

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ
DAELN - DEPARTAMENTO ACADÊMICO DE ELETRÔNICA
CURSO DE ENGENHARIA ELETRÔNICA

XIE JIAYU

**SISTEMA DE IDENTIFICAÇÃO ENTOMOLÓGICA FORENSE:
UMA ABORDAGEM USANDO LABVIEW**

TRABALHO DE CONCLUSÃO DE CURSO

CURITIBA
2018

XIE JIAYU

**SISTEMA DE IDENTIFICAÇÃO ENTOMOLÓGICA FORENSE:
UMA ABORDAGEM USANDO LABVIEW**

Trabalho de Conclusão de Curso apresentado ao Curso de Engenharia Eletrônica da Universidade Tecnológica Federal do Paraná, como requisito parcial para a obtenção do título de Bacharel.

Orientador: Prof. Rubens Alexandre de Faria, PhD
Universidade Tecnológica Federal do Paraná

Coorientadora: Dra. Ozana M. Andrade Maia
Pesquisadora/pós-doutora - CPGEI/UTFPR

CURITIBA
2018

XIE JIAYU

SISTEMA DE IDENTIFICAÇÃO ENTOMOLÓGICA FORENSE: UMA ABORDAGEM USANDO LABVIEW

Este Trabalho de Conclusão de Curso de Graduação foi apresentado como requisito parcial para obtenção do título de Engenheiro Eletrônico, do curso de Engenharia Eletrônica do Departamento Acadêmico de Eletrônica (DAELN) outorgado pela Universidade Tecnológica Federal do Paraná (UTFPR). A aluna foi arguida pela Banca Examinadora composta pelos professores abaixo assinados. Após deliberação, a Banca Examinadora considerou o trabalho aprovado.

Curitiba, 06 de dezembro de 2018.

Prof. Dr. Robinson Vída Noronha
Coordenador de Curso
Engenharia Eletrônica

Profª. Drª. Carmen Caroline Rasera
Responsável pelos Trabalhos de Conclusão de Curso
de Engenharia Eletrônica do DAELN

BANCA EXAMINADORA

Prof. Rubens Alexandre de Faria, PhD
Universidade Tecnológica Federal do Paraná
Orientador

Prof. Dr André Eugênio Lazzaretti
Universidade Tecnológica Federal do Paraná

Prof. Dra. Ozana M. Andrade Maia
Coorientadora

Prof. Dr. Gustavo Benvenuti Borba
Universidade Tecnológica Federal do Paraná

A folha de aprovação assinada encontra-se na Coordenação do Curso de Engenharia Eletrônica.

Dedico este trabalho a todos aqueles que acreditaram em mim um dia.

AGRADECIMENTOS

Agradeço, primeiramente, a Deus pela oportunidade de vida. Por me deixar errar para poder apreciar o perdão. Pelas portas que abriu para mim ao longo dos 5 anos de engenharia. Pelas experiências únicas das quais aproveitei o máximo possível. Por me guiar mesmo em dias tenebrosos. Enfim, por nunca ter me deixado sozinha.

Agradeço a meus pais pelo suporte dado e pelos ensinamentos da vida, por terem me passado virtudes como persistência e humildade. Agradeço aos meus irmãos pelo companheirismo e pelo apoio moral que sempre me deram, em especial à minha irmã que, mesmo longe, sempre ouviu meus desabafos em dias tristes, mas que vibrava com intensidade maior ainda quando compartilhava projetos meus que deram êxito. Obrigada por me motivarem a querer ser uma pessoa melhor todos os dias.

Ao professor Rubens Alexandre de Faria pelo apoio fornecido, seja técnico, profissional ou moral. Obrigada, principalmente, pelas orientações transmitidas sobre a vida, dicas e conselhos que não se encontram em manual ou *datasheet* algum. Agradeço também pela oportunidade em trabalhar no projeto microvestígios forenses/CAPES, cujo esforço resultou neste trabalho.

À professora Ozana M. Andrade Maia pela amizade, apoio e torcida. Sua ajuda na área entomológica foi crucial para poder entender melhor a importância desse projeto.

Ao professor André Lazzaretti pela oportunidade em trabalhar com o LabVIEW no projeto da COPEL, experiência que me permitiu avançar e usar essa ferramenta poderosa neste trabalho.

Ao professor Gustavo Borba por fazer eu me apaixonar ainda mais pela área de processamento de imagens. Sua excelente didática e domínio de conteúdo ampliaram o meu conhecimento nesse campo de pesquisa.

Aos amigos que sempre me apoiaram e sustentaram quando faltaram forças para continuar. Vocês me levantaram todas as vezes que caí. Obrigada por serem simplesmente vocês. Amei cada dia conviver com cada um de vocês até agora. E lutarei para que nossa amizade permaneça forte independente da distância geográfica.

“ Conheça todas as teorias, domine todas as técnicas, mas ao tocar uma alma humana, seja apenas outra alma humana. ” (JUNG, Carl)

RESUMO

JIAYU, Xie. Sistema de Identificação Entomológica Forense: uma abordagem usando LabVIEW. 2018. 43 f. Trabalho de Conclusão de Curso – Curso de Engenharia Eletrônica, Universidade Tecnológica Federal do Paraná. Curitiba, 2018.

A entomologia forense é um dos recursos disponíveis mais precisos, no que tange à identificação de origem para determinados crimes, como homicídios, locais de desova, tráfico de drogas e contrabando. Contudo, a falta de um banco de dados entomológico policial completo e um sistema de reconhecimento automático para identificação deste tipo de vestígio, têm retardado a solução de inúmeros crimes de alta repercussão social. O grupo de Engenharia Forense Edmund Loccard da UTFPR, conveniado ao Departamento de Polícia Federal e CAPES desenvolveu um protótipo de banco de dados entomológicos regional, onde são aplicados métodos de reconhecimento de padrões de imagem. Este trabalho teve como objetivo desenvolver e testar um sistema de identificação automática para estas espécies de insetos, utilizando especificamente a ferramenta LabVIEW. Foram implementados diversos algoritmos à esta aplicação, onde foram comparadas imagens de fragmentos parciais de insetos, coletados como vestígios em locais de crime, ao banco de dados. O objetivo final é a identificação de origem da droga apreendida e a consequente identificação de grupos suspeitos. Os resultados obtidos revelaram que foi possível estabelecer um *matching* entre a base de imagens treinadas com imagens de teste, bem como imagens contendo apenas fragmentos de insetos.

Palavras-chave: Entomologia Forense. Microvestígios forenses. Reconhecimento de Imagem. Classificação de objetos. LabVIEW.

ABSTRACT

JIAYU, Xie. Entomological Identification System: an approach using LabVIEW. 2018. 43 f. Trabalho de Conclusão de Curso – Curso de Engenharia Eletrônica, Universidade Tecnológica Federal do Paraná. Curitiba, 2018.

Forensic entomology is one of the most accurate resources available for identifying origin of certain crimes, such as homicides, spawning grounds, drug trafficking, and smuggling. However, the lack of a complete police entomological database and an automatic recognition system to identify this type of traces have delayed the solution of countless crimes of high social repercussion. Forensic Engineering group Edmund Loccard of UTFPR, convened to the Department of Federal Police and CAPES developed a prototype of a regional entomological database, where methods of image pattern recognition are applied. This work aimed to develop and test an automatic identification system for these insect species, specifically using the LabVIEW tool. Several algorithms were implemented in this application, where images of partial fragments of insects, collected as traces at crime sites, were compared to the database. The ultimate goal is to identify the origin of the drug seized and the consequent identification of suspicious groups. The results showed that it was possible to establish a matching between the base of images trained with test images as well as images containing only fragments of insects.

Keywords: Forensic Entomology. Micro-traces. Image Recognition. Object Classification. LabVIEW.

LISTA DE FIGURAS

Figura 1 – Níveis taxonômicos usados na classificação dos seres vivos.	16
Figura 2 – Composição de um inseto genérico.	18
Figura 3 – Exemplo de um <i>Pentatomidae</i>	19
Figura 4 – Exemplo de um <i>Formicidae</i>	19
Figura 5 – Exemplo de um <i>Buprestidae</i>	20
Figura 6 – Etapas de reconhecimento de padrões.	21
Figura 7 – Exemplo de classificação supervisionada e não-supervisionada.	22
Figura 8 – Diagrama macroscópica do trabalho.	24
Figura 9 – Janelas de inicialização do LabVIEW: diagrama de blocos e painel frontal.	25
Figura 10 – Processo de aquisição de imagens, sob ângulos de visualização diferentes.	27
Figura 11 – Diagrama de configurações do classificador.	28
Figura 12 – Exemplo de amostras de imagens tratadas dos insetos. Respectivamente, da esquerda para direita: <i>Buprestidae</i> , <i>Formicidae</i> e <i>Pentatomidae</i>	29
Figura 13 – Diagrama do algoritmo estruturado para desenvolvimento do sistema de identificação entomológico.	30
Figura 14 – Parte 1 do código em LabVIEW.	31
Figura 15 – Parte 2 do código em LabVIEW.	31
Figura 16 – Parte 3 do código em LabVIEW.	32
Figura 17 – Painel Frontal correspondente ao diagrama de blocos construído no LabVIEW.	32
Figura 18 – Resultado usando o classificador construído no LabVIEW.	35
Figura 19 – Imagem original de entrada para teste do sistema de identificação.	35
Figura 20 – Resultado usando o classificador construído no LabVIEW.	36
Figura 21 – Imagem original de entrada para testes no sistema de identificação.	36
Figura 22 – Resultado da saída de cada bloco do algoritmo estruturado na Figura 13. Da esquerda para a direita: (1) Imagem original; (2) Conversão para 8 bits; (3) Filtro mediana; (4) Dilatação; (5) Rotular objetos; (6) Remoção de partículas indesejáveis; (7) Equalização; (8) Classificação.	37
Figura 23 – Fragmentos de <i>Pentatomidae</i>	38

LISTA DE TABELAS

Tabela 1 – Resultado do treinamento usando o método de k-vizinhos mais próximos. .	33
Tabela 2 – Informações das classes treinadas.	33
Tabela 3 – Distância entre as classes treinadas.	34
Tabela 4 – Resultados do classificador quanto a fragmentos de <i>Pentatomidae</i> na entrada.	38
Tabela 5 – Resultado do classificador, para a métrica da soma.	39

LISTA DE ABREVIATURAS E SIGLAS

KN	<i>K-Neighbor</i> (Em português, K-vizinhos)
KNN	<i>K-Nearest Neighbor</i> (Em português, K-vizinhos mais próximos)
ROI	<i>Region Of Interest</i> (Em português, Região De Interesse)
RGB	<i>Red-Green-Blue</i> (Em português, Vermelho-Verde-Azul)
VI	<i>Virtual Instrument</i>
LabVIEW	<i>Laboratory Virtual Instrument Engineering Workbench</i>

SUMÁRIO

1 – INTRODUÇÃO	13
1.1 JUSTIFICATIVA	13
1.2 OBJETIVO	14
1.2.1 Objetivo Geral	14
1.2.2 Objetivos Específicos	14
1.3 RESULTADOS ESPERADOS	14
1.3.1 Tecnológicos	14
1.3.2 Científicos	15
1.4 ORGANIZAÇÃO DO TRABALHO	15
2 – FUNDAMENTAÇÃO TEÓRICA	16
2.1 TAXONOMIA	16
2.2 INSETOS	17
2.2.1 Diferenciação quanto à cutícula	17
2.2.2 Diferenciação quanto à coloração dos insetos	18
2.2.3 Fisiologia em geral	18
2.3 <i>HEMIPTERA-PENTATOMIDAE</i>	18
2.4 <i>HYMENOPTERA-FORMICIDAE</i>	19
2.5 <i>COLEOPTERA-BUPRESTIDAE</i>	19
2.6 RECONHECIMENTO DE PADRÕES	20
2.7 CLASSIFICAÇÃO DE IMAGENS	21
2.8 NOÇÃO DE DISTÂNCIA NA CLASSIFICAÇÃO	22
2.9 MÉTODO DE CLASSIFICAÇÃO KNN	23
2.10 FILTRO DE MEDIANA	23
3 – METODOLOGIA	24
3.1 VISÃO GERAL DO LABVIEW	25
3.2 MOTIVO DA ESCOLHA DO LABVIEW	26
3.3 AQUISIÇÃO DE IMAGENS	26
4 – DESENVOLVIMENTO	28
4.1 CONFIGURAÇÕES DO CLASSIFICADOR	28
4.2 SISTEMA DE IDENTIFICAÇÃO	29
4.2.1 Algoritmo em LabVIEW	30
5 – TESTES E RESULTADOS	33
5.1 RESULTADOS DO CLASSIFICADOR	33

5.1.1	Teste 1	34
5.1.2	Teste 2	36
5.2	RESULTADOS UTILIZANDO FRAGMENTOS DE INSETOS	38
5.3	TESTES COM OUTROS CENÁRIOS DE CLASSIFICAÇÃO	39
5.3.1	Quanto ao parâmetro k	39
6	– CONCLUSÃO	40
6.1	TRABALHOS FUTUROS	41
6.2	CONSIDERAÇÕES FINAIS	41
	Referências	42

1 INTRODUÇÃO

Diante de um mundo onde crimes estão sendo aperfeiçoados cada vez mais, faz-se necessário que a perícia forense esteja atualizada, no que diz respeito a recursos tecnológicos, com a finalidade de estar um passo à frente do criminoso.

O estudo de microvestígios e vestígios de maneira geral, entre eles os geológicos, pedológicos, paleontológicos, palinológicos e entomológicos, aplicado a investigações forenses, encontra-se em estado incipiente no Brasil. Alguns trabalhos isolados demonstraram inequivocamente o alto potencial de elucidação de crimes. Devido à particularidades da estrutura, morfologia, dispersão e abundância de vestígios minerais, vegetais, animais e antrópicos, pode-se gerar importantes subsídios a investigações civis e criminais.

Em muitas situações reais atendidas pelas equipes de perícia criminal são encontrados resíduos e marcas de diferentes categorias em suportes variados, que podem prestar como indicadores de *modus operandi*, trajetórias feitas pelos que perpetraram as ações ou mesmo como indicadores discriminantes dos próprios criminosos.

Um dos exemplos de como a polícia pode estar à frente dos criminosos é a criação de um sistema de identificação entomológica. Tal sistema é mais uma ferramenta para que peritos criminais, ou profissionais da área de segurança pública, obtenham informações acerca dos insetos, desconhecidos à primeira vista, tais como a família a qual pertencem. De posse destes dados, pode-se então identificar a origem de certos produtos em trâmite, através da rastreabilidade dos insetos identificados.

1.1 JUSTIFICATIVA

No Brasil, há escassez de literatura (trabalhos científicos/acadêmicos), bem como o desenvolvimento de projetos que contemplem a utilização de microvestígios para fins forenses na perícia criminal. Desta forma, o trabalho tem como proposta a criação de um sistema automático para a identificação de insetos pertencentes, primeiramente, relativos à presente em material vegetal entorpecente, o qual pode ser utilizado, adaptado e inovado posteriormente por especialistas na área de entomologia e perícia forense ou qualquer pessoa interessada em obter respostas do sistema de identificação. A importância deste tema é sustentado no fato de que o inseto pode, muitas vezes, delimitar a região onde um cadáver esteve, ou mesmo a origem de um lote de narcóticos. Deste modo, os peritos podem tomar as devidas providências a instruir a investigação criminal.

Adicionalmente, aproveitar-se-á a oportunidade da existência de um conjunto de imagens entomológicas, em um banco de dados em formação, para teste em uma plataforma científica bastante versátil e avançada, que é o LabVIEW. As suas extensas funcionalidades tornam possível a construção de uma interface gráfica acessível ao usuário. Em outras palavras,

esta plataforma pode agrupar especialistas de várias áreas, tornando possível uma maior interdisciplinaridade.

1.2 OBJETIVO

1.2.1 Objetivo Geral

Desenvolver um sistema protótipo de identificação entomológica para aplicação forense, utilizando Labview.

1.2.2 Objetivos Específicos

- Classificação entomológica dos insetos contidos no banco de dados;
- Aquisição das imagens de insetos capturados em locais de crime, ou fragmentos destes, para padronização e comparação com o banco de dados;
- Aplicação de filtros para pré-processamento de imagens inseridas no LabVIEW;
- Treinamento do classificador do LabVIEW para estabelecer confiabilidade quanto à comparação desses fragmentos com as imagens do banco de dados.

1.3 RESULTADOS ESPERADOS

Ao fim do projeto, espera-se obter resultados quanto à fidedignidade do sistema de reconhecimento de um microvestígio entomológico, utilizando a ferramenta LabVIEW, quando da comparação com o banco de dados previamente catalogado. Em outras palavras, aspira-se encontrar um percentual de *matching* entre o inseto da imagem de entrada com os já treinados no classificador. Vale ressaltar que o reconhecimento aqui é de 1 para n . Ou seja, existem n classes de imagens na base de treinamento e verifica-se a existência de similaridade entre uma nova imagem e as classes da base. Caso o inseto inserido não atender a nenhuma das classes treinadas, ainda assim receberá uma classificação (aquela que for mais semelhante), mas sua identificação será baixíssima, visto que tal inseto não se assemelha com nenhum de seus semelhantes do grupo para o qual foi atribuído. Por isso, a análise deve ser feita em conjunto, isto é, tanto em cima da classe atribuída, quanto da identificação feita pelo classificador.

1.3.1 Tecnológicos

Esta ferramenta de reconhecimento automático não existe na polícia brasileira até os dias de hoje. Desta forma, há relevância tecnológica em se desenvolver tal projeto, contribuindo não só academicamente, mas também com intenção de tornar um protótipo inicial para uso forense.

1.3.2 Científicos

Os resultados obtidos alavancarão o projeto CAPES com a UTFPR, fomentando o grupo de engenharia forense já existente no PPGEB – Programa de Pós-graduação em Engenharia Biomédica.

1.4 ORGANIZAÇÃO DO TRABALHO

O trabalho foi dividido em 6 capítulos cujo conteúdo está assim separado:

- **Capítulo 1:** Apresenta a introdução ao tema, das motivações do trabalho, dos objetivos almejados e dos resultados esperados. Enseja-se, com esse capítulo, passar um quadro geral sobre as justificativas da realização desse projeto.
- **Capítulo 2:** Neste capítulo será explanada uma breve revisão teórica acerca de vários tópicos tratados ao longo do trabalho, perpassando desde a morfologia dos insetos em geral até uma breve introdução sobre reconhecimento de padrões, métodos de classificação e processamento de imagens em geral.
- **Capítulo 3:** Apresenta a metodologia adotada para a realização do projeto, bem como uma visão geral sobre a ferramenta LabVIEW e do motivo pelo qual optou-se pelo uso dela.
- **Capítulo 4:** Discussão sobre o desenvolvimento do projeto, tanto da montagem do classificador quanto do sistema de identificação propriamente dito, parte em que será analisado também o diagrama de funcionamento do código elaborado no LabVIEW.
- **Capítulo 5:** Esse capítulo foca em mostrar os testes realizados e os resultados obtidos, bem como a discussão das repercussões geradas a partir dos testes. Em seguida, é exposta também uma comparação entre o método de classificação adotada e os outros disponíveis.
- **Capítulo 6:** Apresenta as discussões em geral e conclusão, salientando a possibilidade de trabalhos futuros e as considerações finais levantadas.

2 FUNDAMENTAÇÃO TEÓRICA

A seguir será apresentado brevemente sobre alguns dos conceitos e tópicos referenciados ao longo do trabalho, tais como a taxonomia adotada, características gerais e diferenciais de insetos e métodos de classificação.

2.1 TAXONOMIA

A classificação taxonômica surgiu da necessidade que o ser humano tem de separar tudo em grupos, desde vegetais, fungos, bactérias até insetos. Tal necessidade é bem mais primitiva e onipresente do que aparenta ser, tendo em vista que o homem tem a característica inerente de querer separar objetos em grupos. Os livros estão separados por categoria em uma livraria, os produtos estão separados em prateleiras no supermercado, doces estão separados dos salgados em padarias e assim por diante.

Para que exista uma divisão, é preciso definir uma série de critérios para a tomada de decisões em agrupar os semelhantes. Esses critérios podem variar de indivíduo para indivíduo e, por isso, há a necessidade de uniformizar tal classificação para que estudos e pesquisas conduzidas nessa área possam ser usadas por todos os cientistas. A classificação universal adotada para classificar os seres vivos segue a escala da Figura 1.

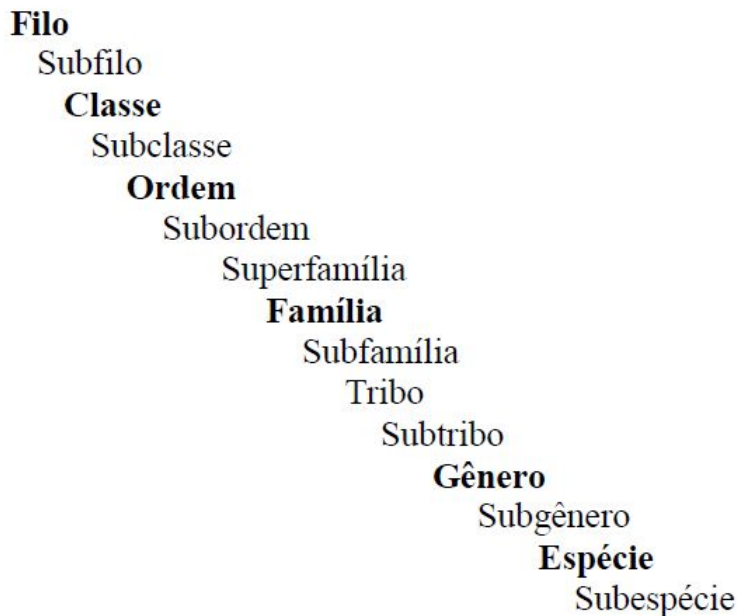


Figura 1 – Níveis taxonômicos usados na classificação dos seres vivos.

Fonte: Adaptado de (LEITE; SÁ, 2010).

2.2 INSETOS

Dentro da classificação apresentada na Figura 1, os insetos estão contidos dentro do reino Animal, pertencentes ao filo *Arthropoda* e da classe *Insecta* (LEITE; ASSOCIADO, 2011).

Segundo (LEITE; ASSOCIADO, 2011), existem muitas razões para estudar sobre os insetos, dentre as quais podem ser citadas:

- Os insetos existem desde o Período Devoniano.
- Habitam em quase todos os ambientes do planeta Terra, com exceção do mar aberto, onde são raramente encontrados. Mesmo na Antártica, foram encontrados cerca de 40 espécies dessa classe.
- São predominantes no reino animal, tendo representatividade de 80%, sem mencionar as espécies ainda desconhecidas (uma estimativa de 2 a 5 milhões).

Como citado em (LEITE; ASSOCIADO, 2011), os seres vivos pertencentes à classe *Insecta* possuem algumas características em comum pelas quais foram classificadas como sendo desse grupo em específico:

- 1) Corpo dividido em cabeça, tórax e abdome.
- 2) Um par de antenas.
- 3) Um par de mandíbulas.
- 4) Dois pares de maxilas (maxila e lábio).
- 5) Tórax apresentando três pares de patas e geralmente dois pares de asas.
- 6) Abdome desprovido de apêndices ambulatórios.
- 7) Abertura genital situada próxima à extremidade anal do corpo.
- 8) Desenvolvimento geralmente por metamorfose (completo ou incompleto).

A Figura 2 mostra as principais partes de um inseto genérico (nesse caso, uma formiga), e de acordo com a diversidade em cada uma dessas partes, bem como da presença ou não de algumas delas, conduzirão à uma classificação mais específica, agrupando-os em famílias e espécies.

2.2.1 Diferenciação quanto à cutícula

A cutícula do inseto, cuja função é determinar a forma do mesmo, garantir relativa impermeabilidade e ligamento de músculos, consiste de uma das principais características que podem diferenciar uma espécie da outra e até de uma parte para outra do inseto. Dentre as principais formações cuticulares podem ser citadas: cerdas ou macrotíquias, que têm origem celular; microtíquias (pelos fixos ou acúleos, sendo acelulares, tais como as asas de díptero); esporas, cuja origem é multicelular e ocorrem frequentemente em patas; e espinhos (acelular). (LEITE; ASSOCIADO, 2011).



Figura 2 – Composição de um inseto genérico.

Fonte: <https://www.sobiologia.com.br/conteudos/Reinos3/Artropodes4.php>

2.2.2 Diferenciação quanto à coloração dos insetos

Como consta em (LEITE; ASSOCIADO, 2011), a coloração dos insetos têm relativa importância na sua sobrevivência, pois podem desempenhar papel de ocultar o ser vivo do ambiente aonde vive (coloração ocultante), de disparar uma alerta de perigo para o próximo (coloração aposemática), ou de prevenção como ocorre em asas de mariposas que têm manchas similares a olhos grandes.

2.2.3 Fisiologia em geral

De forma geral, os artrópodes da classe *Insecta* possuem sistema circulatório aberto, isto é, o sangue flui livremente sobre os órgãos internos. Esse líquido que circula no corpo do animal é chamado de hemolinfa, pois possui função linfático também, além do sangue em si. No entanto, esse sangue não carrega oxigenação para os tecidos (LEITE; ASSOCIADO, 2011).

Quanto ao sistema respiratório, a troca gasosa da maioria dos insetos ocorre por meio de traquéias, podendo variar de tipo dependendo de onde ele vive. Em se tratando da excreção, processo para remover material não digestível, os túbulos de Malpighian realizam um papel fundamental nas excretas líquidas. Seu aparelho digestivo não permite a maioria dos insetos de digerir a celulose, por não produzirem celulase (LEITE; ASSOCIADO, 2011).

2.3 HEMIPTERA-PENTATOMIDAE

Segundo (FERNANDES, 2008), percevejos fitófagos, insetos da ordem *Hemiptera* e da família *Pentatomidae*, são largamente encontrados em colheitas de soja, arroz, café, entre outros, sendo pragas para as plantações, mas ao mesmo tempo sendo vestígios forenses quando é feita a rastreabilidade desses insetos.

A Figura 3 mostra um exemplo de *Pentatomidae* encontrado nas louveiras citadas anteriormente.



Figura 3 – Exemplo de um *Pentatomidae*.

Fonte: Projeto CAPES, 2017.

2.4 HYMENOPTERA-FORMICIDAE

Segundo (SUGUITURU et al., 2013), as formigas apresentam alta diversidade, dominância numérica e de biomassa em quase todos os habitats (exceto nos pólos), sendo por isso fontes importantes em estudos de biodiversidade. Ainda segundo (SUGUITURU et al., 2013), espécies da família *Formicidae* são sensíveis às mudanças ambientais, razão pela qual possuem um papel importante como bioindicadores.

A Figura 4 mostra um representante dessa família.



Figura 4 – Exemplo de um *Formicidae*.

Fonte: Projeto CAPES, 2017.

2.5 COLEOPTERA-BUPRESTIDAE

Segundo (PINHEIRO et al., 2017), os buprestídeos estão entre os principais besouros pragas de eucalipto no Brasil. Dentre seus principais feitiches, estão o ataque a folhas novas e roer galhos tenros. Dentre suas principais características, podem ser citadas: antenas curtas, corpo alongado e afilado na parte posterior, possui um tegumento espesso e duro, muitas vezes com cores brilho metálico.

A Figura 5 mostra um representante dessa família.



Figura 5 – Exemplo de um *Buprestidae*.

Fonte: Projeto CAPES, 2017.

2.6 RECONHECIMENTO DE PADRÕES

Antes de entrar em detalhes sobre o processo de reconhecimento de padrões e da sua vasta aplicabilidade, convém esclarecer algumas definições. Segundo (LOPES, 2012), **padrão** é uma entidade, objeto, processo ou evento, vagamente definido, que pode assumir um nome. **Classe** é um conjunto de padrões que possuem características em comum. Já **característica** ou **atributo** é um dado extraído de uma amostra por meio de medida e/ou processamento e, em geral, são organizadas na forma de um vetor de características.

O reconhecimento de padrões tem invadido muitos campos da rotina da sociedade. Segundo (NOGUEIRA et al., 2006), as aplicações são diversas, principalmente no ramo da segurança de redes e controle de acesso:

- **Reconhecimento de caracteres:** Consiste em extrair características de um conjunto de caracteres e separá-los em classes. Tecnologia largamente usada em reconhecimento automático de placas de carros.
- **Reconhecimento da assinatura manuscrita:** Aplicação usada em sistema de identificação de assinaturas de cheques.
- **Reconhecimento pelo ritmo da digitação:** Assim como a assinatura, o ritmo de digitação, o qual é identificado pela velocidade, espaço de tempo entre o acionamento de cada tecla bem como duração de pressão sobre a mesma e liberação, exhibe o mesmo fator neurofísico que pode ser utilizado na identificação única de um indivíduo, sendo por isso de difícil imitação por um usuário ilegítimo.
- **Reconhecimento pela impressão digital:** É um dos métodos de biometria mais usados em sistemas de segurança.
- **Reconhecimento pela geometria da mão:** Baseia-se no fato de que virtualmente não existem duas pessoas com mãos idênticas e de que o formato da mão não sofre mudanças significativas após certa idade.
- **Reconhecimento pela voz:** Processo de extração automática da informação linguística do sinal de voz.

- **Reconhecimento facial:** Faz-se uso de medidas da face como distâncias entre os olhos, nariz, queixo, boca e linha dos cabelos como meio de verificação.
- **Reconhecimento pela íris e retina**

Segundo (LOPES, 2012), o processo de reconhecimento de padrões requer uma sequência de etapas, como consta na Figura 6. Na etapa de pré-processamento, é feita uma normalização (transformação de escala numérica) para em seguida, extrair o dado (característica) de uma amostra por meio de medida e/ou processamento e fazer o treinamento e classificação.

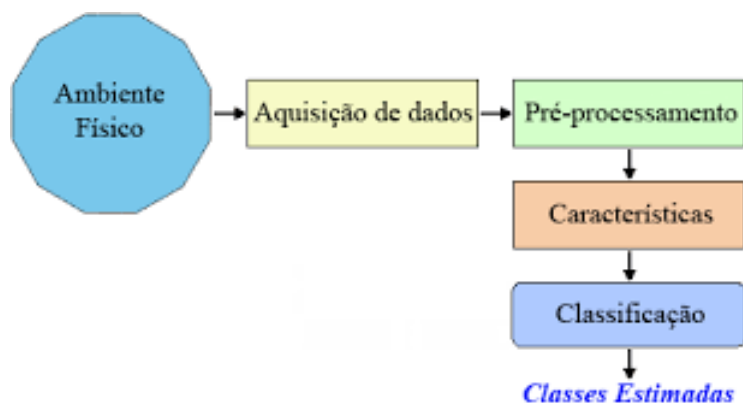


Figura 6 – Etapas de reconhecimento de padrões.

Fonte: (LOPES, 2012).

Numa análise matemática da descrição do reconhecimento de padrão, considera-se como um processo que identifica padrões classificáveis através das transformações matemáticas em quatro espaços: o espaço medido **M**, o espaço padrão **P**, o espaço específico ou característico **F** e o espaço classificado **C** (ANDREWS, 1972).

Segundo (ALMEIDA, 2007), o espaço medido é construído diretamente pelos dados medidos. Após formar o espaço **M**, passa-se para o espaço padrão **P**, que é essencialmente o que domina, seleciona e digitaliza os dados coletados. A dimensão desse espaço deve ser menor ou igual à dimensão do espaço **M**. A etapa seguinte é o espaço característico **F**, no qual é usado para encontrar uma transformação, com o objetivo de reduzir as dimensões do espaço padrão, mas mantendo suas características diferenciáveis para a classificação. Por fim, o algoritmo de classificação define a divisão do espaço característico em regiões separadas onde cada uma delas esteja associada a uma classe.

2.7 CLASSIFICAÇÃO DE IMAGENS

A classificação de imagens, segundo (FREITAS; PANCHER, 2011), é o processo pelo qual se extrai informações em imagens para o reconhecimento de padrões e objetos homogêneos que serão utilizados para um determinado fim, como a de identificação de insetos, tema desse trabalho. Para tal propósito, faz-se uso de classificadores para a identificação de objetos.

Os classificadores podem ser do tipo **pixel a pixel** ou **por regiões**. Como consta em (FREITAS; PANCHER, 2011), o caso do pixel a pixel usa de forma individual a informação

espectral de cada pixel na busca por regiões homogêneas. Já no caso dos classificadores por região, é utilizado, além da informação espectral de cada pixel, a relação espacial de vizinhança entre os pixels.

Segundo (MOREIRA, 2005), a classificação é subdividida em supervisionada e não-supervisionada. Na classificação supervisionada, selecionam-se representantes de cada uma das classes que se deseja classificar e, portanto, conhecem-se os grupos que estão sendo procurados. Já na classificação não-supervisionada, não se conhece nada a respeito dos padrões nem a quantidade de classes que estão sendo treinadas. Sua análise é feita por agrupamentos, ou *clusters*. A Figura 7 exemplifica a diferença entre os dois tipos de aprendizado.

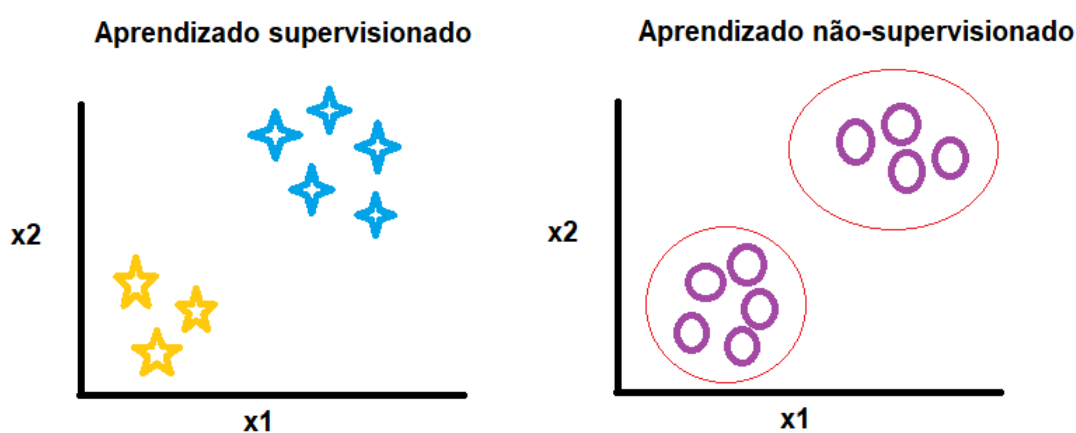


Figura 7 – Exemplo de classificação supervisionada e não-supervisionada.

Fonte: Autoria própria.

2.8 NOÇÃO DE DISTÂNCIA NA CLASSIFICAÇÃO

Tanto para aprendizado não-supervisionado, onde não há mecanismo de validação, quanto para o aprendizado supervisionado, a noção de distância é a base principal para a classificação, segundo (WEIHS; SZEPANNEK, 2009).

A distância Euclidiana é usada mais em classificação supervisionada, na qual cada pixel é incorporado a um agrupamento por meio da análise da medida de semelhança com o uso da distância Euclidiana. O classificador realiza então a tarefa de comparar essa distância do pixel em análise com a distância média de cada agrupamento. A decisão feita em cima dessa análise é incorporar o pixel novo ao agrupamento que apresentar a menor distância Euclidiana. A métrica Euclidiana, segundo (LOPES, 2012), atribui mesma importância a cada dimensão ou característica.

Uma outra métrica existente é a de Manhattan, que é uma simplificação da distância Euclidiana e usada em aprendizados não-supervisionados, na grande maioria das vezes. Essa distância é apenas a somatória das distâncias dadas em relação à quantidade de coordenadas apresentadas.

2.9 MÉTODO DE CLASSIFICAÇÃO KNN

O método dos K vizinhos mais próximos (KNN) funciona basicamente da seguinte forma: o classificador KNN procura por k elementos do conjunto de treinamento que estejam mais próximos de um elemento que não pertença ao conjunto de treinamento, ou seja, que tenham a menor distância, a fim de determinar a classe deste elemento desconhecido. Esse método possui apenas um parâmetro livre, a variável k , o qual é controlado pelo usuário a fim de obter uma melhor classificação (FARIA; MONTEIRO, 2015).

2.10 FILTRO DE MEDIANA

Segundo (RIBEIRO, 2006), os filtros de suavização são usados para a remoção de ruídos que geralmente são introduzidos durante o processo de conversão analógica-digital, excluindo assim ruídos de alta frequência. Existem técnicas de suavização conservativa e não-conservativa. Essas últimas eliminam detalhes como linhas finas e curvas agudas, gerando assim o efeito *blurring*, que é deixar a imagem com um borrado suave.

Os filtros de mediana reduzem o *blurring* e preserva a *edging* (a forma), apesar de causar danos nas linhas finas e curvas agudas. Seu funcionamento consiste em substituir o pixel em análise pelo valor médio de seus vizinhos, dependendo do tamanho da janela empregada.

3 METODOLOGIA

Este trabalho é pioneiro e preliminar na área da perícia criminal brasileira. Desta forma, a primeira etapa consiste em catalogar as imagens colhidas de insetos que servirão de base para o banco de dados. Através de um grupo de profissionais especializados em identificação de insetos regionais será executada a classificação por família e espécie, levantando características de verossimilhança entre cada grupo. Em seguida, tais imagens serão inseridas em um tipo de classificador na ferramenta LabVIEW (National Instruments), a fim de executar um pré-processamento e o consequente treinamento do sistema de reconhecimento. Vale ressaltar que a implementação dos filtros para suavização das figuras, bem como a eliminação de ruído e a remoção de objetos indesejáveis constituem de um passo importante antes de aplicar o classificador que irá executar as devidas identificações. Posteriormente, serão executados testes em ambiente real, com a coleta de microvestígios (fragmentos entomológicos completos e parciais) resultantes de ações policiais, padronizando a aquisição fotográfica, processamento comparativo no banco de dados e a consequente classificação de *matching* com o padrão. A Figura 8 mostra uma visão macroscópica do trabalho realizado.

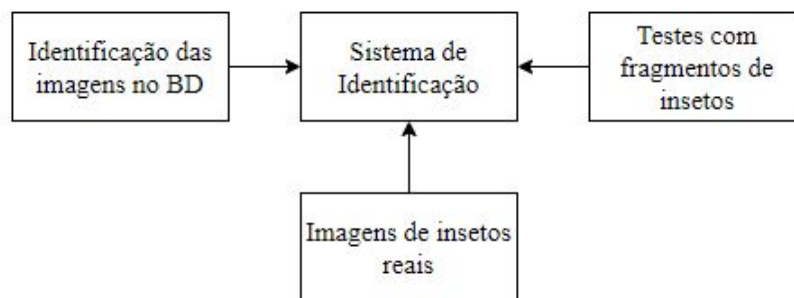


Figura 8 – Diagrama macroscópica do trabalho.

Fonte: Autoria própria.

Com a finalidade de prover soluções para o problema proposto, o trabalho foi dividido em etapas para que os resultados obtidos possam ser mostrados de forma cronológica. A seguir será discutido como foram as fases antes do desenvolvimento do sistema propriamente dito. Primeiramente, o trabalho iniciou-se com a coleta de informações sobre o assunto pertinente, com o objetivo de buscar nas referências teóricas uma base que pudesse viabilizar o projeto. No decorrer dessa etapa, estudaram-se alguns métodos alternativos, como reconhecimento por borda e diferentes ferramentas foram testadas.

Após avaliar que o software LabVIEW poderia satisfazer os objetivos do trabalho, fez-se então uma busca de referencial teórico sobre o funcionamento de processamento de imagem em LabVIEW. Iniciou-se, então, o aprendizado desde a leitura de uma imagem até a execução de operações complexas para extração de informações de interesse nesta ferramenta.

Nesse processo, consultaram-se diversos autores como (KLINGER, 2003), (RELF, 2003) e (BOVIK, 2009).

Em seguida, após o domínio da ferramenta e das técnicas de manipulação de imagem, começou-se a estruturar o classificador e a testar os métodos de treinamento, para então iniciar os testes e simulações com os filtros e para tentativa de melhores resultados.

3.1 VISÃO GERAL DO LABVIEW

O LabVIEW é um ambiente de desenvolvimento gráfico que usa a linguagem G como base de funcionamento. Segundo (INSTRUMENTS, 2015), essa ferramenta oferece ampla variedade de recursos e ferramentas, mas seu principal diferencial em relação às linguagens convencionais de programação é o fato de sua programação se dar por fluxo de dados. Em outras palavras, a estruturação de qualquer algoritmo nesse ambiente de desenvolvimento consiste em conectar ícones gráficos com fios em um diagrama de blocos. Apesar de ser baseada em representação gráfica, a linguagem G possui os mesmos conceitos de programação encontrados nas outras linguagens, tais como estruturas de loop, tratamento de eventos e programação orientada a objeto. A interação com o usuário ocorre numa outra janela de inicialização, o painel frontal (“*Front Panel*”), no qual encontram-se controles e indicadores. A Figura 9 mostra as duas janelas de inicialização, ambientes de trabalho do programador.

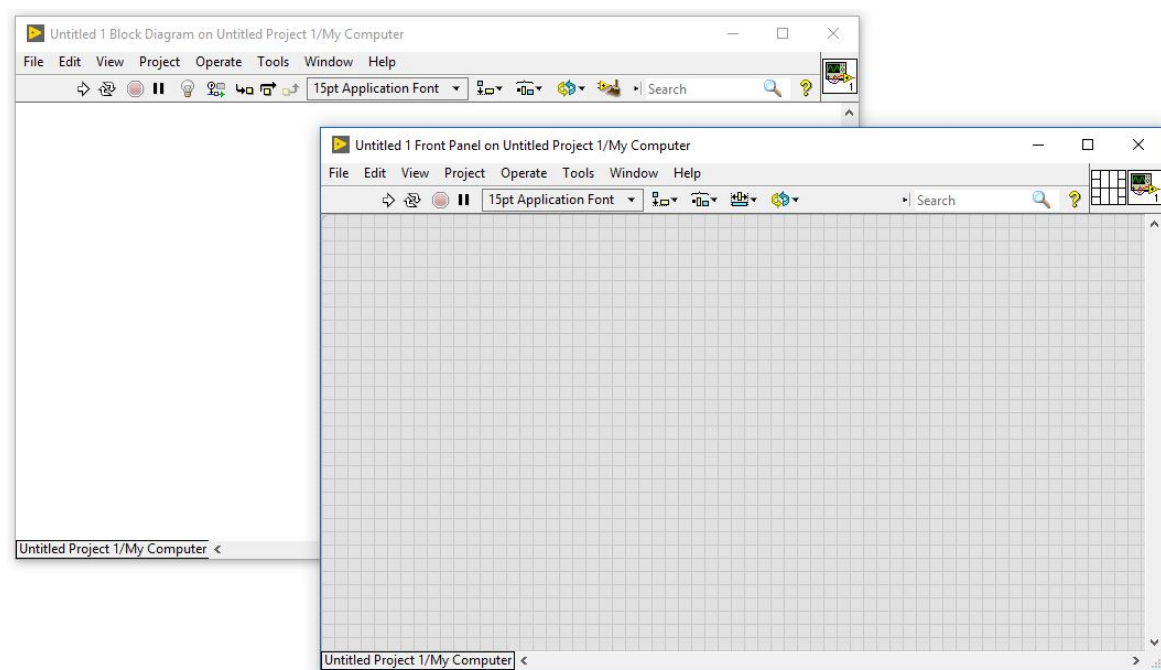


Figura 9 – Janelas de inicialização do LabVIEW: diagrama de blocos e painel frontal.

Fonte: Autoria própria.

3.2 MOTIVO DA ESCOLHA DO LABVIEW

Dentre os benefícios que o LabVIEW apresenta, estão, segundo (INSTRUMENTS, 2015):

- Programação gráfica intuitiva
- Ferramentas interativas de depuração
- Paralelismo automático
- Alto desempenho
- Recurso de combinação da linguagem G com outras linguagens

Além das vantagens citadas, o motivo principal pela opção do LabVIEW perante as demais ferramentas é o fato de possuir como foco principal o dado em si, e não o processo de manipulação dele. A consequência direta dessa vantagem está na simplificação do problema, eliminando grande parte das complexidades e recursos limitados encontrados na programação.

Por apresentar uma interface bem amigável com o usuário, conclui-se que o LabVIEW é a escolha adequada para cumprir os propósitos desse trabalho.

3.3 AQUISIÇÃO DE IMAGENS

As imagens dos insetos foram adquiridos a partir da cessão temporária de uma coleção particular de um entomólogo. Através de um procedimento cuidadoso, foram executadas fotografias em uma escala fixa e com iluminação padronizada. Para a aquisição das imagens, foram utilizados os seguintes materiais:

- 1 Aparelho celular iPhone 7 Plus
- 4 Lâmpadas LED 15W com soquete, fios e tomada
- 1 Base de isopor + foam
- 1 Metro de manta térmica reflexiva
- 1 Armação de PVC
- 1 Tripé reversível ajustável com suporte para celular

A Figura 10 mostra, sob pontos de vista diferentes, como se deu o processo de aquisição de imagens a partir da lista de material apresentada acima.



Figura 10 – Processo de aquisição de imagens, sob ângulos de visualização diferentes.

Fonte: Projeto CAPES, 2017

4 DESENVOLVIMENTO

O presente capítulo consiste em descrever o método usado na construção do classificador, o sistema de identificação criado no LabVIEW, bem como um fluxograma da programação feita em LabVIEW.

4.1 CONFIGURAÇÕES DO CLASSIFICADOR

Primeiramente, antes de treinar as imagens e classificá-las, é preciso configurar o classificador, determinando os padrões que devem ser seguidos para cada amostra de treinamento. O algoritmo das configurações pode ser melhor visualizado na Figura 11.

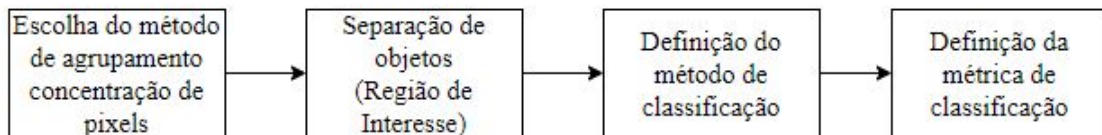


Figura 11 – Diagrama de configurações do classificador.

Fonte: Autoria própria.

Para a etapa de pré-processamento, fase na qual estabelece-se os métodos de agrupamento dos pixels que constituirão a base de treinamento, foi usado o método de **clustering**, ou “nuvens”, o qual consiste em juntar regiões de alta densidade de curvas de contorno dos pixels em uma imagem. Após essa identificação das nuvens de dados, segundo (CRÓSTA, 1992) deve-se tomar uma decisão quanto a quais concentrações devem ser tratados como grupos separados, isto é, definir um limiar. No LabVIEW esse método é automático, não havendo necessidade de estabelecer um valor mínimo e/ou máximo para essa tomada de decisão.

Como as imagens de treinamento já foram tratadas, o ruído é quase insignificante. O fundo é branco e apenas possui o inseto como objeto central, então a tomada de decisão da separação de grupos consiste em procurar por objetos escuros, separando assim o inseto do resto da figura. A Figura 12 mostra algumas imagens tratadas para um representante de cada classe. Pode-se notar que o objeto está bem visível e que interferências de ruídos são quase nulas. Ou seja, a procura pela região de interesse (em inglês, ROI - Region of Interest) limita-se em achar objetos escuros.

O método de classificação usado foi o de k-vizinhos mais próximos (KNN), cuja ideia será discutida e apresentada a seguir. Segundo (BEZERRA, 2006), a fim de estimar a classe de um novo padrão X de entrada, esse algoritmo calcula os k-vizinhos mais próximos a X e a classe que aparecer com maior frequência será a classe de X . O método de KNN é uma versão adaptada do método de NN (vizinho mais próximo), o qual usa apenas o vizinhos mais



Figura 12 – Exemplo de amostras de imagens tratadas dos insetos. Respectivamente, da esquerda para direita: *Buprestidae*, *Formicidae* e *Pentatomidae*.

Fonte: Projeto CAPES, 2017.

próximo como referência para classificar o novo padrão de entrada. A vantagem do algoritmo KNN está no fato deste trazer mais exatidão, visto que leva em conta os padrões ao redor de X , trazendo um desempenho melhor de classificação. Como consta em (BEZERRA, 2006), é sugerido que sejam testados valores diferentes de k para obter uma classificação mais refinada.

Após isso, seguindo o algoritmo da Figura 11, o último passo consiste em estabelecer uma métrica. Dentre as métricas disponíveis no LabVIEW (distância euclidiana, distância da soma e máxima distância), escolheu-se a da soma, também conhecida por métrica Manhattan, a qual propõe que a distância entre dois pontos seja dada como a soma das diferenças entre x e y , como mostra a Equação 1.

$$d = |x_1 - x_2| + |y_1 - y_2| \quad (1)$$

Por fim, depois de estabelecidos os padrões do classificador, o próximo passo é o treinamento propriamente dito. O treinamento foi feito apenas com imagens tratadas do diretório do banco de dados do Projeto CAPES (2017) e os testes executados, tanto para verificar o classificador, como para constituir as imagens de entrada do sistema de identificação, foram feitos com base nas outras imagens (as não tratadas) pertencentes à mesma base do Projeto CAPES (2017).

4.2 SISTEMA DE IDENTIFICAÇÃO

O sistema de identificação construído em LabVIEW segue o algoritmo estruturado na Figura 13. Primeiramente, a figura de entrada do sistema deve ser convertida em uma imagem com representação RGB de 8 bits. Em seguida, aplicam-se filtros na imagem, a fim de prepará-la para o processamento e facilidade na extração das informações de interesse. Com a imagem agora filtrada, efetuam-se operações morfológicas para melhorar a rotulação de objetos

posteriormente. Nessa fase, evita-se, por exemplo, do sistema considerar a cabeça do inseto como um objeto e o abdômen como outro. A etapa de rotular objetos consiste em numerar os objetos encontrados na imagem de entrada, identificando assim a quantidade de insetos da imagem original. Pode haver ainda objetos indesejados na imagem e que o sistema tenha considerado como objeto de interesse. Para contornar isso, aplica-se um filtro de partículas com a finalidade de remover objetos indesejáveis. Por fim, a imagem é equalizada, processo que aumenta o contraste da mesma, para então ser identificada com o uso do classificador já treinado e configurado como discutido na secção anterior.

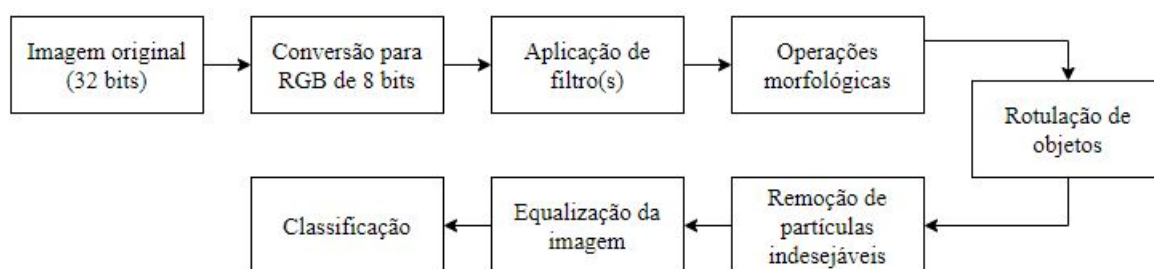


Figura 13 – Diagrama do algoritmo estruturado para desenvolvimento do sistema de identificação entomológica.

Fonte: Autoria própria.

4.2.1 Algoritmo em LabVIEW

A seguir, será comentado como foi executada a estruturação do algoritmo em LabVIEW. Por ser uma linguagem do tipo *dataflow*, então não há como descrever o código completo aqui, mas serão inseridas amostras e, além disso, cada uma dessas partes será comentada, permitindo fácil identificação da função de cada bloco.

Na Figura 14, são mostrados os blocos para a criação da imagem, leitura da imagem e conversão da mesma. Nota-se que, ao fazer a conversão de 32 bits para 8 bits, reduz-se a representatividade da imagem. Isto é, a imagem original possuía $2^{32} = 4.294.967.296$ escalas de cinza para representar cada uma das cores primárias RGB, enquanto na imagem que será processada tem-se, no máximo, apenas $2^8 = 256$ níveis de cinza. Essa conversão modifica a imagem, visto que o pixel com valor mais alto na imagem original é modificado para o valor teto dessa nova escala e, analogamente, o mesmo ocorre com o pixel de valor mais baixo. Após a conversão, o bloco seguinte trata de aplicar um filtro do tipo "mediana de máscara" 5×5 .

A Figura 15 mostra a continuação do código sequencial tratada na primeira parte, conforme a Figura 14. A imagem de saída do filtro da mediana perpassa agora por um processo morfológico de dilatação, para poder agrupar às partes do inseto que ficaram um tanto separadas devido aos efeitos dos blocos anteriores. São feitas 3 iterações de dilatação para que as partes, antes isoladas, agrupem-se formando um único corpo do inseto.

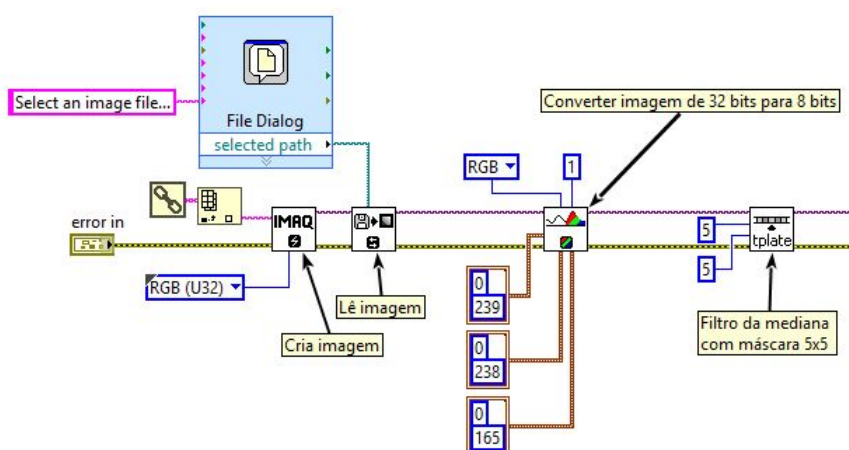


Figura 14 – Parte 1 do código em LabVIEW.

Fonte: Autoria própria.

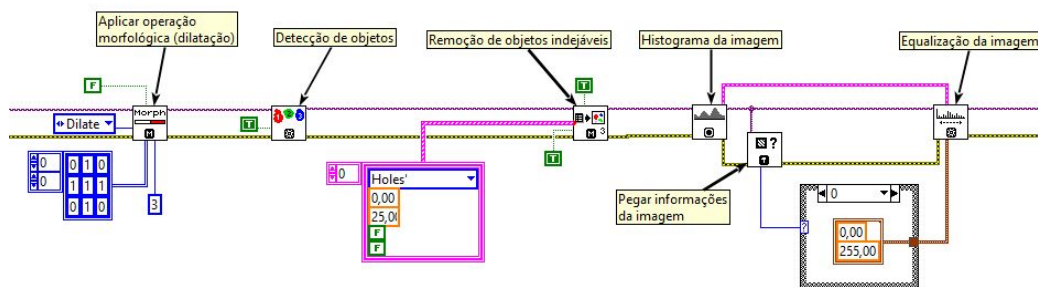


Figura 15 – Parte 2 do código em LabVIEW.

Fonte: Autoria própria.

O bloco seguinte trata da detecção da quantidade de objetos presentes na imagem, para que na etapa seguinte seja feita a remoção dos objetos indesejáveis. Por objetos indesejáveis, determinou-se que seriam orifícios de perímetro entre 0 e 25 unidades de comprimento. Tal intervalo foi especificado manualmente após várias análises de imagens. Observou-se que a média dos orifícios encontrados se encontrava em torno de 50, podendo variar muito dependendo da imagem de entrada. Sendo assim, os testes executados mostraram que um perímetro de valor máximo igual a 25 poderia ser considerado como objeto indesejável.

Após isso, o passo seguinte é equalizar a imagem a fim de aumentar seu contraste. No entanto, para poder realocar os valores dos pixels, é preciso antes passar pelo bloco de calcular o histograma da imagem, como descrito na Figura 15.

A última parte, mostrado na Figura 16, consiste na leitura do classificador e execução da identificação dos objetos rotulados, tendo como referência o classificador já treinado.

Como no LabVIEW existem duas partes correspondentes a um mesmo VI, o diagrama de blocos e o painel frontal, então na Figura 17 é mostrada este último, sendo também uma interface com usuário. Nessa tela, o usuário pode acompanhar onde aconteceu um determinado erro na compilação e execução do código, podendo controlar variáveis e acompanhar os

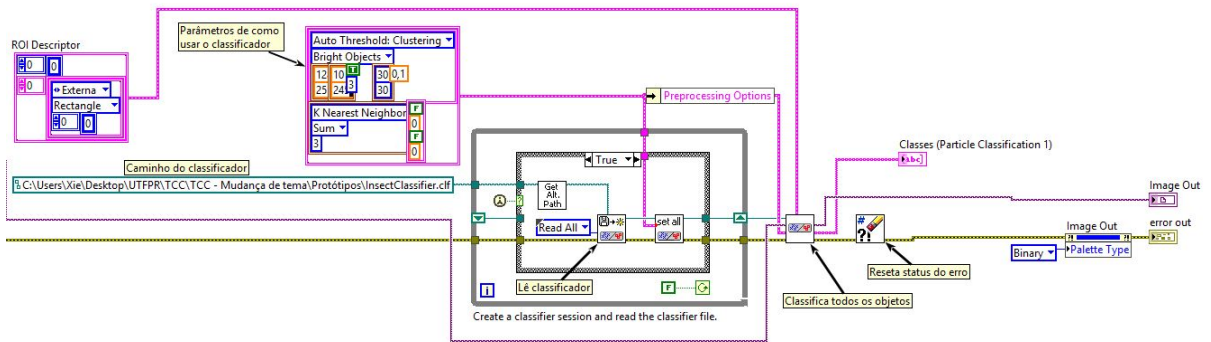


Figura 16 – Parte 3 do código em LabVIEW.

Fonte: Autoria própria.

indicadores que, nesse caso em específico, são a imagem de saída e as classes atribuídas aos insetos localizados na figura de entrada.

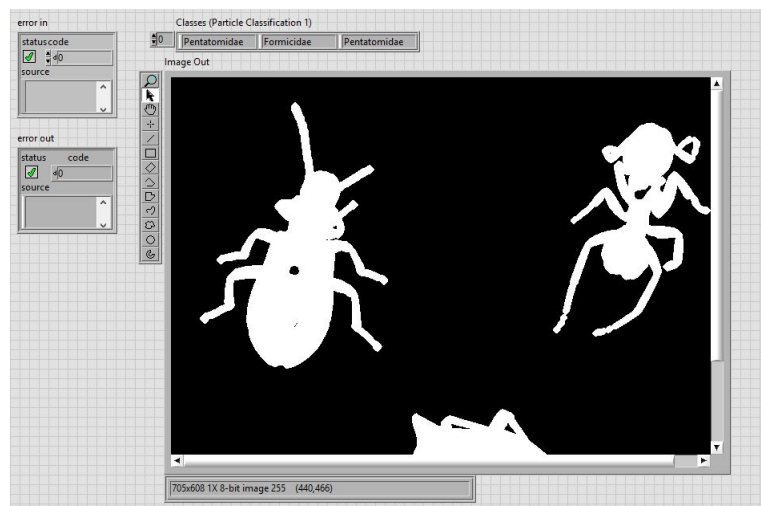


Figura 17 – Painel Frontal correspondente ao diagrama de blocos construído no LabVIEW.

Fonte: Autoria própria.

5 TESTES E RESULTADOS

Esse capítulo trata dos testes realizados, bem como resultados obtidos com o treinamento, classificação e análise das imagens de entrada e saída do sistema de identificação entomológico.

5.1 RESULTADOS DO CLASSIFICADOR

Após estabelecer as configurações do classificador, tem-se o resultado do treinamento mostrado na Tabela 1. Vale ressaltar que o parâmetro k , no método de k -vizinhos mais próximos, foi configurado como 10 para que estes resultados sejam atingidos.

Tabela 1 – Resultado do treinamento usando o método de k -vizinhos mais próximos.

	Buprestidae	Formicidae	Pentatomidae	Total
Buprestidae	126	0	12	138
Formicidae	8	108	2	118
Pentatomidae	4	2	185	191

Fonte: Aatoria Própria.

A Tabela 2 mostra o resultado acerca das informações de cada classe treinada. Dentre essas informações, estão a quantidade de amostras de cada classe, o desvio padrão em relação à média de todas as amostras de cada grupo, a acurácia, que é o grau de certeza de uma amostra ter sido classificada dentro da classe a qual pertence, e valor preditivo, que é a probabilidade que uma amostra classificada em uma determinada classe tem de pertencer a essa classe.

Tabela 2 – Informações das classes treinadas.

Classe	Amostras	Desvio Padrão	Acurácia	Valor Preditivo
Buprestidae	138	0.193249	0.913043	0.913043
Formicidae	118	0.126127	0.981818	0.915254
Pentatomidae	191	0.204148	0.929648	0.968586

Fonte: Aatoria Própria.

A Tabela 3 mostra a distância média entre as classes, isto é, o quão diferente uma é da outra. Vale observar que há muita semelhança entre a família *Buprestidae* com a família *Pentatomidae*, pois a distância entre elas é de apenas 0.45, razão pela qual o classificador poderá atribuir erroneamente uma classe aos insetos de entrada das duas famílias.

Pode-se observar que o classificador KNN não obteve 100% de exatidão na classificação dos insetos, tendo em vista que das 138 imagens de *Buprestidae*, atribuíram-se 12 como se

Tabela 3 – Distância entre as classes treinadas.

	Buprestidae	Formicidae	Pentatomidae
Buprestidae	0.000000	0.708142	0.450540
Formicidae	0.708142	0.000000	0.809340
Pentatomidae	0.450540	0.809340	0.000000

Fonte: Aatoria Própria.

fossem da família *Pentatomidae*, devido à semelhança entre as duas famílias. Do mesmo modo, os insetos da família *Formicidae* resultaram em uma acurácia de 91.53%, enquanto que os da família *Pentatomidae* tiveram apenas 6 imagens classificadas erradas. Ao considerar uma análise macro desse classificador, sua acurácia foi de 94%.

Devido à sua não-exatidão, os testes executados usando esse classificador resultaram em alguns desvios, como era de se esperar. O erro mais comum seria em confundir um *Buprestidae* com um *Pentatomidae*, dependendo da posição na qual se encontra.

5.1.1 Teste 1

Como exemplo, a Figura 18 mostra a classificação de insetos de uma imagem que contém representantes das 3 famílias. É possível notar que o primeiro objeto foi identificado como *Pentatomidae*, enquanto que sua correspondência original seria de um *Buprestidae*. Para fins de comparação, a Figura 19 é a imagem original de entrada no sistema de identificação cujo resultado é a Figura 18.

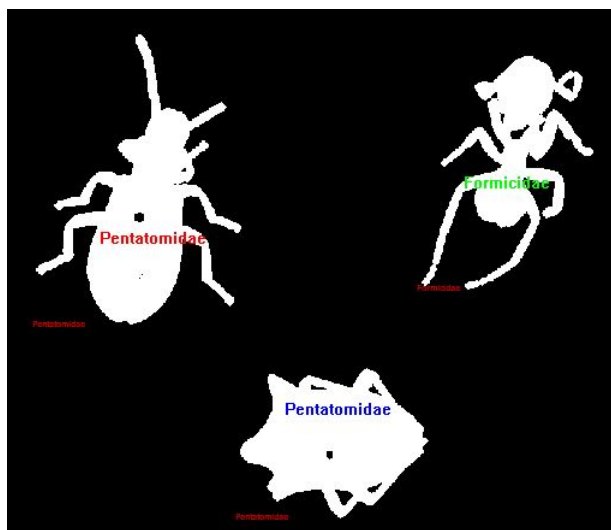


Figura 18 – Resultado usando o classificador construído no LabVIEW.

Fonte: Autoria própria.



Figura 19 – Imagem original de entrada para teste do sistema de identificação.

Fonte: Adaptado de (PROJETO CAPES, 2017).

5.1.2 Teste 2

Como mostra a Tabela 1, o inverso também pode acontecer, isto é, o classificador confundir um *Pentatomidae* com um *Buprestidae*, que é o caso do exemplo da Figura 20. Nesse resultado, o *Formicidae* foi identificado como um *Pentatomidae*, efeito passível de ocorrer, visto que na Tabela 1 aconteceram 2 dos 118 insetos dessa família classificados como *Pentatomidae*. A imagem original de entrada referente a esse resultado é a Figura 21.

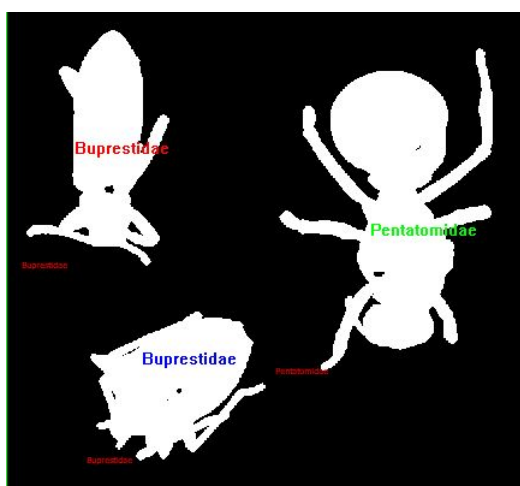


Figura 20 – Resultado usando o classificador construído no LabVIEW.

Fonte: Autoria própria.



Figura 21 – Imagem original de entrada para testes no sistema de identificação.

Fonte: Adaptado de (PROJETO CAPES, 2017).

Por fim, a Figura 22 mostra o resultado de um teste após a passagem de cada um dos blocos do diagrama descrito na Figura 13.

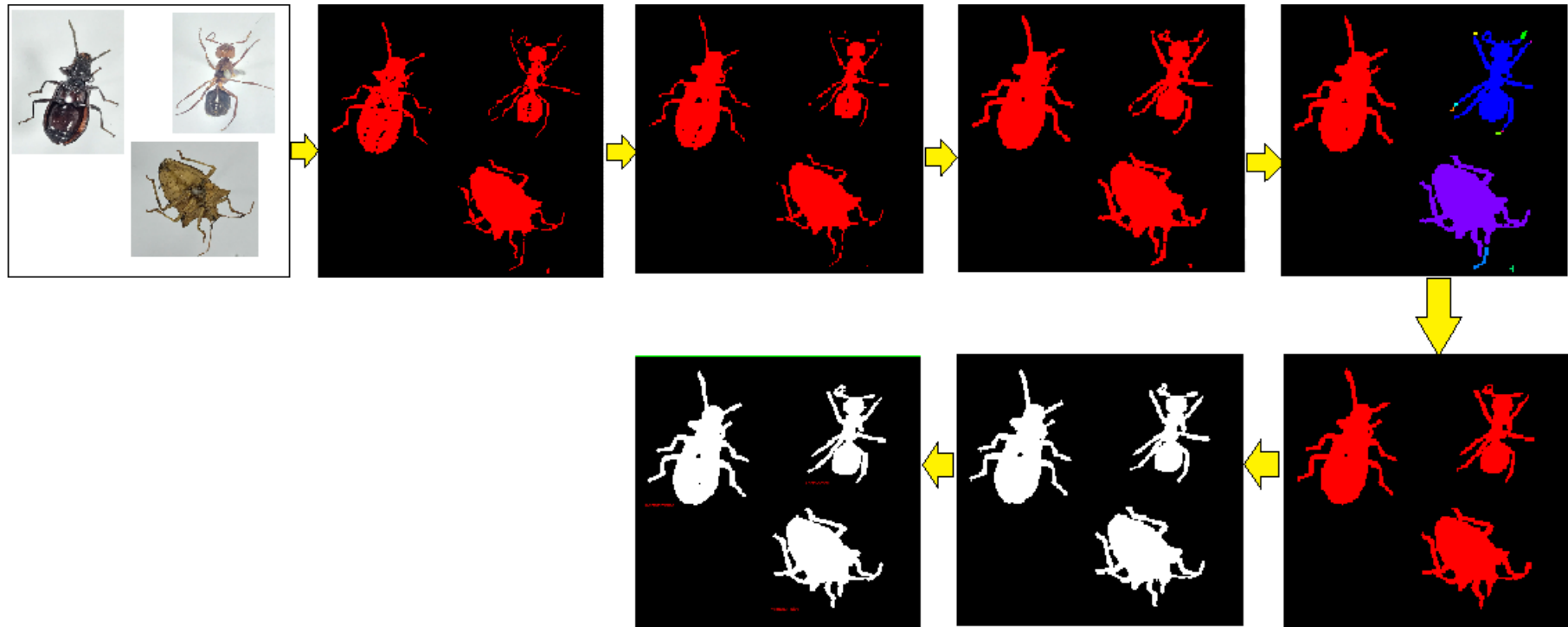


Figura 22 – Resultado da saída de cada bloco do algoritmo estruturado na Figura 13. Da esquerda para a direita: (1) Imagem original; (2) Conversão para 8 bits; (3) Filtro mediana; (4) Dilatação; (5) Rotular objetos; (6) Remoção de partículas indesejáveis; (7) Equalização; (8) Classificação.

Fonte: Autoria própria.

5.2 RESULTADOS UTILIZANDO FRAGMENTOS DE INSETOS

O classificador obteve êxito, também, ao utilizar como objeto de testes alguns fragmentos de *Pentatomidae* (vide Figura 23), como pode ser verificado na atribuição de classe dada de forma correta na Tabela 4.

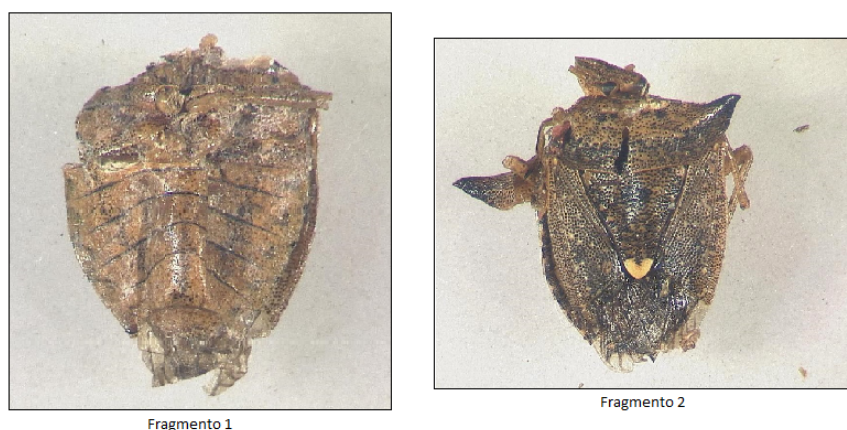


Figura 23 – Fragmentos de *Pentatomidae*.

Fonte: Projeto CAPES, 2017.

Tabela 4 – Resultados do classificador quanto a fragmentos de *Pentatomidae* na entrada.

	Fragmento 1	Fragmento 2
Classe atribuída	<i>Pentatomidae</i>	<i>Pentatomidae</i>
Classificação	1000	1000
Identificação	393.1	462.9
Distância de <i>Pentatomidae</i>	0.5	0.4
Distância de <i>Buprestidae</i>	0.9	0.7
Distância de <i>Formicidae</i>	1.3	1.0

Fonte: Autoria Própria.

Na Tabela 4, foi atribuída uma pontuação de classificação entre 0 a 1000, referente ao grau de certeza do fragmento em questão pertencer a uma classe ao invés de outras. Já a pontuação de identificação, valor entre 0 e 1000, mostra o grau de semelhança entre a amostra sob análise com membros da classe à qual foi atribuída.

Percebe-se que a classificação obteve o resultado esperado, apesar da baixo nível de certeza na identificação, justamente pelo fato dos insetos da Figura 23 serem apenas fragmentos de pentatomídeos, comprometendo assim a sua identificação perante o restante de sua classe. No entanto, o processo de classificação se encontra correto, pois a distância destes fragmentos em relação às duas outras classes era relativamente maior em relação à da *Pentatomidae*.

5.3 TESTES COM OUTROS CENÁRIOS DE CLASSIFICAÇÃO

Os resultados apresentados e discutidos anteriormente foram construídos a partir do estabelecimento de alguns parâmetros do classificador, como o uso do método KNN com $k = 10$ e a métrica da soma (ou Manhattan, como é comumente conhecido). A seguir, serão considerados outros cenários de como configurar o classificador.

5.3.1 Quanto ao parâmetro k

A Tabela 5 mostra alguns testes onde foi variado o parâmetro k para o mesmo método de KNN, porém preservando a métrica Manhattan utilizada.

Tabela 5 – Resultado do classificador, para a métrica da soma.

	$k = 3$	$k = 10$	$k = 15$
Acurácia do classificador	0.99	0.94	0.93
Acerto Teste 1	2/3	2/3	2/3
Acerto Teste 2	1/3	1/3	1/3
Fragmento 1	<i>Pentatomidae</i>	<i>Pentatomidae</i>	<i>Pentatomidae</i>
Fragmento 2	<i>Pentatomidae</i>	<i>Pentatomidae</i>	<i>Pentatomidae</i>

Fonte: Autoria Própria.

Os resultados apresentados se encontram na Tabela 5, onde permitem determinar que o parâmetro k não altera os acertos dos testes utilizados na classificação, Figuras 18 e 20, respectivamente. Apesar de obter uma acurácia melhor do classificador quando $k = 3$, observou-se que essa variável não possui peso significativo no momento de identificar os segmentos de *Pentatomidae*, como mostra a Tabela 5, em que tal resultado permaneceu correto e inalterado para os três valores de k .

6 CONCLUSÃO

Conclui-se que a metodologia proposta no LabVIEW é uma solução promissora e que pode atender de forma satisfatória os requisitos funcionais de um sistema de identificação entomológica, a partir do momento que se incrementa o número de amostras no banco de dados. Esta ferramenta oferece uma interface amigável e de fácil manuseio, além da programação utilizando diagramas de blocos. Isto permite que profissionais de diversas áreas possam interagir no sistema, enriquecendo-o através da inserção de mais amostras, executando melhorias nos filtros de forma a otimizar a extração de objetos, ou mesmo executando exaustivos testes apenas.

Para o desenvolvimento desse sistema, configurou-se primeiro o classificador a fim de treinar as imagens tratadas do Projeto CAPES, 2017. Em seguida, o programa desenvolvido em LabVIEW seguiu uma sequência de fluxo de processamento, iniciando com a abertura de uma imagem, leitura, tratamento com o uso de filtros, até a extração de objetos para poder classificá-los. Vale ressaltar que os testes de verificação dos filtros aplicados foram de suma importância para que se pudesse escolher a melhor combinação, a fim de realçar as informações desejáveis para a identificação.

Os resultados apresentados mostraram que tais filtros e métodos foram adequados, culminado em uma classificação bastante exata. Há erros de interpretação entre um *Pentatomidae* com um *Buprestidae* e vice-versa, devido à proximidade entre as duas famílias, como já apresentado na Tabela 3. Um dos motivos pelos quais o sistema pode resultar em erros de identificação é durante a etapa de remoção de partículas indesejáveis. O algoritmo pode remover, por exemplo, parte das antenas do pentatomídeo, deixando-o com um semblante mais parecido a um buprestídeo.

O sistema também foi testado utilizando fragmentos de insetos coletados em locais de crime atendidos pela Polícia Federal, conforme os dois casos apresentados na Tabela 4. A classificação de classe dada como 100% pertencente à família *Pentatomidae*, apesar da identificação (o quão semelhante é perante os outros membros do grupo para o qual foi atribuído) ter resultado em uma pontuação baixa, motivo justificável pelo fato de se tratar apenas de fragmentos, não constituindo assim um inseto completo a ponto do classificador extrair todas as características inerentes a um inseto pertencente a tal grupo.

Tendo em vista que o método utilizado para o classificador foi o de KNN, isto é, leva em consideração os k vizinhos mais próximos do ponto em análise, foram executados experimentos averiguando o impacto do parâmetro k nos testes e resultados obtidos. Observou-se que a alteração do valor k não interfere na identificação dos insetos, como pode ser visto na Tabela 5, apesar de resultar em um classificador com maior acurácia para o cenário de $k = 3$.

6.1 TRABALHOS FUTUROS

Considera-se imprescindível a adição de mais amostras no banco de dados, a fim de compor uma base treinada mais rica e refinada para cada uma das ordens e famílias de insetos. Espera-se com isso obter uma diferença mais significativa entre as famílias *Buprestidae* e *Pentatomidae* e, conseqüentemente, uma maior acurácia no resultado do classificador.

Além disso, outro ponto a se investir seria em uma versão melhorada dos filtros, tornando-os mais sensíveis a determinadas partes do corpo de um inseto, para que não haja remoção de parcelas do mesmo. Espera-se que essa melhoria possa trazer mais fidelidade ao objeto detectado em relação à imagem original de entrada.

Por fim, há ainda uma sugestão de implementar mais métodos de classificação e treinamento para melhor averiguar aquele que seria mais conveniente para o problema proposto.

Pode-se, ainda, transformar esse trabalho, cujo teor é mais teórico e exploratório, em um software para reconhecimento de microvestígios em locais de crime, para utilização contínua das autoridades policiais, auxiliando na instrução da investigação criminal, ou mesmo outros profissionais interessados em obter mais informações sobre um determinado inseto de forma rápida, permitindo-os a seguir com seus trabalhos a partir da classificação fornecida pelo sistema.

6.2 CONSIDERAÇÕES FINAIS

Observou-se que o sistema de identificação desenvolvido é expansível, tanto em número de amostras por classe, como a quantidade de classes em si, cumprindo um dos principais requisitos do projeto. Outro ponto a considerar é a acurácia obtida na classificação dos insetos e fragmentos destes, mesmo com uma quantidade relativamente baixa de amostras para o treinamento do classificador.

Com esse trabalho, provou-se também que a metodologia proposta, com o uso do LabVIEW, tem alto potencial de ser a solução certa para problemas em geral que envolvam reconhecimento de padrões e identificação de objetos.

Referências

- ALMEIDA, F. R. do V. **Análises Estatísticas e Reconhecimento de Padrão Aplicados em Diagnósticos de Defeitos em Rolamentos Através da Análise de Vibração**. Tese (Doutorado) — UNIVERSIDADE FEDERAL DE ITAJUBÁ, 2007. Citado na página 21.
- ANDREWS, H. C. **Introduction to mathematical techniques in pattern recognition**. [S.l.]: Wiley-Interscience New York, 1972. v. 152. Citado na página 21.
- BEZERRA, M. E. R. Métodos baseados na regra do vizinho mais próximo para reconhecimento de imagens. 2006. Citado 2 vezes nas páginas 28 e 29.
- BOVIK, A. C. **The essential guide to image processing**. [S.l.]: Academic Press, 2009. Citado na página 25.
- CRÓSTA, A. **Processamento digital de imagens de sensoriamento remoto. São Paulo, 173P**. Tese (Doutorado) — Tese-Doutorado)-Instituto de Geociências, Universidade de Campinas, 1992. Citado na página 28.
- FARIA, M. M.; MONTEIRO, A. M. Investigação sobre técnicas de detecção de intrusões em redes de computadores com base nos algoritmos knn e k-means. 2015. Citado na página 23.
- FERNANDES, A. P. Análise comparativa da espermiogênese normal e diapáusica em percevejos fitófagos (hemiptera, pentatomidae)-revisão. **Universitas: Ciências da Saúde**, v. 1, n. 1, p. 87–95, 2008. Citado na página 18.
- FREITAS, M. I. C. de; PANCHER, A. M. **CLASSIFICAÇÃO DE IMAGENS**. 2011. Disponível em: <http://www.rc.unesp.br/igce/planejamento/download/isabel/sist_inf_geografica/Aula2011/classificacao_imagens.pdf>. Acesso em: 27 Novembro 2018. Citado na página 21.
- INSTRUMENTS, N. **Os benefícios da programação gráfica do LabVIEW**. 2015. Disponível em: <<http://www.ni.com/white-paper/14556/pt/>>. Acesso em: 29 Novembro 2018. Citado 2 vezes nas páginas 25 e 26.
- KLINGER, T. **Image processing with LabVIEW and IMAQ Vision**. [S.l.]: Prentice Hall Professional, 2003. Citado na página 25.
- LEITE, G. L. D.; ASSOCIADO, I. Entomologia básica. **Tórax**, v. 17, p. 20, 2011. Citado 2 vezes nas páginas 17 e 18.
- LEITE, G. L. D.; SÁ, V. G. M. de. Apostila: Taxonomia, nomenclatura e identificação de espécies. **Universidade Federal de Minas Gerais, Instituto de Ciências Agrárias**, p. 50, 2010. Citado na página 16.
- LOPES, F. M. Introdução ao reconhecimento de padrões e aplicações em problemas de bioinformática. **Grupo de Pesquisa em Bioinformática e Reconhecimento de Padrões da UTFPR**, 2012. Citado 3 vezes nas páginas 20, 21 e 22.
- MOREIRA, M. A. **Fundamentos do sensoriamento remoto e metodologias de aplicação**. [S.l.]: UFV, 2005. Citado na página 22.

- NOGUEIRA, A. et al. Um overview sobre reconhecimento de padrões. **II Simpósio de Excelência em Gestão e Tecnologia, Rio de Janeiro**, 2006. Citado na página 20.
- PINHEIRO, A. et al. Ocorrência de buprestídeos (coleoptera) em vegetação de cerrado, no município de santo antônio do leverger, mt. 2017. Citado na página 19.
- RELF, C. G. **Image acquisition and processing with LabVIEW**. [S.l.]: CRC press, 2003. Citado na página 25.
- RIBEIRO, B. **Suavização de Imagens - Image Smoothing**. 2006. Disponível em: <<http://www.ic.uff.br/~aconci/suavizacao.pdf>>. Acesso em: 30 Novembro 2018. Citado na página 23.
- SUGUITURU, S. S. et al. Diversidade e riqueza de formigas (hymenoptera: Formicidae) em remanescentes de mata atlântica na bacia hidrográfica do alto tietê, sp. **Biota Neotropica**, Instituto Virtual da Biodiversidade (BIOTA/FAPESP), p. 141–152, 2013. Citado na página 19.
- WEIHS, C.; SZEPANNEK, G. Distances in classification. In: SPRINGER. **Industrial Conference on Data Mining**. [S.l.], 2009. p. 1–12. Citado na página 22.