

COMBATE AO ABIGEATO A PARTIR DE DETECÇÃO DE PESSOAS EM IMAGENS AÉREAS OBTIDAS POR VANT

Ânderson Fiscoeder Soares¹

Leonardo Bidese De Pinho²

Resumo:

Detectar pessoas em imagens digitais é uma tarefa recorrente e importante para o desenvolvimento de várias aplicações. O presente trabalho tem o objetivo de propor um método para identificação de pessoas em imagens aéreas com a utilização de VANT para dar suporte ao produtor e aos órgãos especializados no combate ao abigeato. O treinamento dos descritores em cascata foi efetuado com a utilização de um software que apresenta os parâmetros implementados na biblioteca OpenCV em uma interface gráfica, facilitando assim a manipulação dos mesmos. Para fins de teste foram criados classificadores Haar e LBP com 8, 15 e 20 estágios, todos fazendo uso dos mesmos conjuntos de imagens positivas e negativas. Dados os resultados o classificador que obteve melhor taxa de acerto foi o descritor Haar de 15 estágios. Com os testes realizados pôde-se notar na prática que efetivamente classificadores Haar demandam mais tempo para treinamento, porém a taxa de acerto dos mesmos é melhor. Também observou-se que a quantidade de estágios de uma cascata deve ser muito bem dimensionada, já que classificadores com 20 estágios obtiveram resultados muito insatisfatórios.

Palavras-chave: AgroTIC, VANT, Visão computacional.

Modalidade de Participação: Iniciação Científica

COMBATE AO ABIGEATO A PARTIR DE DETECÇÃO DE PESSOAS EM IMAGENS AÉREAS OBTIDAS POR VANT

¹ Aluno de graduação. anderson.fiscoeder@gmail.com. Autor principal

² Docente. leonardopinho@unipampa.edu.br. Orientador

Combate ao Abigeato a partir de detecção de pessoas em imagens aéreas obtidas por VANT

1 INTRODUÇÃO

Segundo Ballard e Brown (1982), “Visão Computacional” é a maneira como a máquina enxerga o meio a sua volta, através da extração de informações a partir de imagens, sensores e outros dispositivos para que possa reconhecer, manipular e pensar sobre objetos presentes nessas imagens. A maior parte dos problemas de Visão Computacional envolve detecção e reconhecimento de objetos em imagens com a intenção de se obter informações úteis a fim de poder aprimorar o conhecimento ou gerar uma solução para um problema específico. A detecção de características, também chamadas de *features*, em imagens consiste na extração de propriedades que sejam relevantes para a problemática em questão, sendo estas bordas, cantos, gradientes ou padrões, podendo ser representadas por descritores. A classificação das características obtidas pelos descritores para uma máquina é algo significativamente complexa, visto que exige uma expressiva quantidade de informações disponíveis. A classificação pode ser feita de maneira supervisionada ou não supervisionada, sendo que a primeira utiliza uma base de dados com os padrões a serem reconhecidos previamente cadastrados, enquanto que a segunda assume que a máquina é capaz de identificar autonomamente as classes (CRÓSTA, 1992).

No contexto de agricultura e pecuária, nota-se a crescente utilização de Veículos Aéreos não Tripulados (VANT), sobretudo para fins de aerolevantamentos, os quais buscam auxiliar no monitoramento das plantações e rastreamento de rebanhos (MAKAROV et al., 2015). Nos haras da região sul do Rio Grande do Sul, especificamente, na BR-153, rodovia que liga os municípios de Bagé e Aceguá, ocorre o monitoramento dos animais através de rondas periódicas dos funcionários, visto que os cavalos Puro Sangue Inglês (PSI) podem ser avaliados em milhões e uma prática de abigeato – furto de animais de carga e de abate de propriedades rurais – em um desses locais acarretaria em um prejuízo significativo para o proprietário. No Rio Grande do Sul, embora sejam vários os relatos de ocorrências, poucos casos tiveram indiciamentos de envolvidos.

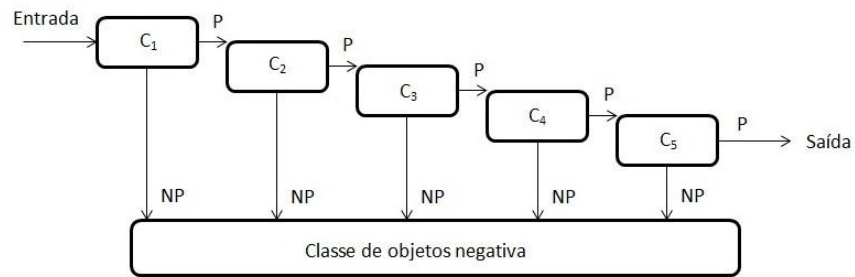
A partir deste cenário, percebe-se a relevância de um estudo que busque propor e avaliar métodos eficazes de tratamento de imagens aéreas para identificação de pessoas em uma propriedade rural, com uma coleta semiautônoma de imagens por meio de câmeras acopladas a VANT. Neste sentido, o presente trabalho objetiva: identificar o estado-da-arte sobre VANT e o seu uso em propriedades rurais; selecionar ferramentas *open source* que implementem abordagens de Visão Computacional eficazes para a detecção de objetos alvo (pessoas) em imagens aéreas obtidas com sensores do espectro visível (RGB); e mensurar a acurácia de diferentes tipos de descritores de características.

O restante do trabalho está organizado como segue: na seção 2 é apresentada a metodologia adotada no trabalho; na seção 3 são expostos resultados preliminares e uma discussão sobre estes; e na seção 4 são elencadas as principais conclusões parciais obtidas.

2 METODOLOGIA

Foram criados classificadores em cascata que se fundamentam na união de diversos classificadores para ter uma maior chance de sucesso (VIOLA; JONES, 2001). A Figura 1 mostra o funcionamento de uma cascata de classificadores.

Figura 1 - Cascata de classificadores.

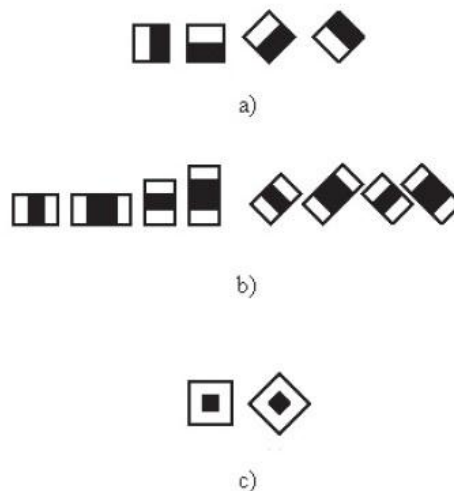


Fonte: Próprio do autor (2018).

Segundo Zhang e Ma (2012), os primeiros classificadores de uma cascata são aqueles mais rápidos que eliminarão com facilidade objetos que são negativos, os classificadores superiores terão uma maior complexidade e serão mais lentos, porém eles confirmarão se realmente os objetos são positivos. No presente trabalho foram construídos descritores baseados em Haar e *Local Binary Pattern* (LBP).

Características Haar baseiam-se no cálculo eficiente das diferenças de pixels que estão presentes na imagem (VIOLA; JONES, 2004). A Figura 2 apresenta os tipos de características Haar, sendo que o cálculo é feito através da soma dos pixels da área preta que é subtraído da soma dos pixels das áreas brancas, assim, percorrendo toda a imagem.

Figura 2 - Características: a) de borda; b) de linha; e c) de ponto.



Fonte: Adaptado de Lienhart, Kuranov e Pisarevsky (2002).

Já as características LBP se baseiam em textura que é definida como a variação espacial na intensidade dos *pixels* e calcula para cada *pixel* uma configuração de acordo com a sua vizinhança (PIETIKÄINEN et. al, 2011). O descritor utiliza uma vizinhança 3x3 e é feita uma limiarização de todos os *pixels* com o *pixel* central, efetuando oito subtrações e atribuindo um quando o resultado for positivo e zero caso contrário. A Figura 3 apresenta as etapas do classificador LBP.

Figura 3 - Etapas de extração de características LBP.

6	5	2
7	6	1
9	3	7

(a)

1	0	0
1		0
1	0	1

(b)

1	2	4
8		16
32	64	128

(c)

1	0	0
8		0
32	0	128

(d)

Fonte: Adaptado de Ojala, Pietikäinen e Harwood (1996, p. 5)

O método foi desenvolvido em linguagem de programação Python, já que esta possui suporte para biblioteca de Visão Computacional OpenCV. O método consiste em um *script* que recebe um conjunto de imagens como entrada, após ele aciona os classificadores desenvolvidos e gera como saída um subconjunto de imagens onde houve detecção de objetos alvo, nesse caso, pessoas.

Para efetivamente realizar o treinamento dos classificadores foi utilizado o *software* “Cascade-Trainer-GUI” (AHMADI, 2016), o qual utiliza as aplicações disponíveis na biblioteca OpenCV, porém facilita a manipulação das mesmas, visto que fornece uma interface gráfica para o usuário.

3 RESULTADOS e DISCUSSÃO

Foram gerados classificadores de 8, 15 e 20 estágios para Haar e LBP e utilizado como *dataset* o “INRIA person DataSet” (DALAL; TRIGGS, 2005). Para a criação foi utilizado um conjunto de 300 imagens positivas que são aquelas que contém os objetos alvo e 1.200 imagens negativas que são as que não contém objetos alvo. A Tabela 1 apresenta o tempo de treinamento dos classificadores Haar e LBP em segundos, de acordo com a quantidade de estágios. É importante notar que o treinamento do LBP é mais rápido, o que evidencia uma maior demanda de recursos computacionais disponíveis quando se quer trabalhar com descritor Haar.

Tabela 1 -Tempo de treinamento dos classificadores

Nº Estágios	Haar	LBP
8	448 s	7 s
15	1233 s	54 s
20	2409 s	385 s

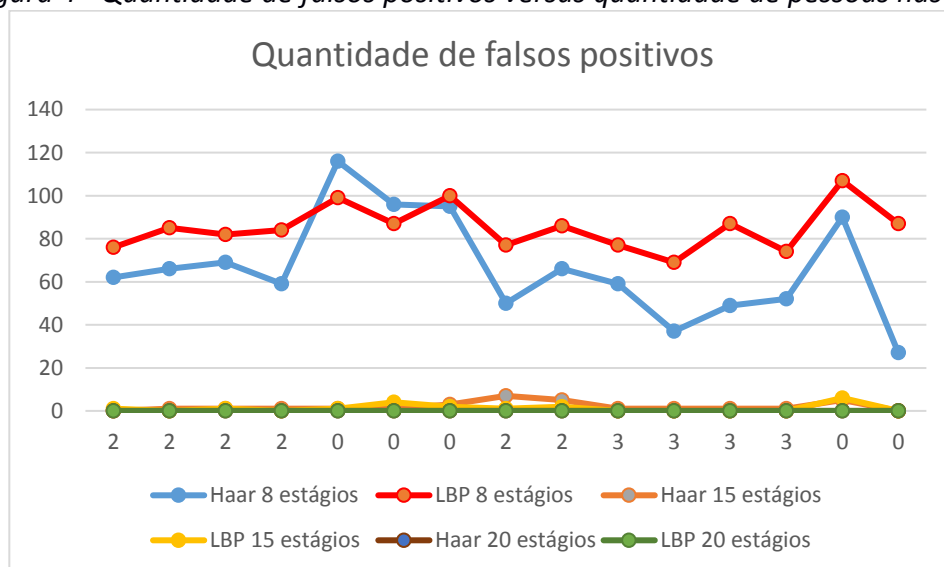
Fonte: do autor, 2018.

Para a realização dos testes dos classificadores foram utilizadas como entradas 10 fotos positivas e 5 negativas. As imagens positivas foram extraídas do *dataset* denominado “DroneFace” (HSU; CHEN, 2017), escolhidas de acordo com a maior distância entre as pessoas e a câmera, posteriormente submetidas a um redimensionamento para 491 x 368 *pixels*. As imagens negativas foram coletadas aleatoriamente no Google.

Os resultados dos testes realizados são apresentados na Figura 4. Para cada classificador, independentemente do número de estágios, foram utilizados os mesmos dados de entrada e de treinamento.

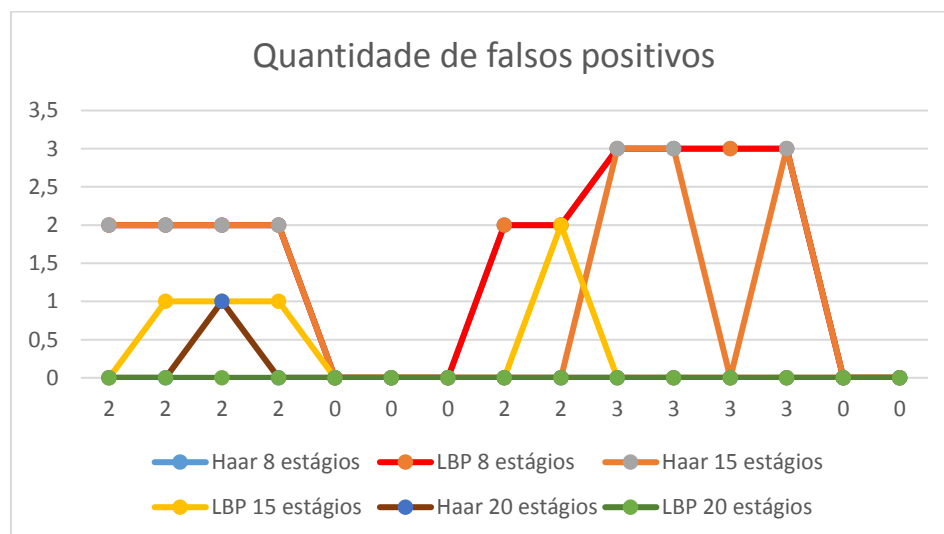
Especificamente, percebe-se que a Figura 4 mostra os resultados para o classificador Haar e LBP de 8, 15 e 20 estágios, com o eixo y apresentando a ocorrência de falsos positivos e o eixo x a quantidade de pessoas presentes nas imagens. No classificador de 8 estágios tanto para Haar quanto LBP, a falta de estágios suficientes para ter uma maior acurácia gerou diversos falsos positivos, que correspondem à detecção de objetos que não interessam no trabalho. Os classificadores com 20 estágios não geraram nenhum tipo de falso positivo, porém isso ocasiona a ocorrência de muitos falsos negativos, que são a não detecção de objetos de interesse. Isso pode ser verificado na Figura 5 que apresenta no eixo y o número de detecção de pessoas nas imagens e no eixo x a quantidade de pessoas presentes nas mesmas.

Figura 4 - Quantidade de falsos positivos versus quantidade de pessoas nas imagens.



Fonte: do autor, 2018.

Figura 5 - Quantidade de falsos positivos versus quantidade de pessoas nas imagens.



Fonte: do autor, 2018.

4 CONSIDERAÇÕES FINAIS

Os testes preliminares realizados no presente trabalho evidenciaram que as técnicas de Visão Computacional e o treinamento de classificadores em cascata tem grande potencial para auxiliar no combate ao crime de abigeato em propriedades rurais, visto que a detecção de pessoas feita pelos classificadores Haar de 15 estágios em imagens com ângulo superior a 90° foi efetiva. Vale ressaltar que as configurações utilizadas nos treinamentos dos classificadores foram as apresentadas de maneira padrão pelo *software* utilizado, ou seja, para casos específicos os classificadores necessitam de manipulação nos parâmetros para que se obtenha resultados satisfatórios.

REFERÊNCIAS

- AHMADI, A. Cascade Trainer GUI. 2016. Disponível: <<http://amin-ahmadi.com/cascade-trainer-gui/>>. Acesso em: 05/09/2018.
- BALLARD, D. H.; BROWN, C. M. **Computer Vision**. 1. ed. Prentice Hall, 1982.
- CRÓSTA, A. P. **Processamento Digital de Imagens de Sensoriamento Remoto**. Campinas, IG/UNICAMP, 1992.
- DALAL, N.; TRIGGS, B. Histograms of oriented gradients for human detection. In: **2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)**. [S.l.: s.n.], 2005.
- HSU, H; CHEN, K. DroneFace: An Open Dataset for Drone Research. In: **MMSys'17 Proceedings of the 8th ACM on Multimedia Systems Conference**. 2017.
- LIENHART, R.; KURANOV, A.; PISAREVSKY, V. Empirical analysis of detection cascades of boosted classifiers for rapid object detection. **Pattern Recognition**, 2002.
- MAKAROV, M. et al. Octorotor uavs for radar applications: Modeling and analysis for control design. **Workshop on RED-UAS**, IEEE, v. 2, 2015. 571p.
- OJALA, T.; PIETIKÄINEN, M.; HARWOOD, D. A comparative study of texture measures with classification based on featured distributions. **Pattern recognition**, v. 29, p. 51–59, 1996.
- PIETIKÄINEN, M. et al. **Computer Vision Using Local Binary Patterns**. Springer, 2011.
- VIOLA, P.; JONE, M. J. Robust real-time face detection. **International Journal of Computer Vision**, 2004.
- VIOLA, P.; JONES, M. Rapid object detection using a boosted cascade of simple features. **IEEE Computer Society Conference on Computer Vision and Pattern Recognition**, 2001.
- ZHANG, C.; MA, Y. **Ensemble Machine Learning: Methods and Applications**. Springer, 2012.