

IDENTIFICAÇÃO DE PRAGAS EM ARMADILHAS INTELIGENTES UTILIZANDO PDI E AM

THAINAN BYSTRONSKI REMBOSKI¹; WILLIAM DALMORRA DE SOUZA¹;
PAULO ROBERTO FERREIRA JR.¹; MARILTON SANCHOTENE DE AGUIAR¹

¹Universidade Federal de Pelotas – {tbremboski, wddsouza, paulo, marilton}@inf.ufpel.edu.br

1. INTRODUÇÃO

O manejo integrado de pragas (MIP) é uma técnica que procura manter as populações de pragas abaixo do nível que possam causar dano econômico. Este controle pode ser feito por meio de insetos (controle biológico), uso de feromônios, retirada e queima da parte do vegetal afetado, entre outros. O MIP caracteriza-se como uma alternativa que procura integrar metodologias para desenvolver um controle de pragas que seja prático, eficiente, econômico, que proteja a saúde pública e ambiental, visando diminuir o uso de agrotóxicos causadores da contaminação dos alimentos e do lençol freático quando aplicados indiscriminadamente (ELLIOTT et al., 1995; SHELTON; BADENES-PEREZ, 2006). Uma técnica bastante empregada no monitoramento da população para o controle de insetos-praga é baseada na utilização de armadilhas, com feromônios ou atrativos alimentares que atraem os insetos. Ao entrarem nas armadilhas, distribuídas na área da plantação, os insetos são incapazes de deixá-la, ficando capturados em seu interior (LÓPEZ et al., 2012).

O método usual para coletar dados das armadilhas, determinando o número de indivíduos capturados, consiste em levantamentos de campo regulares com a observação visual de armadilhas por um operador humano. A periodicidade das visitas do operador é geralmente semanal, podendo ser mais espaçadas dependendo do tamanho da propriedade (TIRELLI et al., 2011). Este método tem alguns inconvenientes, nos quais pode-se destacar: o trabalho intensivo, insalubre e oneroso de deslocar um humano pela área da plantação; e, não se pode obter informações em tempo real e sincronizadas de todas as armadilhas, uma vez que elas vão sendo visitadas ao longo da semana (LÓPEZ et al., 2012).

Uma alternativa para facilitar e agilizar esse processo de monitoramento é utilizar armadilhas inteligentes que, dentre outras funcionalidades, sejam capazes de gerar imagens e processá-las, realizando a identificação dos insetos capturados. A tarefa de identificar insetos através do uso de imagens não é trivial, pois conta com diversos fatores que podem afetar esse processo. Um desses fatores é a possível sobreposição de insetos, pois, como as armadilhas são posicionadas em um ambiente externo, outros insetos maiores, bem como folhas de árvores e outros objetos, podem se tornar um obstáculo entre a câmera e a base interior da armadilha, não possibilitando a visão dos insetos menores através da lente da câmera.

Outro fator que dificulta bastante o processo de identificação é a grande variedade de insetos diferentes que são capturados juntamente dos insetos de interesse. Apesar das armadilhas possuírem feromônios para atração dos insetos de interesse, isso não impede que outros insetos fiquem presos na armadilha, o que é bastante comum de acontecer. Ainda, dada a abertura das asas do inseto, sua forma pode variar significativamente, gerando uma gama de possibilidades a

serem consideradas na identificação, tornando-se, assim, um agravante extra nesse processo.

Os insetos de interesse do presente artigo são as moscas-das-frutas *Ceratitis capitata* e *Anastrepha fraterculus*, que estão entre as principais pragas presentes na fruticultura mundial. As moscas-das-frutas dos gêneros *Anastrepha* e *Ceratitis* apresentam uma larga distribuição geográfica no Brasil, sendo a *Ceratitis capitata*, também conhecida como mosca-do-mediterrâneo, a única espécie do gênero registrada no Brasil, onde já se encontra amplamente disseminada e adaptada a ambientes rurais e urbanos.

Além disso, do ponto de vista econômico da fruticultura mundial, a mosca-do-mediterrâneo é a mais importante dentre as moscas-das-frutas, por ser a espécie mais cosmopolita e pela sua capacidade de transmitir a bactéria *Escherichia coli* para frutos comerciais. Também, a presença de moscas-das-frutas faz com que os países importadores de frutos criem barreiras fitossanitárias, dificultando a exportação de frutos para estes países (MONTES; RAGA, 2006; SELA et al., 2005).

Com base no contexto apresentado, este artigo propõe um algoritmo para identificação dos insetos *Anastrepha fraterculus* e *Ceratitis capitata*, utilizando-se de técnicas de processamento digital de imagens (PDI) e aprendizagem de máquina (AM), capturados em armadilhas convencionais. A proposta de um algoritmo eficiente para identificação dos insetos alvo é o primeiro passo no sentido de desenvolver uma armadilha inteligente, capaz de capturar imagens, identificando e contando os insetos capturados, de forma automática, periódica e remota.

2. METODOLOGIA

A abordagem proposta foi dividida em três etapas, que são: segmentação da imagem, extração de características e classificação. Na etapa de segmentação, partindo de uma imagem como a Figura 1a, diversos métodos de segmentação de imagens foram aplicados com o intuito de separar os objetos de interesse (insetos-praga) do fundo da imagem.

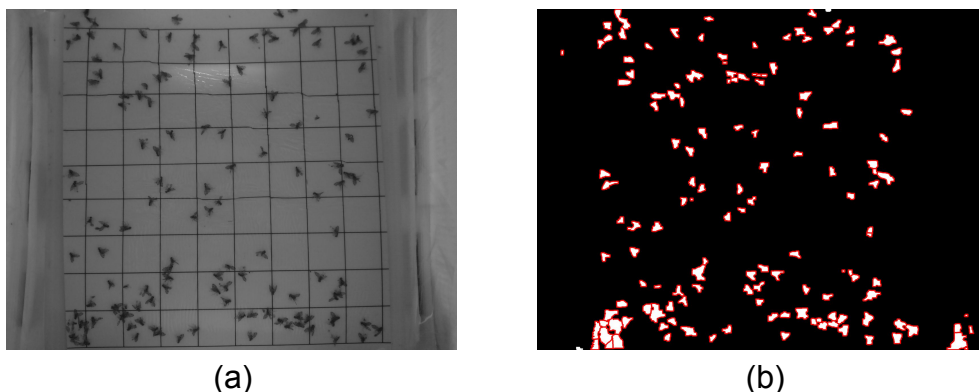


Figura 1: Imagem a) da base da armadilha e b) após a segmentação.

Dentre os métodos usados, destacam-se: algoritmo de *thresholding* global e adaptativo; operações morfológicas de abertura (*opening*) para remoção de objetos indesejados da imagem (ruído); testes condicionais baseados na excentricidade e verticalidade das regiões segmentadas para remover objetos que diferem da forma de um inseto; e, por fim, aplicação do algoritmo de *watershed*

para subdividir regiões maiores (aglomerado de insetos) em regiões menores (insetos individuais) para facilitar a sua identificação. O resultado final da aplicação desses métodos pode ser observado na Figura 1b.

Na etapa de extração de características, para cada região segmentada da Figura 1b, foi aplicado o algoritmo de extração de características SURF, gerando uma lista de descritores de cada região. Esses descritores foram clusterizados, utilizando o algoritmo *K-means*, com a finalidade de agrupar descritores que apresentam características semelhantes. Após a clusterização, cada região segmentada passa a ser representada por um vetor de características, onde cada posição do vetor representa um *cluster* e o valor de cada posição representa quantos descritores desta região pertencem a esse *cluster*. Esse método de extração de características é chamado de *bag-of-words* pois é bastante semelhante ao modo como é utilizado para classificação de texto.

Na etapa de classificação, os vetores de características foram utilizados como dados de entrada para os seguintes algoritmos de aprendizado supervisionado: árvores de decisão (AD), *Gaussian Naive Bayes* (GNB), *K-Nearest Neighbors* (KNN) e *Support Vector Machines* (SVM). Os dados de entrada foram divididos em conjuntos de treinamento (67%) e de teste (33%).

3. RESULTADOS E DISCUSSÃO

A base de treino para a realização deste trabalho consistiu de 99 imagens, com resolução de 2592x1944 *pixels*. As coletas foram realizadas por uma câmera de 5 *megapixels* em uma placa *Raspberry Pi 3*. Tanto a câmera quanto a placa foram acopladas a uma armadilha entomológica, desenvolvida de forma paralela a este trabalho. O conjunto de exemplos de entrada, construído a partir das imagens da base de treino, é constituído de 4338 exemplos, onde foi definido um conjunto de classes composto pelas seguintes classes: *Ceratitis capitata*, *Anastrepha fraterculus* e resíduo.

No que tange aos parâmetros dos algoritmos utilizados, para o algoritmo SVM foi utilizado o *kernel* "rbf" ($\gamma = 0.1$) pois, dentre os outros *kernels*, este apresentou melhores resultados. No KNN, foi utilizado $K = 3$, pois este apresentou melhores resultados que o padrão ($K = 5$), para este conjunto de dados. Para os demais algoritmos, foram utilizados os parâmetros padrões da ferramenta utilizada. A Tabela 1 apresenta um comparativo das médias de todos algoritmos.

Tabela 1: Comparação entre as métricas de todos algoritmos.

Algoritmo	Precisão	Recall	Métrica F1	Acurácia
AD	0,80±0,02	0,79±0,06	0,79±0,04	80,37%
GNB	0,76±0,01	0,73±0,16	0,73±0,08	75,41%
KNN	0,82±0,12	0,70±0,19	0,73±0,05	74,65%
SVM	0,86±0,02	0,86±0,04	0,86±0,03	86,38%

A informação de que foram encontrados exemplares de *Ceratitis capitata* e *Anastrepha fraterculus* em armadilhas espalhadas pela plantação é muito importante para que possa ser realizado o controle adequado. Logo, uma informação incorreta, onde o algoritmo classifica um desses insetos como se fosse ruído pode ser muito grave. Dessa forma, é desejável que tanto a precisão

quanto o *recall* possuam valores elevados, fazendo com que se reduza o número de falsos positivos e falsos negativos encontrados. Com base nos resultados apresentados, o algoritmo SVM se mostrou superior em todas as métricas.

4. CONCLUSÕES

O presente artigo apresenta uma proposta de algoritmo que realiza a identificação dos insetos capturados no interior de armadilhas, utilizando técnicas de processamento digital de imagens e de aprendizagem de máquina.

Esta abordagem proposta é dividida em três etapas: i) segmentação da imagem, onde separam-se as áreas de interesse do fundo da imagem; ii) extração de características, onde se empregou o modelo *bag-of-words* para a geração dos vetores de características; e, iii) classificação, onde foram utilizados os algoritmos de aprendizado supervisionado (SVM, KNN, árvore de decisão e *Gaussian Naive Bayes*), onde SVM apresentou resultados superiores em relação aos demais algoritmos.

Como trabalhos futuros, propõe-se o melhoramento da base de treino com a adição de mais imagens. Também propõe-se a aplicação de outros algoritmos de classificação como CNN (*Convolutional Neural Network*) e algoritmos de *ensemble* (conjunto), como *AdaBoost*.

Por fim, pode-se incrementar o método ao utilizar informações temporais (históricas) de imagens anteriores, com o intuito de acompanhar o crescimento da população de insetos-praga no interior da armadilha, de forma a evitar problemas como a sobreposição de insetos, auxiliando a identificação.

5. REFERÊNCIAS BIBLIOGRÁFICAS

- ELLIOTT, N.; FARRELL, J.; GUTIERREZ, A.; VAN LENTEREN, J. C.; WALTON, M.; WRATTEN, S.; DENT, D. **Integrated Pest Management**. Berlim: Springer Science and Business Media, 1995.
- LÓPEZ, O.; RACH, M. M.; MIGALLON, H.; MALUMBRES, M. P.; BONASTRE, A.; SERRANO, J. J. Monitoring pest insect traps by means of low-power image sensor technologies. **Sensors**, Basel, v.12, n.11, p.15801-15819, 2012.
- MONTES, S.; RAGA, A. Eficácia de atrativos para monitoramento de *Ceratitis capitata* (diptera: Tephritidae) em pomar de citros. **Arquivos do Instituto Biológico**, São Paulo, v.73, n.1, p.317-323, 2006.
- SELA, S.; NESTEL, D.; PINTO, R.; NEMNY-LAVY, E.; BAR-JOSEPH, M. Mediterranean fruit fly as a potential vector of bacterial pathogens. *Applied and Environmental Microbiology*, Washington DC, v.71, n.7, p.4052-4056, 2005.
- SHELTON, A.; BADENES-PEREZ, F. Concepts and applications of trap cropping in pest management. **Annual Review of Entomology**, Palo Alto, v.51, n.1, p.285-308, 2006.
- TIRELLI, P.; BORGHESE, N.; PEDERSINI, F.; GALASSI, G.; OBERTI, R. Automatic monitoring of pest insects traps by zigbee-based wireless networking of image sensors. In: **INSTRUMENTATION AND MEASUREMENT TECHNOLOGY CONFERENCE (I2MTC)**, 5., Binjiang, Hangzhou, China, 2011. Proceedings... New York: IEEE, 2011. p.1-5.