



---

*Research article*

## Investigation of ant cuticle dataset using image texture analysis

Noah Gardner<sup>1</sup>, John Paul Hellenbrand<sup>2</sup>, Anthony Phan<sup>1</sup>, Haige Zhu<sup>1</sup>, Zhiling Long<sup>4</sup>,  
Min Wang<sup>3</sup>, Clint A. Penick<sup>2</sup> and Chih-Cheng Hung<sup>1,\*</sup>

<sup>1</sup> Laboratory of Machine Vision and Security Research, College of Computing and Software Engineering, Kennesaw State University, Marietta GA, USA

<sup>2</sup> College of Science and Mathematics, Kennesaw State University, Kennesaw, GA, USA

<sup>3</sup> Department of Mathematics, Kennesaw State University, Kennesaw, GA, USA

<sup>4</sup> Department of Informatics and Mathematics, Mercer University, Atlanta, GA, USA

\* **Correspondence:** Email: chung1@kennesaw.edu.

Academic Editor: Pasi Fränti

**Abstract:** Ant cuticle texture presumably provides some type of function, and therefore is useful to research for ecological applications and bioinspired designs. In this study, we employ statistical image texture analysis and deep machine learning methods to classify similar ant species based on morphological features. We establish a public database of ant cuticle images for research. We provide a comparative study of the performance of image texture classification and deep machine learning methods on this ant cuticle dataset. Our results show that the deep learning methods give higher accuracy than statistical methods in recognizing ant cuticle textures. Our experiments also reveal that the deep learning networks designed for image texture performs better than the general deep learning networks.

**Keywords:** texture analysis; image processing; classification; machine learning; ant cuticle images; ecology

---

### 1. Introduction

Insects compose half of biodiversity and rank among the most dominant organisms in terrestrial ecosystems [47]. A key factor for the ecological success of insects is their exoskeleton, also known as cuticle. The cuticle protects insects from predation, provides structural support, prevents desiccation, and serves as a canvas for advertising visual and chemical signals [17]. Earlier research has focused on the macrostructures and internal chemical components that describe the functionality of the exoskeleton [18]. More recent work is being done to understand the functional aspects of external

cuticle micro sculpturing [41,52].

Due to the extensive number of insect species, manual exploration of insect-based information is difficult and often requires expert training. Therefore, automated entomology has attracted biologists and computer scientists, and is expected to be a major contribution to the future of insect-based research [40]. One of the most commonly used data types for insect analysis is image data. To develop an image-based system for insect analysis, we can utilize existing work in general image analysis methods.

We study ants (*Formicidae*) as they display an extreme diversity of cuticle micro sculpturing across all subfamilies. Microsculpture ranges from parallel longitudinal ridges to deep oval impressions to erratic protuberances. Microsculpture has arisen convergently and independently throughout ant's evolutionary history, which suggests that these are complex traits undergoing selection. Cuticle microsculpture on ants may help increase strength and rigidity, resist abrasion, increase internal and external surface area, resist microbial growth, and cultivate beneficial anti-biotic producing bacteria [5, 11, 30]. These specific functions may be associated with certain sculpturing types. To analyze those functions, we first segment the image and group similar textures for further analysis.

There is a large variety of ant species and most species are diverse in terms of size, shape, behaviors, and cuticle textures [22]. Based on our initial observation of the cuticle sculpturing in ants, it appears that some type of textural structures can be identified in images. Therefore, it is worthwhile to explore image texture analysis for further knowledge discovery. Image texture analysis has been used in image processing for many decades [28]. Due to the diversity of images on scales, smoothness, coarseness, microtexture, and macrotexture occurring in images, it is a difficult problem. With a rapid development of deep machine learning, deep neural networks have been frequently used in image texture analysis in other fields [34, 35, 49].

Our contribution is to establish an ant image dataset which can be used to explore the categorization of different ant cuticle textures using image texture analysis methods. We document how the dataset of ant cuticle images is established. Our experiment shows that the deep learning networks designed for image texture analysis give a better accuracy than classical K-means clustering.

This paper is organized as below: Section 2 reviews the microsculpturing identification in ants, and image texture analysis and insect classification. Section 3 describes the dataset preparation and image texture analysis methods used in our experiments. Section 4 gives the experimental setup and methods used. Section 5 shares visualization of both raw and processed images, and analysis of experimental results. The conclusion then follows.

## 2. Related work

In this section we review related work on microsculpturing identification, image texture analysis, and insect classification.

### 2.1. Microsculpturing identification

Because ants lack bold coloration and complex wing venation patterns used to identify other insect species, microsculpturing has taken a prominent role in ant identification. To aid in species identification, ant taxonomists have developed nearly 100 terms to describe fine differences in microsculpturing patterns [3, 22]. Despite the prominent role of microsculpturing in ant identification, however, almost nothing is known about how these patterns evolved or their function. Similar

microsculpturing patterns in other organisms have been found to reduce abrasion, increase stiffness, and perform as structural antibiotics [8,19,23,46,57]. Developing a better understanding of the function of microsculpturing in ants could provide insight into their evolution and biology as well as lead to the development of bio-inspired technologies [44].

One challenge for functional investigations of cuticle microsculpturing is the complexity of terms used to describe these patterns. Complex terminology may be useful for describing fine differences between closely related species, but broader comparisons require a simplified classification scheme. We recently developed a classification system for microsculpturing that consists of only five categories: smooth, striate, punctate, reticulate, and tuberosus [26]. We used these categories to classify 11,722 ant species, and found that the earliest ants were smooth and that cuticle microsculpturing independently evolved later in multiple lineages [26]. A simple classification scheme divides ants into two categories: smooth and rough. Because microsculpturing has evolved numerous times in ants, this binary classification scheme is sufficient.

The next challenge is to find out how to apply these schemes to large numbers of species. Insect researchers have access to millions of images available through digital museum collections [1]. Going through these digital collections manually is not feasible for large projects. There are over 14,000 described species of ants, and each species has multiple worker and reproductive castes that may vary in their microsculpturing patterns. Even within an individual ant, microsculpturing typically varies among and within body segments. Machine learning offers a solution to this problem through developing tools that automate trait classification. Automated approaches would allow researchers to leverage existing image repositories to extract large trait datasets that can then be used to detect evolutionary patterns.

## 2.2. Image texture analysis

Pattern recognition methods have been developed for textural images based on the spatial relationship of gray levels of pixels encoded in the texture [21, 28]. Those methods extract texture features and the classification and segmentation methods are used for the interpretation either using the supervised or unsupervised approaches [14]. The interpretation method is usually modified from the traditional pattern recognition methods. For example, the gray-level co-occurrence matrix (GLCM) and local binary patterns (LBP) are two common methods for extracting features [20,42].

Classical image texture analysis methods can be grouped into four categories: statistical, structural, model-based, and transform-based methods [2]. The statistical method is often used for image classification after feature extraction. Markov random field models have also been used for textural image interpretation [10, 24]. Transform-based methods use some functions to decompose an image texture into a set of basic feature images. Gabor filters and wavelet expansions are two of the widely used approaches [4].

Liu et al. gave a comprehensive survey on textural characterization in which they call this feature extraction as texture representation [34]. Their survey on bag of words and convolutional neural networks (CNNs) [32] provides information on spatial relationship concepts. The K-views was developed for taking the spatial information in the clustering approach [28].

## 2.3. Deep learning for texture analysis

In the past decade, deep learning using CNNs has emerged as the mainstream technology for image analysis. Following this trend, various related network architectures have been designed to specifically

characterize textural images. Among these, the Fisher-vector CNN descriptor (FV-CNN) proposed by Cimpoi et al. [9] is widely accepted as one of the pioneering works. It applies Fisher-vector pooling to deep features obtained via a CNN pre-trained using the ImageNet [32] to obtain encoded features for texture classification. FV-CNN is capable of achieving higher classification accuracy comparing to traditional hand-crafted texture features. However, CNN-based deep learning is only used in its feature extraction module. Its encoding and classification modules still use traditional learning methods, thus not supporting an *end-to-end* deep learning. In the context of texture image classification, an end-to-end deep learning model refers to a neural network architecture which consists of all the modules (feature extraction, encoding and classification) between the initial input data and the final output classification result, with all the modules being trained simultaneously.

To achieve an end-to-end learning, Zhang et al. [56] proposed the deep texture encoding network (DeepTEN), in which a novel texture encoding layer is added to a standard CNN architecture. Then, Xue et al. [53] constructed the deep encoding pooling network which improves over DeepTEN by integrating local spatial characteristics into the texture representation. Based on these two methods, Hu et al. [27] further developed the multi-level texture encoding and representation (MuLTER) network, which embeds a learnable encoding module at each convolutional layer so that encoding is performed for both low-level and high-level features, yielding a multi-level texture representation.

Other network architectures for end-to-end texture learning are also available. For example, in the deep multiple-attribute-perceived network (MAP-Net) [55], multiple perceptual attributes are progressively learned in a mutually reinforced manner through multiple branches. In the deep structure-revealed network (DSR-Net) [54], inherent structural representation for a texture pattern is obtained by employing a primitive capturing module to learn spatial primitives and a dependency learning module to capture the dependency among the primitives. In [38], a residual pooling layer consisting of a residual encoding module and an aggregation module is used to generate discriminative features of low dimensions. In [43], a histogram layer is designed to compute local spatial distribution of CNN features. In [6], an innovative aggregation module is presented to exploit statistical self-similarity across layers. All these architectures customize the standard CNN structure to accomplish the characterization of certain spatial, visual, and statistical traits unique to textural images.

Compared with the traditional method in which a kernel must be designed by an engineer for extracting features, the deep learning networks can automatically extract the features through the training. In addition, the deep learning networks can achieve the higher accuracy than that of the classical approaches, although the deep networks require much more data for training.

#### 2.4. Classification of insects

Proposed insect classification methods seek to classify insects at different hierarchical levels, such as species, genus, family, and order. Additionally, some methods may classify insects at a combination of different hierarchical levels. Insect classification methods can be applied to a variety of fields. In agriculture, insect classification methods can be used to identify the presence of pest insects in crops, which can inform crop managers in their choice of pesticides and help prevent crop loss [31, 35].

Feng et al. [13] apply an automated system to classify moth images based on semantic related visual attributes, which are defined as a pattern on the moth wings. Feng et al. [13] use a custom texture descriptor based on the combination of GLCM and *scale-invariant feature transform* (SIFT) features [16, 37]. The method is used to classify 50 different moth species across 8 families [13]. Their results [13] suggest that traditional feature extraction techniques for the semantic visual attributes of

the moth wings are sufficient for training a classifier to classify an image between 10 randomly selected moth species. Marques et al. [39] propose an ensemble-based method to classify a large number of ant genera. They use ant head, profile, and dorsal images sourced from AntWeb in order to classify 44,806 specimens with at least one picture into 57 genera. Although the images used in their work are also from AntWeb, our goal is the classification of the ant cuticle micro sculpturing into broad groups based solely on the ant head images.

Urteaga et al. [51] use machine learning methods in order to classify images between two different scorpion species: *Centruroides limpidus* and *Centruroides noxius*. After applying background distinction based on dynamic color threshold, they extract features from the separated scorpion image such as aspect ratio, rectangularity, and compactness. They apply three different models to classify the image as one of the species: artificial neural network, regression tree, and random forest classifiers [51]. Their results show that after background removal, characteristics from the entire body of the scorpion can be used to create a binary classifier that can classify the image as one of the two species.

Lim et al. [33] apply a CNN-based algorithm for insect classification. Lim et al. [33] classify a subset of insect species and families based on the classes available in the ImageNet dataset. ImageNet is a widely used dataset of images labeled by experts with millions of images and thousands of categories [12]. In the ImageNet dataset, there are some categories that specify the class of the insect on a species level, e.g. *monarch butterfly* and *ringlet butterfly* as well as some categories that specify the class of the insect on a family level, e.g. *ant*, *fly*, and *bee* [12]. Lim et al. [33] use a modified AlexNet architecture and experiment with different numbers of kernels and their effect the performance of the model. Glick et al. [15] employ a similar approach by classifying 277 insect classes from ImageNet using a hierarchical convolutional neural network. The results from Lim et al. [33] and Glick et al. [15] suggest that a CNN is capable of differentiating between different hierarchical classes of insects.

### 3. Ant cuticle image dataset

In this section, we describe the creation of the custom dataset used in this research. We use ant head images from AntWeb [45] and define two categories for them based on the appearance of the cuticle texture: *rough* and *smooth*. The original AntWeb dataset was extended with these categories for each applicable ant specimen. The extended dataset is used for classifying the cuticle micro sculpturing into the broad groups. Due to the large number of ant species, it is quite difficult to obtain lab-quality specimen photos. For each species used, there exists only one ant head image. We also refer to these ant head images, which describe the pose of the ant, as ant cuticle images, due to the clear cuticle texture present on the head. Some randomly selected images from each category are shown in Figures 1 and 2.

To begin, a master spreadsheet was created with the 2,499 different ant species to be identified for the primary dataset. The team was trained to identify cuticle sculpturing through a process which consisted of one 45-minute introductory lesson explaining the project and texture categories. Then, the team was given a training set of photos to identify from the genus *Polyrhachis*. The sculpture identification protocol describes the two primary categories: rough and smooth.

Initially, the sculpture identification protocol had five subcategories of cuticle texture: smooth, punctate, striate, reticulate, and tuberos. For simplicity, we worked only with the two main categories, and the other four categories were combined to the single rough category. The training set identifications were reviewed together as a group by the assistants. Once training was complete,



**Figure 1.** Examples of smooth cuticle texture ant images in the dataset after center cropping, from AntWeb [45].

assistants were assigned the same genera of ants to identify independently each week. A weekly meeting was held to discuss identifications and assign new ones. These identifications were collected in the master spreadsheet and the identifications were assigned to individual ant species on a majority basis.

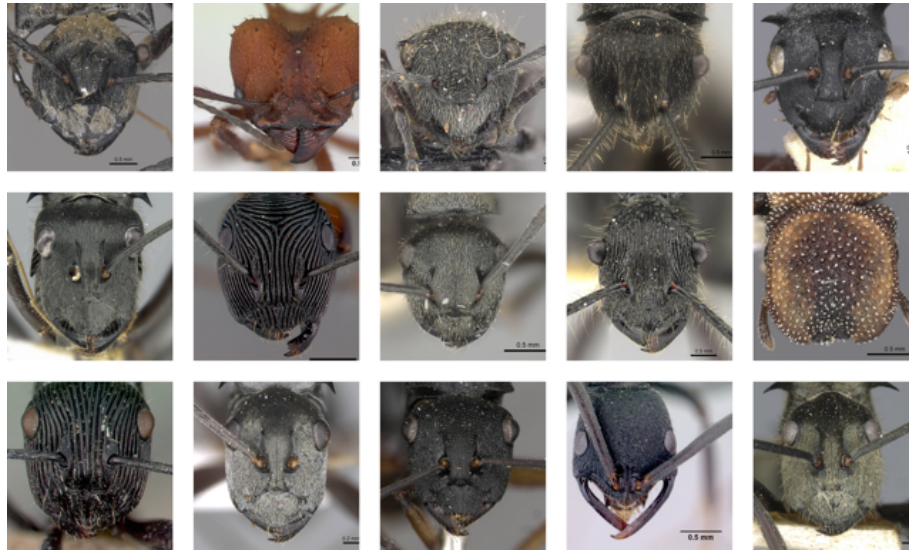
To collect the images from AntWeb\*, the assistants followed the taxonomy information available in the master spreadsheet to the appropriate AntWeb page. In many cases, there are multiple ant head images of the same species, and occasionally there are multiple image resolution available from a single image. To simplify the data collection process, the assistants were instructed to download the first ant head image of the species being identified in the highest resolution possible. Each image was named with an identifier that corresponds with the row number in the master spreadsheet. The same ant head images that were downloaded in the data collection phase were the same ones used in the sculpture identification protocol. Ant species which did not have any images of the head were excluded from the dataset. Additionally, ant species which only had a head image of a queen ant were excluded from the dataset.

Ant specimen images in AntWeb [45] are created by different photographers and therefore have different attributes, such as environment, resolution, and lighting. In the ant head images, the ant head is in the center of the image and the body is pointing away from the camera. The focus of the ant head image is centered on the head, with the background and image artifacts from the ant body typically blurred. In most ant head images, there is a bar which indicates the scale of the image due to the variety in the sizes of different ant species. In a few ant head images, there exists some text denoting the specimen identifier and other information. In terms of texture, some ant specimens are very old, so their head images have other abnormalities such as cracks in the cuticle and the presence of dust.

The completed dataset<sup>†</sup> contains 2,499 images as a 4-dimensional array with shape of (samples, rows, columns, channels), of which there are 3 channels extracted from the RGB images. 1072 samples of rough textured ant cuticle textures comprise 43% of the dataset. The remaining 1427 samples of

\*<https://www.antweb.org/>

<sup>†</sup>The dataset is available on GitHub <https://github.com/ngngardner/cuticulus>



**Figure 2.** Examples of rough cuticle texture ant images in the dataset after center cropping, from AntWeb [45].

smooth textured ant cuticle textures comprise 57% of the dataset. Due to the variety of the ant head image attributes, we apply simple preprocessing before the images are used in our model. We want the images to have a uniform size for simplicity in our classification process. Since the ant head images are typically centered in the image, we apply a center crop to each image to create a square image of the same size. Once the image is square, we resize each image to a fixed size of  $256 \times 256$  pixels. We leave other discrepancies in the images untouched. A summary of the resulting dataset is shared in Table 1.

**Table 1.** Summary of the dataset.

Total images	2499
Rough images	1072
Smooth images	1427
Image size	$256 \times 256$

#### 4. Methods tested

The K-means algorithm is one of widely used clustering algorithms in the pattern recognition [36]. It is a single pixel based classification algorithm for images. Based on the concept of K-means, K-views was developed, which uses characteristic views for image texture representation [29]. The K-views algorithm is suitable for classifying image textures that have basic local patterns repeated in a periodic manner. In contrast to the K-means, K-views looks at the neighboring pixels for providing the spatial relationship for classification.

For the feasibility study in these ant image datasets, we tested the statistical clustering methods and deep learning methods separately. Both K-means and K-views were used for statistical clustering methods. The experiments may indicate how well the statistical approach and deep learning networks

will do in this type of textural images. A pre-processing step of the dataset for the statistical clustering methods is shown below. This pre-processing will highlight some of characteristic features in textures, such as edge-like in the image. This step is useful in feeding the datasets to the K-views algorithms. In order to classify an image with K-views or K-means in which we count the number of pixels assigned to each cluster. A brief procedure for K-views experiments is given below for training and classification.

**Step 1:** Pre-processing the dataset: apply the Gabor filter to transform the gray-scale ant image into four Gabor-filtered channels, subtract the gray-scale image from each channel, and take the absolute value of the resulting images. The theta  $\theta$  parameters in the Gabor filters used are multiples of  $\pi/4$  radians starting from 0 to  $\pi$  radians. The Gabor filter lambda  $\lambda$  parameter is equal to  $\pi$ .

**Step 2:** Initialize  $k = 3$  clusters; one for background, and the other two as rough and smooth.

**Step 3:** Extract patches for each cluster for the clustering.

**Step 4:** Calculate the mean of each cluster once the training converges.

**Step 5:** Use the mean of the cluster for the classification. To determine if an image is rough or smooth, we choose the cluster with the majority number of pixels. The number of pixels assigned to the background cluster are ignored for classification.

Steps 2 to 4 are for training and step 5 is for the classification. For the K-means experiments, it will be the same as the K-views except no patches are used.

The first deep learning model used in our experiment is *visual geometry group* (VGG), a convolutional neural network that takes advantage of very small convolutional filters in a deep network architecture [48]. We compare four architectures of VGG11, VGG13, VGG16, and VGG19. The primary difference between the architectures is the number of layers in each model. The second deep learning model used in our experiments is *residual network* (ResNet), a deep network architecture that includes shortcut connections between layers (residual connections) [25]. We compare three architectures of ResNet18, ResNet50, and ResNet101. Again, the primary difference between the architectures is the number of layers in each model.

For the initial ResNet models, we have two versions: randomized and pretrained. The randomized version is the same architecture, but the weights are randomly initialized. The pretrained version has weights from training on the CIFAR dataset, an image dataset with 1000 classes. Since the pretrained weights are readily available, we also evaluate a fine-tuned model. For VGG, we only use the randomized version. The base VGG architecture also has an output layer of size 1000. Since we are working with a binary classification problem, we modify the architecture for all models to have an output layer of size 2. Each model is trained over 100 epochs, using stochastic gradient descent with momentum. The batch size is set to 16 images. We apply a learning rate of 0.001 and momentum parameter of 0.9.

In addition to the classical texture analysis methods and the deep learning models for general image analysis, we also examine deep learning analysis methods specifically designed for textural images. For this examination, we adopt the deep residual pooling (DRP) network [38]. The DRP framework consists of a unique residual pooling layer, which is formed by a residual encoding module followed by an aggregation module. The residual encoding module extracts relevant spatial information, while the aggregation module performs averaging to obtain orderless low-dimension features for classification.



According to [38], the DRP network can be used either with a single residual pooling layer, applied right before the classifier layer, or with multiple residual pooling layers, applied to the output of multiple convolutional layers before feeding the combined pooling results into the classifier layer. Additionally, an auxiliary classifier layer may also be used. In our experiments, we follow the same strategy. For each scenario, we conduct two experiments. One experiment employs weights that are randomly initialized, and the other uses weights that are fine tuned over those from a pretrained model.

To handle the imbalance of the dataset, we apply undersampling for each class for the training dataset [50]. By using random stratified sampling, we construct a training set with 800 images per class. The remaining images are randomly split between test and validation, which turns out to roughly a 60%/20%/20% train, test, and validation data split. With 272 rough samples and 627 smooth samples left over after the stratified split, the test dataset has roughly 136 rough samples and 313 smooth samples. Since these leftover samples are split with code by 50% there will be some rounding variance and therefore the test dataset built at run-time will not always have exactly the same number of samples.

We evaluate the performance of the models using accuracy, precision, F1 score and Matthew's correlation coefficient (MCC) [7]. We also apply Grad-CAM with manual inspection to visualize the activation weights for classified images to show which features lead to the classification result. Finally, we apply t-SNE to visualize the separation learned for the model to further analyze the classifications made by the model.

## 5. Results

Experiments were run on an Ubuntu 18.04 LTS Lambda Labs GPU server. The server contained 8 NVIDIA GeForce RTX 2080 Ti graphics cards with 12GB of memory each. The server used an Intel Xeon Silver 4116 with 48 total threads and maximum frequency of 3.000 GHz, and has 256 GB of RAM. K-means algorithm comes from the scikit-learn<sup>‡</sup> library. The K-views implementation was based on the K-means implementation from scikit-learn. VGG and ResNet models come from the torchvision<sup>§</sup> library. DRP model code was tweaked from the original author's uploaded code on GitHub<sup>¶</sup>.

The results show that the fine-tuned ResNet models outperform the VGG and randomly initialized ResNet models on the task of ant head image classification. The DRP models perform better than the the general deep learning models VGG and ResNet on the task of ant head image classification. The fine-tuned DRP single-layer model with auxiliary classifier performed the best with an average F1 score of 0.92. We further analyze the separation learned by both ResNet101 models in the following section.

We share the results on each algorithm specified in Section 4. Both the K-means and K-views algorithms are unsupervised. All of the deep learning methods used are supervised. The results are shown in Table 2. It should be noted that due to the class imbalance in the dataset, the F1 score is the preferable metric to the accuracy.

In this section, we provide visualization of the fine-tuned ResNet101 model and the randomly initialized ResNet101 model using t-SNE dimensionality reduction. To visualize the deep extracted features, we modify each model to obtain the embeddings of the second to last layer. We plot side-by-side the ground truth and predicted labels for each model. Figure 3 shows the results of the trained

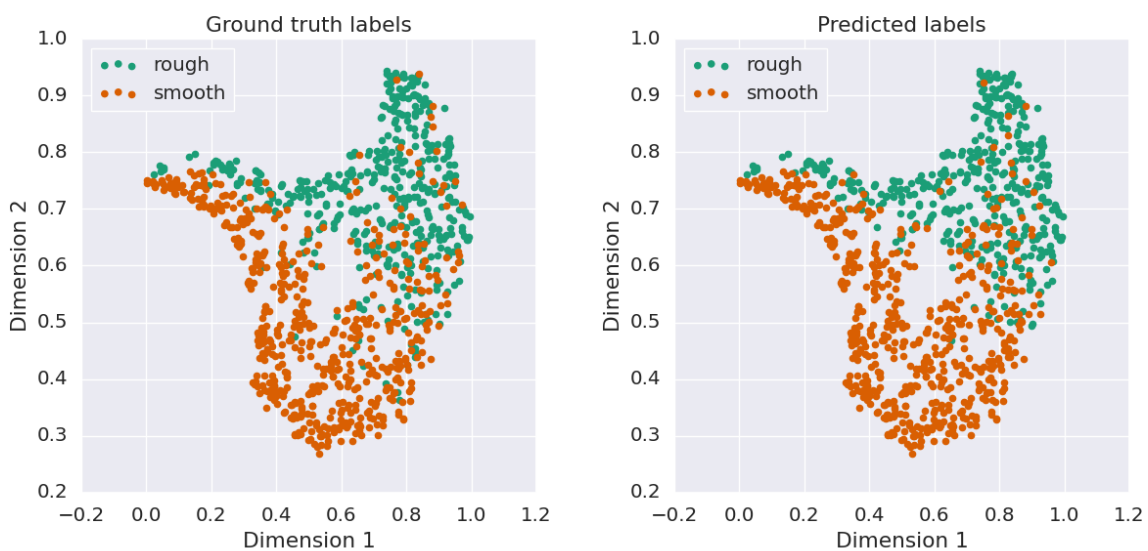
<sup>‡</sup><https://scikit-learn.org/>

<sup>§</sup><https://pytorch.org/vision>

<sup>¶</sup><https://github.com/maoshangbo/DRP-Texture-Recognition>

**Table 2.** Experimental results of different algorithms

Model	Accuracy	Precision	Recall	F1	MCC
K-means	0.47	0.07	0.35	0.11	-0.10
K-views (15 × 15)	0.56	0.58	0.56	0.57	0.12
K-views (17 × 17)	0.62	0.79	0.59	0.68	0.26
K-views (19 × 19)	0.62	0.80	0.59	0.68	0.26
K-views (25 × 25)	0.52	0.71	0.51	0.59	0.03
ResNet18	0.80	0.72	0.67	0.69	0.55
ResNet18 (fine-tuned)	0.88	0.87	0.77	0.82	0.73
ResNet50	0.78	0.59	0.67	0.62	0.47
ResNet50 (fine-tuned)	0.87	0.83	0.78	0.80	0.71
ResNet101	0.77	0.63	0.62	0.62	0.46
ResNet101 (fine-tuned)	0.89	0.87	0.78	0.82	0.74
VGG11	0.80	0.73	0.66	0.69	0.55
VGG13	0.82	0.69	0.71	0.69	0.56
VGG16	0.80	0.65	0.67	0.65	0.52
VGG19	0.81	0.67	0.68	0.64	0.53
DRP Multi-layer	0.87	0.89	0.92	0.91	0.70
DRP Multi-layer (fine-tuned)	0.88	0.89	0.93	0.91	0.72
DRP Single-layer	0.88	0.90	0.94	0.92	0.73
DRP Single-layer (fine-tuned)	0.88	0.90	0.93	0.91	0.74
DRP Single-layer Auxiliary	0.87	0.89	0.92	0.90	0.70
DRP Single-layer Auxiliary (fine-tuned)	0.89	0.90	0.94	0.92	0.75

**Figure 3.** t-SNE visualization of the embeddings of the second to last layer of the randomly initialized ResNet101 model trained on dataset.

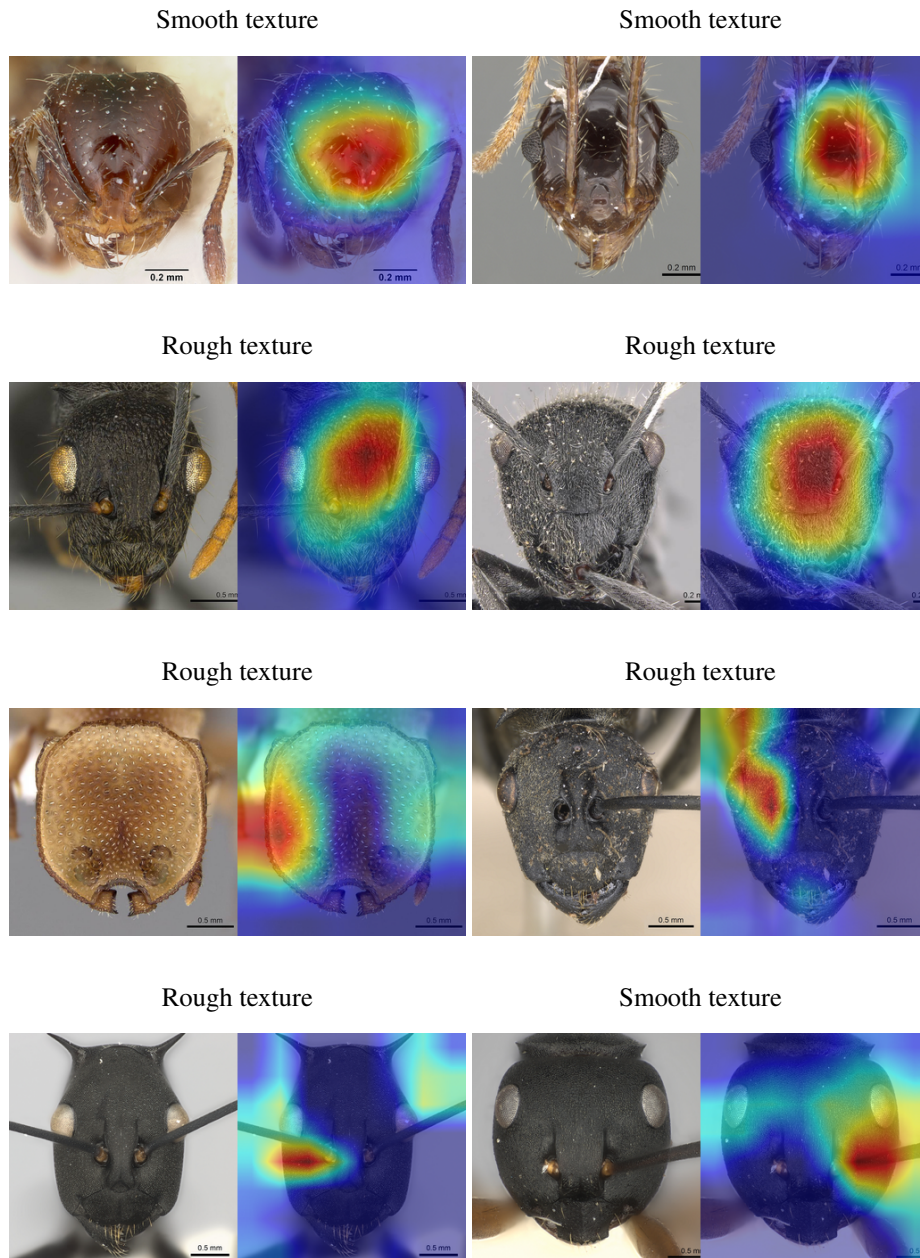
randomly initialized model and Figure 4 shows the results of the fine-tuned model. Based on the visual results of the t-SNE visualization, we can see that the fine-tuned model learned a stronger separation of the two classes, which reinforces the results that the fine-tuned model received a higher average accuracy.



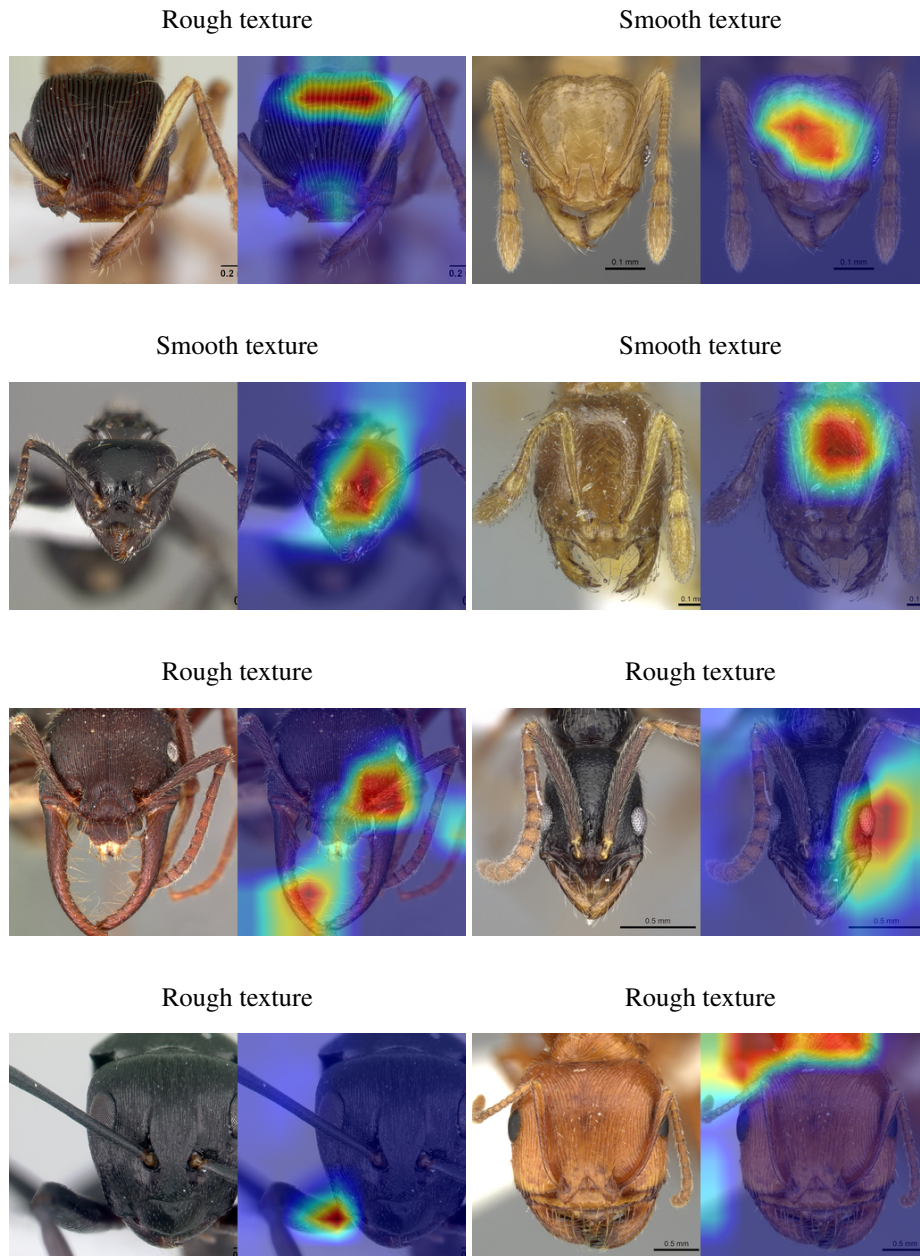
**Figure 4.** t-SNE visualization of the embeddings of the second to last layer of the fine-tuned ResNet101 model trained on dataset.

Next, we provide some visual analysis of some correctly and incorrectly classified images using GradCAM. In the expected case, the features that are used to compute the classification are the same as the features used by the assistants in the sculpture identification process. In general, the features used by the assistants are the textures of the cuticle on the ant head. In the non-expected case, the features used to compute the classification are not from the head, for example, from the background, extraneous text, or the body of the ant. We used randomly selected images from the dataset and the fine-tuned ResNet101 model to perform the analysis. We show the GradCAM results in Figures 5 and 6. The left image shows the preprocessed image input to the model. The right image shows the GradCAM output based on the classification. Four specimens were selected randomly from each category and subcategory.

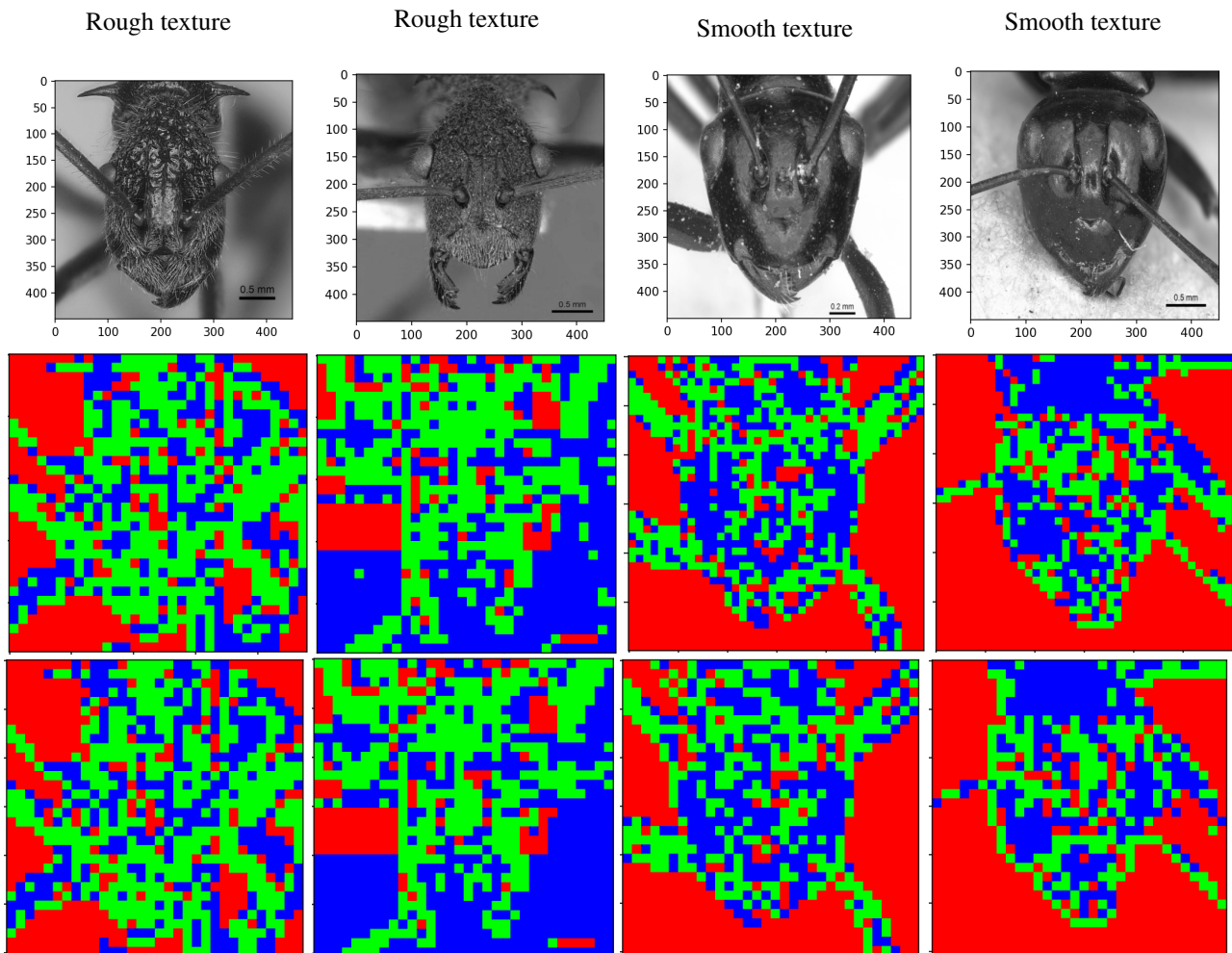
Correctly classified images which use the expected features show the ideal performance of the model. Incorrectly classified images which use the expected features should be further analyzed. In essence, the model in this situation knows *where* to look, but not *what* to look for. In Figure 6, there are results where the features activated are mostly in the correct location on the ant head, and the rough texture is clearly visible, yet the model predicts the incorrect class *smooth*. Similarly in Figure 6, there are results where the features activated are also mostly in the correct location, yet the model predicts the incorrect class *smooth*. In this case, it may be due to the pose of the ant being slightly different from the average pose. In the incorrectly classified images with non-expected features, analysis shows that the model is unable to find *where* to look, and obtains feature information from other parts of the ant or the background. Cases where the image was correctly classified using the non-expected features can basically be seen as noise. In order to further analyze this class, we should introduce some parameter such as model confidence to examine further.



**Figure 5.** GradCAM visualization for correctly classified images from fine-tuned ResNet101. The red sections of the heatmap show which features had higher activation for the classification.



**Figure 6.** GradCAM visualization for incorrectly classified images from fine-tuned ResNet101. The red sections of the heatmap show which features had higher activation for the classification.



**Figure 7.** Raw images are shown with corresponding segmented outputs from K-views algorithm. The second row shows the results of using size of  $15 \times 15$  patches. The third row shows the results of using size of  $19 \times 19$  patches. In the images, the color red represents the background class, green represents the smooth class, and the blue represent the rough class.

Some classified results from the K-views are shown in Figure 7. A training set of 550 patches for background, 455 patches for rough, and 493 patches for smooth are used with patch sizes of  $15 \times 15$  and  $19 \times 19$ . These patches are extracted from two rough and two smooth images with sizes of  $608 \times 608$ . The K-views algorithm converges early although we set the maximum number of iterations to be 1000. These experimental results are preliminary due to the minimal number of source images used and future work will explore a more expansive approach to the problem.

## 6. Conclusions

We have shown in this work that a deep learning approach and classical image texture analysis methods can be used to automatically categorize ants based on their cuticle texture, therefore supporting research on the evaluation of the cuticle micro sculpturing function in future work. Our categorization system is novel in the field of automated insect identification due to the broad number of species captured by it. A model that is pre-trained on a diverse image task such as ResNet can be transferred to our domain of texture analysis<sup>||</sup>. However, a deep learning algorithm created specifically for the domain of texture analysis had an F1 score of 92% in our experiments. In future work, we will continue to explore texture classes present in the dataset which are not captured by the binary class system. We will also investigate application of texture-based deep machine learning in image texture analysis.

## Conflict of interest

All authors declare no conflicts of interest in this paper.

## References

1. R. Beaman and N. Cellinese, Mass digitization of scientific collections: New opportunities to transform the use of biological specimens and underwrite biodiversity science, *ZooKeys*, **209** (2012), 7–17. <https://doi.org/10.3897/zookeys.209.3313>
2. M. H. Bharati, J. J. Liu and J. F. MacGregor, Image texture analysis: methods and comparisons, *Chemometr. Intell. Lab.*, **72** (2004), 57–71. <https://doi.org/10.1016/j.chemolab.2004.02.005>
3. B. Bolton, *Identification guide to the ant genera of the world*, Harvard University Press, Cambridge, Mass, 1994.
4. A. C. Bovik, M. Clark and W. S. Geisler, Multichannel texture analysis using localized spatial filters, *IEEE T. Pattern Anal.*, **12** (1990), 55–73. <https://doi.org/10.1109/34.41384>
5. A. Brückner, M. Heethoff and N. Blüthgen, The relationship between epicuticular long-chained hydrocarbons and surface area - volume ratios in insects (Diptera, Hymenoptera, Lepidoptera), *PLOS ONE*, **12** (2017), e0175001, <https://doi.org/10.1371/journal.pone.0175001>
6. Z. Chen, F. Li, Y. Quan, Y. Xu and H. Ji, Deep Texture Recognition via Exploiting Cross-Layer Statistical Self-Similarity, in *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, (2021), 5231–5240.

<sup>||</sup>All code is publicly available on GitHub <https://github.com/ngngardner/cuticulus>

7. D. Chicco, M. J. Warrens and G. Jurman, The matthews correlation coefficient (mcc) is more informative than cohen's kappa and brier score in binary classification assessment, *IEEE Access*, **9** (2021), 78368–78381. <https://doi.org/10.1109/ACCESS.2021.3084050>
8. K. K. Chung, J. F. Schumacher, E. M. Sampson, R. A. Burne, P. J. Antonelli and A. B. Brennan, Impact of engineered surface microtopography on biofilm formation of *Staphylococcus aureus*, *Biointerphases*, **2** (2007), 89–94. <https://doi.org/10.1116/1.2751405>
9. M. Cimpoi, S. Maji and A. Vedaldi, Deep filter banks for texture recognition and segmentation, in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, (2015), 3828–3836. <https://doi.org/10.1109/CVPR.2015.7299007>
10. G. R. Cross and A. K. Jain, Markov random field texture models, *IEEE T. Pattern Anal.*, **PAMI-5** (1983), 25–39. <https://doi.org/10.1109/TPAMI.1983.4767341>
11. C. R. Currie, Coevolved Crypts and Exocrine Glands Support Mutualistic Bacteria in Fungus-Growing Ants, *Science*, **311** (2006), 81–83. <https://doi.org/10.1126/science.1119744>
12. J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li and L. Fei-Fei, ImageNet: A large-scale hierarchical image database, in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, (2009), 248–255.
13. L. Feng and B. Bhanu, Automated identification and retrieval of moth images with semantically related visual attributes on the wings, in *2013 IEEE International Conference on Image Processing*, (2013), 2577–2581.
14. K. Fukunaga, *Introduction to statistical pattern recognition*, Elsevier, 2013.
15. J. A. Glick and K. Miller, Insect Classification With Heirarchical Deep Convolutional Neural Networks, *Convolutional Neural Networks for Visual Recognition*, (2016).
16. C. C. Gotlieb and H. E. Kreyszig, Texture descriptors based on co-occurrence matrices, *Computer Vision, Graphics, and Image Processing*, **51** (1990), 70–86. [https://doi.org/10.1016/S0734-189X\(05\)80063-5](https://doi.org/10.1016/S0734-189X(05)80063-5)
17. P. J. Gullan and P. S. Cranston, *The Insects: an Outline of Entomology*, Wiley, Hoboken, 2009.
18. S. Gunderson and R. Schiavone, The insect exoskeleton: A natural structural composite, *JOM*, **41** (1989), 60–63. <https://doi.org/10.1007/BF03220386>
19. Z. Han, H. Feng, W. Yin, S. Niu, J. Zhang and D. Chen, An Efficient Bionic Anti-Erosion Functional Surface Inspired by Desert Scorpion Carapace, *Tribol. T.*, **58** (2015), 357–364. <https://doi.org/10.1080/10402004.2014.971996>
20. R. M. Haralick, K. Shanmugam and I. Dinstein, Textural Features for Image Classification, *IEEE Transactions on Systems, Man, and Cybernetics*, **SMC-3** (1973), 610–621. <https://doi.org/10.1109/TSMC.1973.4309314>
21. R. M. Haralick and L. G. Shapiro, Image segmentation techniques, *Computer vision, graphics, and image processing*, **29** (1985), 100–132. [https://doi.org/10.1016/S0734-189X\(85\)90153-7](https://doi.org/10.1016/S0734-189X(85)90153-7)
22. R. A. Harris, A Glossary Of Surface Sculpturing, *Occasional Papers in Entomology*, **28** (1979), 1–31.
23. J. Hasan, H. K. Webb, V. K. Truong, S. Pogodin, V. A. Baulin, G. S. Watson, et al., Selective bactericidal activity of nanopatterned superhydrophobic cicada *Psaltoda claripennis*



- wing surfaces, *Appl. Microbiol. Biot.*, **97** (2013), 9257–9262. <https://doi.org/10.1007/s00253-012-4628-5>
24. M. Hassner and J. Sklansky, The use of markov random fields as models of texture, *Image Modeling*, (1981), 185–198. <https://doi.org/10.1016/B978-0-12-597320-5.50015-2>
  25. K. He, X. Zhang, S. Ren and J. Sun, Deep Residual Learning for Image Recognition, in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, (2016), 770–778. <https://doi.org/10.1109/CVPR.2016.90>
  26. J. Hellenbrand and C. Penick, Ant cuticle sculpturing: diversity, classification, and function, submitted for publication.
  27. Y. Hu, Z. Long and G. AlRegib, Multi-Level Texture Encoding and Representation (Multer) Based on Deep Neural Networks, in *2019 IEEE International Conference on Image Processing (ICIP)*, (2019), 4410–4414. <https://doi.org/10.1109/ICIP.2019.8803640>
  28. C.-C. Hung, E. Song and Y. Lan, *Image Texture Analysis: Foundations, Models and Algorithms*, Springer International Publishing, Cham, 2019.
  29. C.-C. Hung, S. Yang and C. M. Laymon, Use of characteristic views in image classification, in *16th International Conference on Pattern Recognition, ICPR 2002*, (2002), 949–952. <https://doi.org/10.1109/ICPR.2002.1048462>
  30. R. A. Johnson, A. Kaiser, M. Quinlan and W. Sharp, Effect of cuticular abrasion and recovery on water loss rates in queens of the desert harvester ant *Messor pergandei*, *J. Exp. Biol.*, **214** (2011), 3495–3506. <https://doi.org/10.1242/jeb.054304>
  31. T. Kasinathan and S. R. Uyyala, Machine learning ensemble with image processing for pest identification and classification in field crops, *Neural Computing and Applications*, **33** (2021), 7491–7504. <https://doi.org/10.1007/s00521-020-05497-z>
  32. A. Krizhevsky, I. Sutskever and G. E. Hinton, ImageNet classification with deep convolutional neural networks, *Commun. ACM*, **60** (2017), 84–90. <https://doi.org/10.1145/3065386>
  33. S. Lim, S. Kim and D. Kim, Performance effect analysis for insect classification using convolutional neural network, in *2017 7th IEEE International Conference on Control System, Computing and Engineering (ICCSCE)*, (2017), 210–215. <https://doi.org/10.1109/ICCSCE.2017.8284406>
  34. L. Liu, J. Chen, P. Fieguth, G. Zhao, R. Chellappa and M. Pietikäinen, From BoW to CNN: Two Decades of Texture Representation for Texture Classification, *Int. J. Comput. Vision*, **127** (2019), 74–109. <https://doi.org/10.1007/s11263-018-1125-z>
  35. L. Liu, R. Wang, C. Xie, P. Yang, F. Wang, S. Sudirman, et al., PestNet: An End-to-End Deep Learning Approach for Large-Scale Multi-Class Pest Detection and Classification, *IEEE Access*, **7** (2019), 45301–45312. <https://doi.org/10.1109/ACCESS.2019.2909522>
  36. S. Lloyd, Least squares quantization in pcm, *IEEE T. Inform. Theory*, **28** (1982), 129–137. <https://doi.org/10.1109/TIT.1982.1056489>
  37. D. G. Lowe, Distinctive Image Features from Scale-Invariant Keypoints, *Int. J. Comput. Vision*, **60** (2004), 91–110. <https://doi.org/10.1023/B:VISI.0000029664.99615.94>
  38. S. Mao, D. Rajan and L. T. Chia, Deep residual pooling network for texture recognition, *Pattern Recogn.*, **112** (2021), 107817. <https://doi.org/10.1016/j.patcog.2021.107817>

39. A. C. R. Marques, M. M. Raimundo, E. M. B. Cavalheiro, L. F. P. Salles, C. Lyra and F. J. V. Zuben, Ant genera identification using an ensemble of convolutional neural networks, *PLOS ONE*, **13** (2018), e0192011. <https://doi.org/10.1371/journal.pone.0192011>
40. M. Martineau, D. Conte, R. Raveaux, I. Arnault, D. Munier and G. Venturini, A survey on image-based insect classification, *Pattern Recogn.*, **65** (2017), 273–284. <https://doi.org/10.1016/j.patcog.2016.12.020>
41. S. Muthukrishnan, S. Mun, M. Y. Noh, E. R. Geisbrecht and Y. Arakane, Insect Cuticular Chitin Contributes to Form and Function, *Curr. Pharm. Design*, **26** (2020), 3530–3545. <https://doi.org/10.2174/1381612826666200523175409>
42. T. Ojala, M. Pietikäinen and T. Mäenpää, Gray Scale and Rotation Invariant Texture Classification with Local Binary Patterns, in *European conference on Computer Vision - ECCV 2000*, (2000), 404–420. Springer, Berlin, Heidelberg. [https://doi.org/10.1007/3-540-45054-8\\_27](https://doi.org/10.1007/3-540-45054-8_27)
43. J. Peeples, W. Xu and A. Zare, Histogram Layers for Texture Analysis, *IEEE Transactions on Artificial Intelligence*, **3** (2021), 541–552. <https://dx.doi.org/10.1109/TAI.2021.3135804>
44. C. Penick, G. Cope, S. Morankar, Y. Mistry, A. Grishin, N. Chawla, et al., The comparative approach to bio-inspired design: integrating biodiversity and biologists into the design process, *Intergrative and Comparative Biology*, (2022).
45. V. Perrichot and B. Fisher, AntWeb: digitizing Recent and fossil insects for an online database of the ants of the world, in *Digital Fossil International Conference*, (2012).
46. C. J. C. Rees, Form and function in corrugated insect wings, *Nature*, **256** (1975), 200–203. <https://doi.org/10.1038/256200a0>
47. A. Sheikh, N. Rehman and R. Kumar, Diverse adaptations in insects: A review, *Journal of entomology and zoology studies*, **5** (2017), 343–350.
48. K. Simonyan and A. Zisserman, Very Deep Convolutional Networks for Large-Scale Image Recognition, *arXiv:1409.1556 [cs]*.
49. X. Sun, *Robust texture classification based on machine learning*, PhD thesis, Deakin University, 2014.
50. H. Tiittanen, L. Holm, P. Törönen, H. Tiittanen, L. Holm and P. Törönen, Novel split quality measures for stratified multilabel cross validation with application to large and sparse gene ontology datasets, *Applied Computing and Intelligence*, **2** (2022), 49–62. <https://doi.org/10.3934/aci.222003>
51. J. C. Urteaga-Reyesvera and A. Possani-Espinosa, Scorpions: Classification of poisonous species using shape features, in *2016 International Conference on Electronics, Communications and Computers (CONIELECOMP)*, (2016), 125–129. <https://doi.org/10.1109/CONIELECOMP.2016.7438563>
52. G. S. Watson, J. A. Watson and B. W. Cribb, Diversity of Cuticular Micro- and Nanostructures on Insects: Properties, Functions, and Potential Applications, *Annu. Rev. Entomol.*, **62** (2017), 185–205. <https://doi.org/10.1146/annurev-ento-031616-035020>
53. J. Xue, H. Zhang and K. Dana, Deep Texture Manifold for Ground Terrain Recognition, in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, (2018), 558–567. <https://dx.doi.org/10.1109/CVPR.2018.00065>

54. W. Zhai, Y. Cao, Z.-J. Zha, H. Xie and F. Wu, Deep Structure-Revealed Network for Texture Recognition, in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, (2020), 11007–11016. <https://dx.doi.org/10.1109/CVPR42600.2020.01102>
55. W. Zhai, Y. Cao, J. Zhang and Z.-J. Zha, Deep Multiple-Attribute-Perceived Network for Real-World Texture Recognition, in *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, (2019), 3612–3621. <https://dx.doi.org/10.1109/ICCV.2019.00371>
56. H. Zhang, J. Xue and K. Dana, Deep TEN: Texture Encoding Network, in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, (2017), 2896–2905. <https://dx.doi.org/10.1109/CVPR.2017.309>
57. H. Zhiwu, Z. Junqiu, G. Chao, W. Li and L. Ren, Erosion Resistance of Bionic Functional Surfaces Inspired from Desert Scorpions, *Langmuir*, **28** (2012), 2914–2921. <https://doi.org/10.1021/la203942r>



AIMS Press

©2022 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)