$ $\sigma = 0.8$)					
	V1	V2	V3	V4	V5
FN	21	17	10	4	2
FP	1	3	3	9	15
TP	0	4	11	17	19
TN	47	45	45	39	33
TPR	0	0,19	0,524	0,81	0,905
SPC	0,979	0,94	0,938	0,81	0,688
PPV	0	0,57	0,786	0,65	0,559
NPV	0,691	0,73	0,818	0,91	0,943
FPR	0,021	0,06	0,063	0,19	0,313
FNR	1	0,81	0,476	0,19	0,095
ACC	0,681	0,71	0,812	0,81	0,754
FDR	1	0,43	0,214	0,35	0,441
FOR	0,309	0,27	0,182	0,09	0,057

Fonte: Autor.

Tabela 39 – Erros de classif. (SVM-HIK) para o conjunto (A) com $\Delta = 3$ e $\sigma = 0.9$.

SVM HIK Conj. A ($\Delta = 3$ $\sigma = 0.9$)					
	V1	V2	V3	V4	V5
FN	21	9	6	3	1
FP	2	6	7	25	33
TP	0	12	15	18	20
TN	46	42	41	23	15
TPR	0	0,571	0,714	0,86	0,95
SPC	0,958	0,875	0,854	0,48	0,31
PPV	0	0,667	0,682	0,42	0,38
NPV	0,687	0,824	0,872	0,88	0,94
FPR	0,042	0,125	0,146	0,52	0,69
FNR	1	0,429	0,286	0,14	0,05
ACC	0,667	0,783	0,812	0,59	0,51
FDR	1	0,333	0,318	0,58	0,62
FOR	0,313	0,176	0,128	0,12	0,06

Fonte: Autor.

Tabela 40 – Erros de classif. (SVM-HIK) para o conjunto (A) com $\Delta = 5$ e $\sigma = 0.8$.

SVM HIK Conj. A ($\Delta = 5$ $\sigma = 0.8$)					
	V1	V2	V3	V4	V5
FN	21	7	5	4	3
FP	1	7	12	18	34
TP	0	14	16	17	18
TN	47	41	36	30	14
TPR	0	0,667	0,762	0,81	0,86
SPC	0,979	0,854	0,75	0,63	0,29
PPV	0	0,667	0,571	0,49	0,35
NPV	0,691	0,854	0,878	0,88	0,82
FPR	0,021	0,146	0,25	0,38	0,71
FNR	1	0,333	0,238	0,19	0,14
ACC	0,681	0,797	0,754	0,68	0,46
FDR	1	0,333	0,429	0,51	0,65
FOR	0,309	0,146	0,122	0,12	0,18

Fonte: Autor.

Tabela 41- Erros de classif. (SVM-Adaboost) para o conjunto (B) com $\Delta = 3$ e $\sigma = 0.9$.

AdaBoost Conj. B ($\Delta = 3$ $\sigma = 0.9$)					
	V1	V2	V3	V4	V5
FN	37	24	14	9	5
FP	1	1	8	13	14
TP	0	13	23	28	32
TN	33	33	26	21	20
TPR	0	0,35	0,622	0,76	0,865
SPC	0,971	0,97	0,765	0,62	0,588
PPV	0	0,93	0,742	0,68	0,696
NPV	0,471	0,58	0,65	0,7	0,8
FPR	0,029	0,03	0,235	0,38	0,412
FNR	1	0,65	0,378	0,24	0,135
ACC	0,465	0,65	0,69	0,69	0,732
FDR	1	0,07	0,258	0,32	0,304
FOR	0,529	0,42	0,35	0,3	0,2

Fonte: Autor.

Tabela 42 - Erros de classif. (SVM-Adaboost) para o conjunto (B) com $\Delta = 5$ e $\sigma = 0.8$.

AdaBoost Conj. B ($\Delta = 5$ $\sigma = 0.8$)					
	V1	V2	V3	V4	V5
FN	37	18	14	7	2
FP	1	7	8	27	24
TP	0	19	23	30	35
TN	33	27	26	7	10
TPR	0	0,51	0,622	0,81	0,946
SPC	0,971	0,79	0,765	0,21	0,294
PPV	0	0,73	0,742	0,53	0,593
NPV	0,471	0,6	0,65	0,5	0,833
FPR	0,029	0,21	0,235	0,79	0,706
FNR	1	0,49	0,378	0,19	0,054
ACC	0,465	0,65	0,69	0,52	0,634
FDR	1	0,27	0,258	0,47	0,407
FOR	0,529	0,4	0,35	0,5	0,167

Fonte: Autor.

Tabela 43 -Erros de classif. (SVM-HIK) para o conjunto (B) com $\Delta = 3$ e $\sigma = 0.9$.

SVM HIK Conj. B ($\Delta = 3$ $\sigma = 0.9$)					
	V1	V2	V3	V4	V5
FN	36	16	9	7	1
FP	0	9	12	21	31
TP	1	21	28	30	36
TN	34	25	22	13	3
TPR	0,027	0,568	0,757	0,81	0,97
SPC	1	0,735	0,647	0,38	0,09
PPV	1	0,7	0,7	0,59	0,54
NPV	0,486	0,61	0,71	0,65	0,75
FPR	0	0,265	0,353	0,62	0,91
FNR	0,973	0,432	0,243	0,19	0,03
ACC	0,493	0,648	0,704	0,61	0,55
FDR	0	0,3	0,3	0,41	0,46
FOR	0,514	0,39	0,29	0,35	0,25

Fonte: Autor.

Tabela 44 - Erros de classif. (SVM-HIK) para o conjunto (B) com $\Delta = 5$ e $\sigma = 0.8$.

SVM HIK Conj. B ($\Delta = 5$ $\sigma = 0.8$)					
	V1	V2	V3	V4	V5
FN	34	27	15	13	12
FP	0	3	17	14	13
TP	3	10	22	24	25
TN	34	31	17	20	21
TPR	0,081	0,27	0,595	0,65	0,68
SPC	1	0,912	0,5	0,59	0,62
PPV	1	0,769	0,564	0,63	0,66
NPV	0,5	0,534	0,531	0,61	0,64
FPR	0	0,088	0,5	0,41	0,38
FNR	0,919	0,73	0,405	0,35	0,32
ACC	0,521	0,577	0,549	0,62	0,65
FDR	0	0,231	0,436	0,37	0,34
FOR	0,5	0,466	0,469	0,39	0,36

Fonte: Autor.

Tabela 45 - Erros de classif. (SVM-Adaboost) para o conjunto (C) com $\Delta = 3$ e $\sigma = 0.9$.

AdaBoost Conj. C ($\Delta = 3$ $\sigma = 0.9$)					
	V1	V2	V3	V4	V5
FN	20	14	5	4	2
FP	3	5	6	13	17
TP	0	6	15	16	18
TN	26	24	23	16	12
TPR	0	0,3	0,75	0,8	0,9
SPC	0,897	0,83	0,793	0,55	0,414
PPV	0	0,55	0,714	0,55	0,514
NPV	0,565	0,63	0,821	0,8	0,857
FPR	0,103	0,17	0,207	0,45	0,586
FNR	1	0,7	0,25	0,2	0,1
ACC	0,531	0,61	0,776	0,65	0,612
FDR	1	0,45	0,286	0,45	0,486
FOR	0,435	0,37	0,179	0,2	0,143

Fonte: Autor.

Tabela 46 - Erros de classif. (SVM-Adaboost) para o conjunto (C) com $\Delta = 5$ e $\sigma = 0.8$.

AdaBoost Conj. C ($\Delta = 5$ $\sigma = 0.8$)					
	V1	V2	V3	V4	V5
FN	19	14	9	3	1
FP	2	3	5	8	19
TP	1	6	11	17	19
TN	27	26	24	21	10
TPR	0,05	0,3	0,55	0,85	0,95
SPC	0,931	0,9	0,828	0,72	0,345
PPV	0,333	0,67	0,688	0,68	0,5
NPV	0,587	0,65	0,727	0,88	0,909
FPR	0,069	0,1	0,172	0,28	0,655
FNR	0,95	0,7	0,45	0,15	0,05
ACC	0,571	0,65	0,714	0,78	0,592
FDR	0,667	0,33	0,313	0,32	0,5
FOR	0,413	0,35	0,273	0,13	0,091

Fonte: Autor.

Tabela 47- Erros de classif. (SVM-HIK) para o conjunto (C) com $\Delta = 3$ e $\sigma = 0.9$.

SVM HIK Conj. C ($\Delta = 3$ $\sigma = 0.9$)					
	V1	V2	V3	V4	V5
FN	20	19	11	6	1
FP	3	4	6	8	16
TP	0	1	9	14	19
TN	26	25	23	21	13
TPR	0	0,05	0,45	0,7	0,95
SPC	0,897	0,862	0,793	0,72	0,45
PPV	0	0,2	0,6	0,64	0,54
NPV	0,565	0,568	0,676	0,78	0,93
FPR	0,103	0,138	0,207	0,28	0,55
FNR	1	0,95	0,55	0,3	0,05
ACC	0,531	0,531	0,653	0,71	0,65
FDR	1	0,8	0,4	0,36	0,46
FOR	0,435	0,432	0,324	0,22	0,07

Fonte: Autor.

Tabela 48 - Erros de classif. (SVM-HIK) para o conjunto (C) com $\Delta = 5$ e $\sigma = 0.8$.

SVM HIK Conj. C ($\Delta = 5$ $\sigma = 0.8$)					
	V1	V2	V3	V4	V5
FN	20	15	9	5	1
FP	4	5	7	9	27
TP	0	5	11	15	19
TN	25	24	22	20	2
TPR	0	0,25	0,55	0,75	0,95
SPC	0,862	0,828	0,759	0,69	0,07
PPV	0	0,5	0,611	0,63	0,41
NPV	0,556	0,615	0,71	0,8	0,67
FPR	0,138	0,172	0,241	0,31	0,93
FNR	1	0,75	0,45	0,25	0,05
ACC	0,51	0,592	0,673	0,71	0,43
FDR	1	0,5	0,389	0,38	0,59
FOR	0,444	0,385	0,29	0,2	0,33

Fonte: Autor.