



## SISTEMA INTEGRADO DE ARMADILHAS INTELIGENTES E MODELO DE PREDIÇÃO DE INFESTAÇÕES PARA O MANEJO DE PRAGAS

THAINAN BYSTRONSKI REMBOSKI<sup>1</sup>; WILLIAM DALMORRA DE SOUZA<sup>1</sup>;  
PAULO ROBERTO FERREIRA JR.<sup>1</sup>; MARILTON SANCHOTENE DE AGUIAR<sup>1</sup>

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

### 1. INTRODUÇÃO

O manejo integrado de pragas (MIP) é uma técnica que busca manter as pragas sempre abaixo do nível em que são consideradas danosas para as plantações. Este controle pode ser feito por meio de insetos (controle biológico), uso de feromônios, retirada e queima da parte vegetal afetada, entre outros (ELLIOTT et. al., 1995).

O MIP é uma alternativa que procura integrar metodologias para desenvolver um controle de pragas que seja prático, eficiente, econômico e que proteja a saúde pública e ambiental, visando diminuir o uso de agrotóxicos, que causam a contaminação dos alimentos e do lençol freático quando aplicados indiscriminadamente (ELLIOTT et. al., 1995).

Estes agrotóxicos, além disso, podem ser extremamente danosos à saúde, tanto dos consumidores quanto dos responsáveis pela aplicação do produto. Existem diversos estudos relacionados ao uso de pesticidas e enfermidades associadas. Por exemplo, em MOSTAFALOU; ABDOLLAHI (2013), são listadas diversas doenças crônicas e as evidências da relação de seu desenvolvimento devido à exposição a pesticidas.

A implementação dos programas de MIP é baseada no monitoramento das populações de pragas na áreas de produção. Uma técnica bastante empregada no monitoramento de insetos-praga é o uso de armadilhas, as quais podem ser iscadas com atrativos sexuais (feromônios) ou alimentares (LÓPEZ et. al., 2012).

O método usual para coletar dados, 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 maior dependendo do tamanho da propriedade (TIRELLI et. al., 2011). Este método tem alguns inconvenientes, nos quais pode-se destacar: o trabalho intenso, 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.

O presente projeto propõe um sistema de monitoramento de pragas baseado em armadilhas inteligentes. Pretende-se fazer isso tratando as armadilhas como sensores, que compõem uma rede de sensores sem fio e que utilizam o processamento de imagens capturadas por câmeras para realizar a contagem automática de indivíduos.

Além disso, como insetos possuem sangue frio, qualquer variação de temperatura na região afeta fortemente a temperatura corporal desses animais, logo eles são diretamente influenciados pelas alterações climáticas do local onde estão. Além da temperatura, umidade, precipitação, velocidade do vento, entre outras, são variáveis também conhecidas por influenciar a vida dos insetos



(PORTER; PARRY; CARTER, 1991). Com base nessas informações, desenvolveu-se um modelo, através de algoritmos de aprendizagem de máquina, para a predição de possíveis infestações de pragas em lavouras, utilizando as informações de variações meteorológicas da região a partir de um banco de dados com informações históricas, mais especificamente de estações que cobrem microrregiões de algumas plantações no município de Pelotas. Este modelo integrado com o sistema de monitoramento baseado nas armadilhas compõe um sistema de decisão que apoia o produtor na realização do MIP.

Os insetos de interesse deste trabalho são as moscas-das-frutas *Anastrepha fraterculus* e *Ceratitis capitata*, que estão entre as principais pragas presentes na fruticultura mundial. A etapa de predição de infestações também conta com informações do inseto *Grapholita molesta*, portanto, este inseto também é considerado na etapa de predição.

## 2. METODOLOGIA

A metodologia deste projeto está dividida em duas partes diferentes. Primeiro é apresentado o desenvolvimento do algoritmo de identificação de insetos-praga através de imagens e, posteriormente é apresentado o modelo de predição de infestações com base em informações meteorológicas.

O algoritmo de identificação de insetos-praga foi dividido em três etapas principais: segmentação da imagem, extração de características e classificação. Partindo de uma imagem capturada no interior da armadilha, diversos métodos de segmentação de imagens foram aplicados com o intuito de separar os objetos de interesse (insetos-praga) do fundo da imagem. Dentre os métodos usados, destacam-se: algoritmo de *thresholding* global e adaptativo; operações morfológicas de abertura (*opening*); testes condicionais baseados na excentricidade e verticalidade das regiões segmentadas; 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.

Para cada região segmentada, foi aplicado o algoritmo de extração de características SURF (*Speeded-Up Robust Features*), gerando uma lista de descritores de cada região. Esses descritores foram clusterizados, utilizando o algoritmo *Kmeans*. Após a clusterização, cada região segmentada passa a ser representada por um vetor de características. Esse método de extração de características é chamado de *bag-of-words*.

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

Os dados meteorológicos necessários para a implementação do modelo de predição foram disponibilizados pela empresa Embrapa. O banco de dados possui informações como temperatura, umidade, precipitação, velocidade do vento, entre outras, sendo todas conhecidas por afetarem o ciclo de vida e o comportamento de determinados insetos. Estes dados foram coletados ao longo de 5 anos, formando uma base de dados consistente. Para o tratamento dos dados foi realizada a aplicação do processo de mineração de dados, que é dividido em diversas etapas, desde a limpeza dos dados até a análise dos padrões encontrados.



Após a etapa de mineração dos dados, a predição de infestação é realizada para um período de quatro dias a frente, observando as informações dos três dias anteriores. Os algoritmos de classificação utilizados para realizar a predição foram: SVM, AD, *Bernoulli Naive Bayes* (BNB) e *Neural Networks* (NN).

### 3. RESULTADOS E DISCUSSÃO

A base de treino para a realização do trabalho de identificação dos insetos consiste de 99 imagens, com resolução de 2592x1944 pixels. As coletas foram realizadas por uma câmera de 5 megapixels de uma placa Raspberry Pi 3. Tanto a câmera quanto a placa foram acopladas a uma armadilha entomológica, desenvolvida em conjunto com o projeto. O conjunto de exemplos de entrada, construído a partir das imagens da base de treino, é constituído de 4338 exemplos. Com base nos exemplos coletados, as classes definidas para classificação são: *Ceratitis capitata*, *Anastrepha fraterculus* e resíduo. Para a classificação com aprendizado supervisionado, os dados de entrada foram divididos em dois conjuntos: conjunto de treinamento, com 67% dos exemplos; e conjunto de testes, com 33% dos exemplos. A Tabela 1 apresenta os resultados obtidos da classificação.

Tabela 1: Comparação dos resultados dos algoritmos de classificação.

Alg.	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 aos outros algoritmos.

Na etapa de predição, os resultados obtidos pelos algoritmos foram comparados e avaliados utilizando a Curva ROC, que é uma técnica comum para avaliação de modelos de classificação binária. Para validação dos resultados, foram utilizadas a técnica de validação cruzada e foi feita uma análise da matriz de confusão gerada por cada modelo. Os resultados obtidos pela área sob a Curva ROC de cada modelo podem ser observados na Tabela 2.

As curvas ROC apresentadas mostram que as taxas de positivos verdadeiros ultrapassam as taxas de falsos positivos, o que significa uma baixa taxa de falsos negativos. Logo, pode-se afirmar que, de acordo com os resultados apresentados, os modelos são suficientes para previsão de infestações, com o modelo de *Neural Network* atingindo os melhores resultados.



Tabela 2: Área sob a curva ROC dos algoritmos utilizados.

Alg.	<i>Anastrepha fraterculus</i>	<i>Ceratitis capitata</i>	<i>Grapholita molesta</i>
SVM	0,68	0,84	0,77
BNB	0,86	0,85	0,77
AD	0,72	0,83	0,71
NN	0,92	0,97	0,96

#### 4. CONCLUSÕES

O projeto apresentado faz parte de um projeto maior, onde todas as partes serão interligadas, incluindo a comunicação entre as armadilhas, que estão sendo desenvolvidas paralelamente a este projeto.

As contribuições esperadas serão bastante relevantes, pois trarão benefícios tanto para o produtor, reduzindo os custos com os agrotóxicos, quanto para a população em geral, que consumirá produtos mais saudáveis. Além disso, reduzirá o trabalho oneroso e cansativo de realizar a manutenção das armadilhas, pois além de ser um trabalho desgastante, também pode gerar resultados imprecisos, que são causados por fatores humanos.

Como trabalhos futuros, pretende-se realizar coletas de mais imagens, com a finalidade de gerar uma base de dados mais robusta para a tarefa de identificação. Para a predição, pretende-se a captura de novos dados, meteorológicos e populacionais, permitindo o aprimoramento deste modelo com o acompanhamento do crescimento da população dos insetos. Após, os trabalhos serão validados junto à empresa Partamon, parceira deste projeto.

#### 5. REFERÊNCIAS BIBLIOGRÁFICAS

- ELLIOTT, N.; FARRELL, J.; GUTIERREZ, A.; LENTEREN, J. C. van; WALTON, M.; WRATTEN, S.; DENT, D. **Integrated pest management**. [S.l.]: Springer Science & 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.
- MOSTAFALOU, S.; ABDOLLAHI, M. Pesticides and human chronic diseases: evidences, mechanisms, and perspectives. **Toxicology and applied pharmacology**, [S.l.], v.268, n.2, p.157-177, 2013.
- PORTER, J.; PARRY, M.; CARTER, T. The potential effects of climatic change on agricultural insect pests. **Agricultural and Forest Meteorology**, [S.l.], v.57, n.1, p.221- 240, 1991.
- 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.