TECHNICAL WHITE PAPER 1

BriteScan® iVeris®: Development and validation of herb and spice authentication from images using Computer Vision/Machine Learning (CV/ML)

Danica T. Harbaugh Reynaud, Ph.D.
Co-Founder & Chief Executive Officer, BriteScan, LLC

Abstract

In this article, we present an extensive case study using the BriteScan iVeris® device to authenticate the species, plant part, and form of nearly 200 commonly used dried medicinal and culinary herbs and spices. The iVeris is a small portable device that uses Artificial Intelligence (AI) to verify the attributes of sample materials based on images, through connected cloud-based Computer Vision/Machine Learning (CV/ML) algorithms similar to fingerprint or facial recognition. In this study, we developed a robust authentication testing method using the iVeris by building a reference database of more than 12,000 images from over 600 authenticated samples of dried herbs and spices in their commonly traded forms. Samples were authenticated using morphology and/or DNA sequencing, where appropriate. To ensure the scientific validity of the test, we conducted multiple studies to examine the accuracy, specificity, and precision (repeatability and reproducibility). The results indicate that the final model we developed is highly accurate (avg = 99.5%), with high probabilities (avg = 95.0%), zero false positives and only 2.1% false negatives. It is also repeatable (avg = 100%) and reproducible both intra-specifically (avg = 94.7%) and inter-specifically (avg = 94.5%), based on testing a portion of the samples. These include testing non-homogenous and known adulterated samples (i.e., Oregano, *Origanum sp.)*, which were used to examine the ability of the model to flag non-conforming materials, possibly indicating adulteration or substitution or human error, warranting additional testing. To further examine the model, we tested its ability to distinguish a portion of the samples both before and after their images were added to the model. The results showed that the additional images improved the accuracy of the model from an average of 86.8% to 95%, and reducing the false positives from 5% to zero, and false negative rates from 17% to 5%. In conclusion, this study demonstrates that the iVeris system can provide scientifically valid and reliable test results to confirm the species, plant part and form of even the most closely related and similar looking dried herbs and spices. Because of the ease-of-use and rapid test results, extensive validation studies can be performed to ensure high-quality testing that can be used along their entire supply-chains.

Introduction

Ensuring the authenticity of natural products including foods, supplements, and botanicals such as culinary and medicinal herbs and spices is critical for protecting human health, minimizing business risks, complying with federal regulations, and eliminating fraud. Traditional testing approaches are often expensive, time-consuming, technical, and cannot be performed in the field. Because of these challenges, many botanical products are inadequately tested- or not tested at all. The patented BriteScan® iVeris® is the first “digital-lab-in-a-box” designed to address the shortcomings of traditional approaches by harnessing the power of Artificial Intelligence (AI) to allow anyone along the supply chain to perform accurate testing rapidly and inexpensively for a wide range of materials and applications.

AI is a broad term that covers any computer system that can replicate human intelligence, immortalizing and infinitely scaling the knowledge of experts and improving upon it where humans have limitations. BriteScan’s technology uses Pattern Recognition and Computer Vision/Machine Learning (CV/ML), similar to Apple’s Face ID and Fingerprint ID. BriteScan’s CV/ML algorithms automatically evaluate all visual aspects of an image such as color, shape, texture, and size, even those which are not detectable by an unaided human eye. In addition to facial and fingerprint recognition, CV/ML is increasingly being used in a wide array of fields including identification of human diseases (1) and living plant and animal species (e.g. see 2, 3), as well as quality inspection and grading of food and produce (4, 5).

The BriteScan iVeris is a small portable device in the nature of a lightbox approx. 20 cm wide x 18 cm deep x 18 cm high that is used to take high-quality, consistent images of any tangible material that fits in the chamber for verification using CV/ML. It contains white LED lights, a rechargeable battery, and multiple staging areas including a pull-out tray to place materials, and either a built-in camera or receptable on which to place a smartphone camera. The iVeris is connected to BriteScan’s cloud-based software that analyzes the images in approximately one minute. The device is essential in that it eliminates external variables and controls the focal distance, background color, position, and area or volume of material so that even the most subtle differences can be detected. Additionally, the imaging consistency also reduces the number of reference images and samples needed to train the algorithm by as much as 99%, saving valuable time and money.

The iVeris has the ability to test a wide range of non-biological and biological materials in any form including powders, liquids, and solids either processed or whole objects or samples of them that can fit within the chamber. Depending on the placement of the materials in the device, it can accommodate small objects up to approx. 4 cm in width, length, and height, large objects 6 cm high and 7 cm in diameter, or flat or small and processed materials that are 4 cm high, 12 cm long and 10 cm wide or up to a volume of approx. 60 ml. It also has a specialized tray designed to hold approx. 10 ml of liquid or powder that can be placed in a standard weigh boat for easy disposal. Additionally, it can accommodate small objects only a few millimeters in diameter.

Example materials that can be tested with iVeris include natural products such as foods, supplements, herbs and spices, as well as meats, fish, oils and beverages. Other non-biological materials such as rocks including precious gems such as diamonds can be verified using the iVeris. Because the iVeris relies on visual characteristics, it is not suitable for completely clear liquids, nor for microscopic organisms that require a light microscope for visualization. However, simple sample manipulation- such as the addition of reactive dyes- can elucidate visual characteristics that wouldn’t otherwise be detectable. It has demonstrated utility for homogeneous liquids as it detects subtle differences in color, and for verification of powers including those that are white, which present differences in color, light reflectance, and particle shape and size at a pixel-level which can be detected by the iVeris software.

Applications for the iVeris are very broad for botanicals and include direct authentication of species, plant part, and form, detection of adulteration, substitution and filth. It also has the ability to verify material attributes that wouldn't necessarily be performed using morphology- or even a microscope. Other indirect testing of attributes such as origin, quality, density, freshness, aroma, and flavor, genotype chemotype or concentrations of secondary metabolites, volatile oils, and other chemicals may be classified based on their co-occurrence with subtle visual differences. While iVeris cannot directly sequence DNA, genotypes are often the result of geographic isolation where plants have evolved different physical traits. This is analogous to using a fingerprint or facial features rather than DNA to identify a human- all can be highly accurate and reliable, and when used in combination they further increase the accuracy and confidence in identification.

Confirmation of plant origin by the iVeris has been highly successful for a number of commodities including black pepper, coffee, and vanilla beans. Because plants are highly affected by their growing conditions such as climate, soil, water, harvest time, and post-processing procedures- even at the farm level- distinct morphotypes often occur. In some cases, specific samples, containers, lots or objects can be fingerprinted by the iVeris and tracked and traced through the supply chain. Therefore, BriteScan can be used to simultaneously verify multiple attributes of a sample. For example, a single picture of an oregano sample could potentially be used to verify the species, plant part, form, quality, freshness, aroma, flavor, and origin.

Just as the iVeris can indirectly confirm the genotype or provenance based on their effect on morphological characteristics, it can predict the presence or concentration of specific chemical constituents, metals, yeasts and molds. While these subtle differences may not be classifiable by a human, they can be learned and remembered by the CV/ML software. The iVeris has been used to accurately detect and estimate the concentrations of low levels of lead chromate in turmeric powder (as low as 1%), as well as salt concentrations in seasoning blends, and volatile oils in mints. Although the iVeris is a qualitative test, reference images representing ranges of known concentrations can be used to train the algorithms, which in turn can classify test images. Results indicate the probabilities of each potential category so that informed decisions can be made based on internal specifications.

The specificity of the BriteScan iVeris "digital fingerprints" even have the ability to identify lots, allowing for verification of lot-to-lot consistency, or tracking them through the supply-chain. The digital data can be easily integrated into existing software such as IBM's FoodTrust™, TraceGains, TagOne, FoodChain ID, or FoodLogIQ, with or without blockchain technology for increased supply-chain traceability and transparency. With iVeris, consistent quality control testing can be used along the entire supply-chain, from the farmer and the supplier, to the distributer, manufacturer, and ultimately the consumer.

In this article, we present an extensive case study using the BriteScan iVeris to authenticate the species, plant part, and form of nearly 200 commonly used dried medicinal and culinary herbs and spices. In this study, we follow the general steps for development of any botanical identification method: *1) Build a library of authenticated reference materials, 2) Test reference materials using the iVeris device and train the AI model,* and finally *3) Validate appropriate parameters to ensure it is fit-for-purpose.* Our goal is to demonstrate that the BriteScan iVeris is a reliable alternative or addition to traditional herb authentication and quality control testing programs.

Materials and Methods

Methods if identification using CV/ML technology for botanicals (or any other material), is analogous to developing any other qualitative identification method. Below we outline in detail the steps that we performed to develop an iVeris method of species, plant part and form verification of dried botanicals and validate it for acceptability for regulatory compliance or internal manufacturing specifications. We accomplish this with three general steps: *1) Build a library of authenticated reference materials, 2) Test reference materials using proposed method,* and finally *3) Validate appropriate parameters to ensure it is fit-for-purpose*

1) Build a library of authenticated reference materials

In the first phase of this project, we built an extensive library of over 600 authenticated reference materials, representing over 200 dried herbs and spices in their commonly traded forms, including whole, cut, chopped, flaked, minced, and broken of a variety of plant parts including leaves, seeds, roots, fruits and flowers (Appendix 1). Powders were not included in this study, because the parameters required for imaging and analysis are different from less processed materials; other studies have demonstrated the utility of the iVeris to authenticate and detect adulteration in a wide range of botanicals, including powders.

A total of 187 different target herbs and spices, as well as over 50 non-target materials were included in the study (Appendix 1). These were represented by a total of 577 individual target samples and 50 non-target samples. To test the ability of the model to detect adulterated oregano- a significant issue in the herb and spice industry- three samples with confirmed adulteration were included in the target list.

To develop an accurate model and eliminate false positives and negatives, inclusion of a wide range of materials that cover all of the acceptable variation that is practically obtainable within each herb and spice is essential. In an effort to cover a wide range of variation in physical attributes that may exist in the marketplace within each herb and spice due to such factors as origin, freshness, or processing techniques, numerous brands of target and a majority of the non-target materials were purchased from a number of grocery stores (e.g., Raley's, Safeway, Whole Foods, Oliver's Market, Target), herb stores (e.g., Rosemary's Garden, Lhasa Karnak, Penzey's Spices), and online retailers (e.g., Amazon, Mountain Rose Herbs).

In addition to the target materials, creation of an extensive "mismatch" or non-target category to capture potential adulterants and avoid false positives is paramount. In this study, non-target or "mismatch" materials include closely related herb and spice species in the same forms and those that are visual similar, as well as collected fresh plant materials and other common inanimate objects (i.e., rocks, landscape bark, pens) that could be used to produce false positive results. However, "tricking" the test is minimized using the iVeris because there is a photographic record on each results report.

All target materials and a portion of the non-target herbs and spices used in the study were authenticated by an expert botanical taxonomist and geneticist using morphology and/or Sanger DNA sequencing, where appropriate. For those samples requiring sequencing, the samples were first ground manually using a mortar and pestle. DNA was extracted and amplified by Polymerase Chain Reaction (PCR) using primers that target the nuclear ribosomal DNA Internal Transcriber Region. The ITS region has proven to be most useful for species discrimination in past botanical taxonomic research (e.g. see 6). Amplified products were sequenced by Genewiz, Inc and DNA sequences were identifyed by performing phylogenetic analyses to compare the sequences to published sequences available on GenBank online.

2) *Test reference materials using the iVeris device and train the AI model*

In the second phase of the project, the authenticated reference materials were used to train the model using the iVeris cloud-based software. This software uses CV/ML algorithms that automatically evaluates all of the physical attributes of the images, such as the color, size, shape and texture at a pixel level. No user input is necessary to guide the algorithms in what characteristics to learn from. There are four simple steps in training the model: (1) Take multiple images of each authenticated sample, (2) Upload the images into the BriteScan iVeris software, (3) Label them with the category name, and (4) Click the "Train" button.

In order to evaluate the range of materials included in this project- from cut and processed herbs such as oregano to whole long objects such as vanilla beans, the image processing included evaluation of the entire sampling tray (approx. 10 cm x 13 cm) with no cropping necessary. The dimensions of the tray are optimized to fill the entire field-of-view of the camera. The sampling area on the tray includes raised sides in order to hold the materials in place, with a maximum volume of approx. 60 ml. In this project, all measurable materials such as chopped, cut, minced and broken materials and small objects such as seeds used 60 ml. While for all larger whole objects such as roots, leaves, and vanilla beans only one object was used. In the case of vanilla beans which were often longer than the tray, they were consistency cut in two pieces. It is imperative to use a consistent amount of material or number of objects when developing a database and testing against it to have accurate results. Additionally, the orientation of the objects (i.e., horizontal vs. vertical) can affect the results, because the algorithms evaluate all aspects of the image including the shape and amount of white space from the sampling tray below materials that don't completely cover it. The more consist the preparation, the fewer images and samples need to be included in the reference image database.

Once all of the samples were authenticated, a total of 11,219 target and 1505 non-target images were taken with the iVeris fitted with a smartphone camera (Appendix 1). A total of 1-13 samples were imaged per category, with an average of approx. 3. For materials that have little variation within and between samples, especially those that are distinct from other categories (i.e., coriander and annatto seeds) only one or a few were sampled. For those materials with more variation within categories and similar looking to others (i.e., bay leaves, or cut basil, oregano, and parsley), more samples were imaged.

For each sample, an average of approx. 20 images in total were taken. This ensured that all of the variation within each sample could be learned by the algorithms. The most accurate CV/ML models are trained using a wide range of images that represent any and all variation that may exist in test materials. The algorithms can learn and remember the variation, and invariably the addition of more and more images and samples will increase the accuracy and probabilities of identification. One way to increase variability is by taking multiple pictures of the same sample. This can be accomplished by moving around the material that is on a tray manually. Multiple images taken from different parts of a container or lot should be also sampled and imaged, especially in the case of non-homogenous materials and whole objects such as leaves, roots, and large fruits that may have major differences in physical characteristics such as shape and size (i.e., ginseng roots). This is also useful for detection of adulteration that may not be evenly distributed without a container or lot.

In this study, three smartphone cameras (Apple's iPhone8 and iPhoneX, and the Samsung Galaxy S9+) were used to image the materials to build the reference image database. In preliminary studies, even the oldest model (Apple iPhone8) had sufficient quality for accurate authentication. However, the phone with the highest quality camera, the Samsung Galaxy S9+, was used for performing the testing against the database in this study. A new model of the iVeris (v.2) (not used in this study) contains a built-in camera similar to the Samsung Galaxy S9+ to provide additional consistency in imaging, without the need to distribute the same phone model to different users of the same database; it can be expected that the accuracy, specificity, and repeatability using iVeris v.2 would be similar or superior to the version used in this study because of the built-in camera.

As the images are taken of the reference samples using the iVeris, they are uploaded into the cloud-based software accessed on the smartphone web-browser and labeled with the correct category before training the algorithms. In general, categories represent each of the potential test results. In this case, each herb and spice represent a category, such as "Cut Turkish Oregano (*Origanum onites*) Leaf", or "Whole Vanilla (*Vanilla planifolia*) Beans." The iVeris test is qualitative, and each category is mutually exclusive so images can only be labeled with a single category. However, each category can contain multiple attributes, such form, species, and plant part, as we did in this study.

For closed systems where all of the possible test results can be included in the reference database (i.e., origin of vanilla beans where all the regions are known and included as categories), a non-target "mismatch" or "unknown" category is not necessary. However, in the case of dried herbs and spices, there are many hundreds of thousands of plant species that cannot all be included in the database. Therefore, the non-target category containing a wide range of variation in similar looking, closely related, known adulterants, as well as other materials or objects that could be used to "trick" the test should be included in it to flag non-conforming samples due to either adulteration or from variation in the target material that is not already included in the reference database. These samples can then be authenticated with an alternative method and then added to the reference database, if applicable, to improve it.

Once the imaging of the reference materials is completed, the "Train" button on the phone or device is clicked and the BriteScan cloud-based AI software trained the model in a few minutes. Once the model is trained from the AI algorithms, it is then available to start testing new sample materials.

3) Validate appropriate parameters to ensure it is fit-for-purpose

In the final phase of this project, we performed multiple studies to test the performance of the model, to ensure it is fit-for-purpose and scientifically valid for verification of dried herb authenticity, plant part, and form for regulatory compliance. In Table 1 below, we have defined terminology for method validation of qualitative CV/ML identification methods such as the iVeris. These include *Accuracy (Model Accuracy, Classification Accuracy), Precision (Repeatability, Extended Repeatability, Reproducibility),* and *Robustness. Reproducibility* and *Robustness* were not included in our study; however, this may be relevant when multiple labs or users are performing testing. In the next sections, further details about each validation study are provided.

Table 1. Computer Vision/Machine Learning (CV/ML) Qualitative Identification Method Validation Definitions

Target Samples: Samples of known identity that belong to a target category included in the model image database.

Non-Target (Mismatch) Samples: Samples of known identity that do not belong to a target category and may or may not be included in a "Non-Target/Mismatch" category included in the model image database.

True Positives: Target samples that were correctly classified.

True Negatives: Non-Target/Mismatch samples that were correctly classified as Non-Target.

False Positives: Non-Target/Mismatch samples that were incorrectly classified as a Target category.

False Negatives: Target samples that were incorrectly classified as Non-Target/Mismatch.

Accuracy:

(a) **Model Accuracy:** The ability of a model to correctly classify samples from images of specimens of known identity that were *used* to train the model. This is calculated by dividing the true positives (TP) and true negatives (TN) by the total number of samples (TP+TN/Total).

(b) **Classification Accuracy:** The ability of a model to correctly classify samples from images of specimens of known identity that were *not used* to train the model. This is calculated by dividing the true positives (TP) and true negatives (TN) by the total number of samples (TP+TN/Total).

Precision:

(a) **Repeatability:** The consistency in identification and associated probabilities of sample identification from repeat tests of the same image of the specimens.

(b) **Extended Repeatability:** The consistency in identification and associated probabilities of identification from repeat tests of different images of the *same* specimens (*Intra-specimen).* Or, the consistency in identification and associated probabilities of identification from *different specimens in the same category* (*Intra-category).*

(c) **Reproducibility:** Consistency of results obtained from samples of the same specimens by different analysts in different laboratories

Robustness: The ability of the model to correctly identify a sample with deliberate changes to the sample preparation (i.e., amount, orientation), which could result from multiple labs or users are performing testing.

---

1. Accuracy:

In general, accuracy refers to the ability of a model to correctly identify a target sample. There are two components to testing accuracy for an AI test such as performed by the iVeris, both for the *Model* and *Classification* based on the model. The cloud-based CV/ML software automatically evaluates all of the physical attributes from the sample photos. Therefore, it is not possible to enumerate or describe the specific characteristics that the model is using to classify the samples, as in a morphological (i.e. green leaves), genetic (i.e. mutations), or chemical test. As a result, it is essential to ensure that the model created from the characteristics that the algorithms learned from the reference photo library is accurate. A model with a 50% accuracy is useless, while those between 80-100% are typically considered a good model. Before accuracy is calculated, an acceptance probability threshold must be specified. In this study, we calculated the accuracy rates at both 50% and 70% probabilities.

(a) *Model Accuracy*

The *Model Accuracy* refers to the ability of a model to correctly classify samples from images of specimens of known identity that *were used to train the model*. This is calculated by dividing the true positives (TP) and true negatives (TN) by the total number of samples tested (TP+TN/Total). However, in this study, we tested target and non-target samples in separate studies. In this study, one randomly selected specimen from each of the 187 target categories was tested using the iVeris model (Table 1; Appendix 1). From each specimen, a single image that was used in the training set was tested.

(b) Classification Accuracy

The *Classification Accuracy* is determined in the same way as the model accuracy (TP+TN/Total). However, it involves testing images from specimens *not used in the model training set.* For this study we tested a total of 54 samples that includes 40 target and 14 non-target (mismatch) specimens (Table 2) before they were added to the final model (that was tested above). The non-target samples do belong to species that *are* included in the non-target category in the database, however. A single image was taken from one sample of each specimen and tested using the iVeris. This is one of the most critical steps in validating a model.

3. Precision

In general, the *Precision* of a test is its ability to generate the same answer over and over, either from the same or different images of a sample, or from laboratory to laboratory. These all are calculated based on the *same preparation of the sample,* such as the same amount of material and orientation as used to train the model.

(a) Repeatability:

The *repeatability* is the ability of the model to consistently identify a specimen from repeat tests of the *same image*. In this study, a total of six specimens from different categories were tested that covered a wide range of variation, from highly processed materials to large whole objects (Table 3). For each specimen, a single image of each was randomly selected that was included in the training image database and tested using the iVeris a total of five times. The accuracy of the results (TP/Total) was calculated. Additionally, to examine the consistency in the probabilities, we also calculated the average (mean) probability and standard deviation.

(b) Extended Repeatability:

This is the ability of the model to accurately identify a specimen from repeat tests of *different images from different samples taken from the same specimen (Intra-specimen* (Table 4). In this study, the same six specimens included in the *Repeatability* tests, including the adulterated oregano sample, were used. For each specimen, a single image from five randomly selected samples was tested using the iVeris. The accuracy of the results (TP/Total) was calculated. Additionally, to examine the consistency in the probabilities, we also calculated the average (mean) probability and standard deviation. This is a particularly useful test in that it can test the homogeneity of a sample and flag samples that may contain adulterants that are not consistent from sample to sample.

Extended Repeatability also refers to a model's ability to accurately identify *different specimens within the same category (Intra-category).* In this study, a total of 25 samples representing the same six categories used above were included, with two to five samples per category (Table 5). For each specimen, a single image from five randomly selected samples was tested using the iVeris. The accuracy of the results (TP/Total) was calculated. Additionally, to examine the consistency in the probabilities, we also calculated the average (mean) probability and standard deviation. This may be the most critical test, as it helps to identify categories that may be confused with one another or have a wide range of variation within them that may require the addition of more samples and images. Alternatively, it may be used to indicate what categories may be overlapping and should be combined with one another.

Results

1. Accuracy:

(a) *Model Accuracy*

The results from the first *Model Accuracy* study are shown in Table XX below. Because Non-Target samples were not included in this analysis, the accuracy is calculated by TP/Total Samples. At a 50% probability threshold of acceptance, there were no false positives (FP), only one false negative (FN) and a Model Accuracy score of 99.5% (= 186/187). When the probability threshold is increased to 70%, there no FP, only four FN, and a Model Accuracy of 98.4% (= 183/187). Additionally, the average probability of identification (percent probability of top hit) is 95.0% (standard deviation = 7.6%).

Table 1. Results from the *Model Accuracy* study. The percent thresholds indicate the level of probabilities for acceptance.

	# Samples 50% Threshold	**# Samples 70% Threshold**
True Positive (TP)	186	183
False Positive (FP)	0	0
False Negative (FN)	1	4
Total Samples	187	187

(b) Classification Accuracy

The results from the classification accuracy test of 40 target and 14 non-target samples (Appendix 2) before they were added to the training image database are shown in Table 2 below. To calculate the classification accuracy is TP+TN/Total Samples. For TN samples, we included any sample that was identified as a mismatch, no matter what the probability because if the sample is identified as a mismatch at any level it should be rejected. The classification accuracy score at a 50% threshold is 81.5%, and 59.3% at 70% level. The POI was an average of 75.5%, significantly lower than after they were added (Appendix 2), improving to an average of 94.8%. This provides evidence that adding more samples and images to the database can increase the accuracy of the model, and that once a sample is added to the database the sample characteristics are learned so that the same sample is accurately identified when tested against the database.

Table 2. Results from the *Classification Accuracy* study. The percent thresholds indicate the level of probabilities for acceptance.

	# Samples 50% Threshold	**# Samples 70% Threshold**
True Positive (TP)	33 (61.1%)	27 (50%)
True Negative (TN)	11 (20.4%)	5 (9.3%)
False Positive (FP)	5 (9.3%)	5 (9.3%)
False Negative (FN)	5 (9.3%)	11 (20.4%)
Total Samples	54	54

3. Precision

(a) Repeatability:

The results of the *Repeatability* study are in the Table 3 below. For all six categories, they were correctly identified with high probabilities 100% of the time when the same image of the samples was tested five times. Additionally, the Probability of Identification (POI) was identical in each replicate for all samples.

Table 3. Results from the *Repeatability* study.

Category	**Sample ID #**	**True Positives**	**Average Probability of Identification (POI)**	**Standard Deviation (%)**
Dill (Anethum graveolens) Seed Whole	272	5 (100%)	99.3	0.0
Black Cohosh (Actaea racemosa) Root Cut	555	5 (100%)	99.3	0.0
Oregano (Origanum sp.) Leaf Cut with Adulterant	145	5 (100%)	97.2	0.0
Turkish/Mediterranean Oregano (Origanum onites) Leaf Whole/Cut	1205	5 (100%)	99.9	0.0
Muira Puama (Croton echioides) Bark Cut	1111	5 (100%)	99.9	0.0
Bay or Turkish Bay (Laurus nobilis) Leaf Whole	1281	5 (100%)	99.9	0.0

(b) Extended Repeatability:

Examining the repeatability when multiple images are taken from different samples of the same specimen, the repeatability is slightly less than when the same image is tested, as expected. The results of the *intra-specimen* repeatability study are in the Table 4 below. For each specimen we also examined the visual variation within each one as Low, Medium, and High to see if that corresponds to the differences in results between replicate images. As you can see below, all of the replicate images from each specimen samples were correctly identified. For the specimens that had high variation, such as those that had sticks and stems and were not homogeneous from sample to sample within the same container had lower average probabilities and high standard deviations, as expected. This test demonstrated that the model could correctly identify samples when multiple replicates were taken, and that the probability of identification variance can detect non-homogenous samples such as those with adulteration.

Table 4. *Intra-specimen repeatability* test results.

Category	**Specimen ID**	**Visual Variation**	**True Positives**	**Ave Probability**	**Standard Deviation**
Dill (Anethum graveolens) Seed Whole	272	Low	5 (100%)	97.86	2.27
Black Cohosh (Actaea racemosa) Root Cut	555	Medium	5 (100%)	95.82*/98.30 (leaves in one picture 85.9)	5.64*/1.16
Oregano (Origanum sp.) Leaf Cut with Adulterant	145 [adulterated; sticks and stems; not homogeneous]	High	5 (100%)	82.58	17.84
Turkish/Mediterranean Oregano (Origanum onites) Leaf Whole/Cut	1205	Low	5 (100%)	99.64	.38
Muira Puama (Croton echioides) Bark Cut	1111	Medium	5 (100%)	98.08	1.09
Bay or Turkish Bay (Laurus nobilis) Leaf Whole	1281	High	5 (100%)	94.08	9.38

The last study that we conducted was to test the consistency in identification between different specimens within the same category, with the results shown in Table 5 below. In five out of the six categories all of the specimens were correctly identified. One category, the adulterated oregano, had a specimen with very high variability due to sticks and stems so it was not correctly identified. Results also demonstrate that for categories with high levels of visual variability (i.e. different sized or colored objects or materials), the average probability of identification is lower and the standard deviation is higher.

Table 5. Results from the *Inter-specimen* repeatability Study. *Non-homogeneous specimen (i.e. sticks and stems in some samples).

Category	**# Samples**	**Specimen ID #**	**True Positives**	**Visual Variability**	**Ave Probability**	**Standard Deviation**
Dill (Anethum graveolens) Seed Whole	2	272, 634*	2 (100%)	Low /*High	94.85	5.59
Black Cohosh (Actaea racemosa) Root Cut	5	555, 199, 557, 1120, 1129	5 (100%)	Medium	96.78	2.82

Oregano (Origanum sp.) Leaf Cut with Adulterant	3	101, 111*, 145	2 (66.7%)	Low/High*	89.5	10.89
Turkish/Mediterranean Oregano (Origanum onites) Leaf Whole/Cut	5	105, 108, 112, 446, 1205	5 (100%)	Medium	96.96	2.71
Muira Puama (Croton echioides) Bark Cut	5	355, 1111, 1112, 1229, 1279	5 (100%)	High	76.54	26.05
Bay or Turkish Bay (Laurus nobilis) Leaf Whole	5	1281, 1284, 1285, 1286, 1287	5 (100%)	High	97.82	2.44

Discussion & Conclusions

In this study, we demonstrated that the iVeris device can accurately identify a wide range of commonly used medicinal and culinary dried herbs and spices. In fact, the *Model Accuracy* score of 99.5% across all categories in the database is likely higher than many of the traditional methods on the market. Even closely related and similar looking species that are often confused using morphological or chemical analysis were easily distinguished, such as species of oregano and echinacea. In fact, even the oregano specimens with known adulteration were distinguished from non-adulterated oregano with high probability.

The results of the *Classification Accuracy* study also demonstrate that the computer model gets more accurate as more images are included in the reference database. This is similar to a taxonomist getting smarter as they look at more and more reference materials. Unlike a human, however, the computer never forgets and can instantly access all of the reference materials that it has learned from. Additionally, the computer model eliminates human error; if results are inconclusive or unexpected, it may indicate that a sample was not properly prepared, such as the wrong amount was used, or it was in the wrong orientation.

The results in this study show that the repeatability is extremely high, as expected, because the computer model is consistent between replicates. When the same image is tested multiple times, the test results were identical. This indicates that it may not be necessary to validate this parameter in further studies. For the *Extended Repeatability* studies where different images from multiple samples of the same specimen were imaged, it showed that assuming the samples were all properly prepared, any variation that is identified in the results is an indicator of non-homogeneity, which could be the result of adulteration and may require further inspection or testing. Human mediated review of the results can be used to discount spurious results by examining the second or third best identifications.

This study successfully demonstrates how to build a highly accurate model using the iVeris device. It also is a good baseline for how to conduct multiple different validation studies. In this study, our results were based on testing only one or a few samples of each category, versus conducting validation studies for each and every category in the model. This would be ideal for regulatory compliance. However, it is a solid baseline study from which to emulate for other similar qualitative verification methods using the iVeris.

Appendix 1: List of Target Categories, the number of samples and photos in each category, as well as the Specimen ID, Scan ID, Result and Probability from the Model Accuracy study.

Category	# Specimens	# Photos	Sample ID	Scan ID#	Result	Probability
Alfalfa (Medicago sativa) Aerial Parts Cut	5	58	171	3101	Alfalfa (Medicago sativa) Aerial Parts Cut	71.3
Allspice (Pimenta officinalis) Fruit Whole	3	66	173	2996	Allspice (Pimenta officinalis) Fruit Whole	99.9
Amalaki or Amla (Phyllanthus emblica) Fruit Cut or Whole	2	33	526	2992	Amalaki or Amla (Phyllanthus emblica) Fruit Cut or Whole	94.3
American Ginseng (Panax quinquefolius) Root Whole	3	80	298	2986	American Ginseng (Panax quinquefolius) Root Whole	95.5
Angelica (Angelica archangelica) Root Cut	3	57	529	3102	Angelica (Angelica archangelica) Root Cut	96.4
Anise (Pimpinella anisum) Seed Whole	2	66	177	2999	Anise (Pimpinella anisum) Seed Whole	99.3
Annatto (Bixa orellana) Seed Whole	1	32	533	2922	Annatto (Bixa orellana) Seed Whole	100
Arnica (Arnica montana) Flower Whole	2	48	860	2914	Arnica (Arnica montana) Flower Whole	97.6
Artichoke (Cynara scolymus) Leaf Cut	3	37	1126	3054	Artichoke (Cynara scolymus) Leaf Cut	99.6
Ashwagandha (Withania somnifera) Root Cut	4	61	185	3036	Ashwagandha (Withania somnifera) Root Cut	93.4
Asian, Chinese, Korean or Red Ginseng (Panax ginseng) Root Whole	3	44	296	2983	Asian, Chinese, Korean or Red Ginseng (Panax ginseng) Root Whole	63.9
Astragalus (Astragalus mongholicus) Root Sliced (Lengthwise)	2	37	186	2916	Astragalus (Astragalus mongholicus) Root Sliced (Crosswise)	99.2
Astragalus (Astragalus mongholicus) Root Sliced (Crosswise)	3	50	188	2917	Astragalus (Astragalus mongholicus) Root Sliced (Crosswise)	99.7
Barberry (Berberis vulgaris) Root Cut	2	27	191	2921	Barberry (Berberis vulgaris) Root Cut	79.5
Basil (Ocimum basilicum) Leaf Cut	13	123	134	2982	Basil (Ocimum basilicum) Leaf Cut	76.5
Bay or Turkish Bay (Laurus nobilis) Leaf Whole	8	257	193	2969	Bay or Turkish Bay (Laurus nobilis) Leaf Whole	99.8
Bayberry (Myrica cerifera) Root Bark Cut	4	36	545	2976	Bayberry (Myrica cerifera) Root Bark Cut	86.6
Bilberry (Vaccinium myrtillus) Fruit Whole	3	71	195	3033	Bilberry (Vaccinium myrtillus) Fruit Whole	99.9
Birch (Betula sp.) Bark Cut	4	67	553	3090	Birch (Betula sp.) Bark Cut	93.7
Bitter Orange (Citrus aurantium) Peel Cut	2	21	752	2940	Bitter Orange (Citrus aurantium) Peel Cut	99.2
Black Cohosh (Actaea racemosa) Root Cut	5	77	555	2872	Black Cohosh (Actaea racemosa) Root Cut	99.2

Black Peppercorn (Piper nigrum) Fruit Whole	1	41	386	3001	Black Peppercorn (Piper nigrum) Fruit Whole	97.4
Black Seed (Nigella sativa) Seed Whole	1	30	559	2981	Black Seed (Nigella sativa) Seed Whole	99.8
Bloodroot (Sanguinaria canadensis) Root Cut	2	29	564	3019	Bloodroot (Sanguinaria canadensis) Root Cut	99.3
Blue Cohosh (Caulophyllum thalictroides) Root Cut	2	39	204	2930	Blue Cohosh (Caulophyllum thalictroides) Root Cut	98.1
Boneset (Eupatorium perfoliatum) Herb Cut	4	60	305	3103	Boneset (Eupatorium perfoliatum) Herb Cut	95.1
Brahmi (Bacopa monnieri) Herb Cut	1	16	560	2919	Brahmi (Bacopa monnieri) Herb Cut	78.7
Brown Mustard (Brassica juncea) Seed Whole	3	76	357	2925	Brown Mustard (Brassica juncea) Seed Whole	99.7
Buckthorn (Frangula alnus) Bark Cut	2	27	571	2959	Buckthorn (Frangula alnus) Bark Cut	99
Burdock (Arctium lappa) Root Cut	4	80	214	2912	Burdock (Arctium lappa) Root Cut	74.1
Calamus (Acorus calamus) Root Cut	3	51	215	3104	Calamus (Acorus calamus) Root Cut	85.6
Calendula or Marigold (Calendula officinalis) Flower Whole	2	53	216	2927	Calendula or Marigold (Calendula officinalis) Flower Whole	99.5
California Bay (Umbellularia californica) Leaf Whole	3	89	128 2	3055	California Bay (Umbellularia californica) Leaf Whole	99.7
Caraway (Carum carvi) Seed Whole	2	71	218	2929	Caraway (Carum carvi) Seed Whole	88.8
Cardamom (Elettaria cardamomum) Seed Pod Whole	3	97	219	2952	Cardamom (Elettaria cardamomum) Seed Pod Whole	99.4
Cardamom (Elettaria cardamomum) Seed Whole	2	68	221	2953	Cardamom (Elettaria cardamomum) Seed Whole	96
Carob (Ceratonia siliqua) Pod Cut	2	28	587	2932	Carob (Ceratonia siliqua) Pod Cut	96.6
Cascara Sagrada (Rhamnus purshiana) Bark Cut	3	52	224	3010	Cascara Sagrada (Rhamnus purshiana) Bark Cut	95
Cedar (Juniperus monosperma) Berry Whole	1	29	325	3105	Cedar (Juniperus monosperma) Berry Whole	99.8
Celery (Apium graveolens) Seed Whole	3	40	230	2911	Celery (Apium graveolens) Seed Whole	98.5
Ceylon or True Cinnamon (Cinnamomum verum) Bark Stick (3"& 5")	2	76	496	2938	Ceylon or True Cinnamon (Cinnamomum verum) Bark Stick (3"& 5")	99.8
Ceylon or True Cinnamon (Cinnamomum verum) Chips	1	43	610	2939	Ceylon or True Cinnamon (Cinnamomum verum) Chips	95.2
Chamomile (Matricaria chamomilla) Flower Whole	3	37	233	2974	Chamomile (Matricaria chamomilla) Flower Whole	99.8
Chervil (Anthriscus cerefolium) Leaf Cut or Crushed	5	61	862	2910	Chervil (Anthriscus cerefolium) Leaf Cut or Crushed	88
Chia (Salvia hispanica) Seed Whole	1	13	559	3015	Chia (Salvia hispanica) Seed Whole	97.9

Chicory (Cichorium intybus) Root Roasted Granules or Cut	2	50	239	2936	Chicory (Cichorium intybus) Root Roasted Granules or Cut	97.7
Chinese Licorice or Gan Cao (Glycyrrhiza uralensis) Root Sliced	3	47	709	2961	Mismatch	46.4
Chives (Allium schoenoprasum) Leaf Rings	4	103	959	2887	Chives (Allium schoenoprasum) Leaf Rings	99.6
Cinnamon (Cinnamomum sp.) Bark Chips or Chunks	2	161	607	2937	Cinnamon (Cinnamomum sp.) Bark Chips or Chunks	99.9
Cloves (Syzygium aromaticum) Flower Bud Whole	5	58	250	3030	Cloves (Syzygium aromaticum) Flower Bud Whole	99.4
Codonopsis (Codonopsis pilosula) Root Sliced	3	29	252	2943	Codonopsis (Codonopsis pilosula) Root Sliced	99.6
Comfrey (Symphytum officinale) Root Cut	3	51	618	3100	Comfrey (Symphytum officinale) Root Cut	93.5
Coriander (Coriandrum sativum) Seed Whole	3	89	259	2946	Coriander (Coriandrum sativum) Seed Whole	99.6
Corn (Zea mays) Silk Whole	2	25	261	3040	Corn (Zea mays) Silk Whole	99.8
Cornflower (Centaurea cyanus) Flower Whole	1	11	260	2931	Cornflower (Centaurea cyanus) Flower Whole	99.9
Cramp Bark (Viburnum opulus) Bark Cut	4	35	262	3106	Cramp Bark (Viburnum opulus) Bark Cut	72.2
Cumin (Cuminum cyminum) Seed Whole	3	42	626	3093	Cumin (Cuminum cyminum) Seed Whole	97.9
Curry (Murraya koenigii) Leaf Whole or Broken	3	42	265	2980	Curry (Murraya koenigii) Leaf Whole or Broken	98.3
Devil's Claw (Harpagophytum procumbens) Root Cut	5	76	271	2963	Devil's Claw (Harpagophytum procumbens) Root Cut	99
Dill (Anethum graveolens) Seed Whole	2	61	273	2899	Dill (Anethum graveolens) Weed Cut	96.2
Dill (Anethum graveolens) Weed Cut	3	87	273	3041	Dill (Anethum graveolens) Weed Cut	97.6
Dong Quai (Angelica sinensis) Root Cut	3	50	636	3042	Dong Quai (Angelica sinensis) Root Cut	75.9
Echinacea Angustifolia (Echinacea angustifolia) Root Cut	3	40	513	2950	Echinacea Angustifolia (Echinacea angustifolia) Root Cut	97
Echinacea Purpurea (Echinacea purpurea) Root Cut	5	72	514	2951	Echinacea Purpurea (Echinacea purpurea) Root Cut	97.1
Elderberry (Sambucus nigra) Flowers Whole	2	22	279	3018	Elderberry (Sambucus nigra) Flowers Whole	95.9
Elderberry (Sambucus nigra) Fruit Whole	2	70	278	3017	Elderberry (Sambucus nigra) Fruit Whole	99.7
Eleuthero (Eleutherococcus senticosus) Root Cut	3	49	282	3107	Eleuthero (Eleutherococcus senticosus) Root Cut	97.4
Emmer Wheat (Triticum dicoccon) Grain Whole	1	11	872	3044	Emmer Wheat (Triticum dicoccon) Grain Whole	99

Fennel (Foeniculum vulgare) Seed Whole	3	82	286	2957	Fennel (Foeniculum vulgare) Seed Whole	99.1
Fenugreek (Trigonella foenum-graecum) Seed Whole	2	76	287	3043	Fenugreek (Trigonella foenum-graecum) Seed Whole	99.9
Flax (Linum usitatissimum) Seed Whole	2	26	288	2970	Flax (Linum usitatissimum) Seed Whole	99.5
Fo-Ti (Fallopia multiflora) Root Cut	2	33	290	2956	Fo-Ti (Fallopia multiflora) Root Cut	98.8
Forsythia (Forsythia suspensa) Fruit Whole	2	27	657	2958	Forsythia (Forsythia suspensa) Fruit Whole	99.9
Frankincense (Boswellia sacra) Resin Chunks	2	31	659	2923	Frankincense (Boswellia sacra) Resin Chunks	99.1
Galangal (Alpinia officinarum) Rhizome Cut	1	31	660	2888	Galangal (Alpinia officinarum) Rhizome Cut	74.9
Garlic (Allium sativum) Cloves Minced	2	59	467	2886	Garlic (Allium sativum) Cloves Minced	98.7
Ginger (Zingiber officinale) Rhizome Cut or Minced	6	147	667	3039	Ginger (Zingiber officinale) Rhizome Cut or Minced	96
Ginkgo (Ginkgo biloba) Leaf Cut	5	58	295	2960	Ginkgo (Ginkgo biloba) Leaf Cut	93.2
Goji (Lycium barbarum) Berry Whole	2	28	299	2972	Goji (Lycium barbarum) Berry Whole	99.8
Goldenseal (Hydrastis canadensis) Root Cut	2	27	301	2965	Goldenseal (Hydrastis canadensis) Root Cut	98.4
Gotu Kola (Centella asiatica) Herb Cut	3	48	303	3108	Gotu Kola (Centella asiatica) Herb Cut	96.2
Grapefruit (Citrus x paradisi) Peel Cut or Minced	3	65	645	2942	Grapefruit (Citrus x paradisi) Peel Cut or Minced	96.3
Gravel Root (Eutrochium purpureum) Root Cut	3	49	305	2955	Gravel Root (Eutrochium purpureum) Root Cut	94.6
Green Peppercorn (Piper nigrum) Fruit Whole	3	92	866	3002	Green Peppercorn (Piper nigrum) Fruit Whole	99.9
Green Szechuan Peppercorn (Zanthoxylum sp.) Fruit Whole	2	61	865	3038	Green Szechuan Peppercorn (Zanthoxylum sp.) Fruit Whole	97.8
Guarana (Paullinia cupana) Seed Whole	1	10	679	2989	Guarana (Paullinia cupana) Seed Whole	99.4
Guggul (Commiphora wightii) Gum Resin Cut	1	16	680	3045	Guggul (Commiphora wightii) Gum Resin Cut	98.6
Hibiscus (Hibiscus sabdariffa) Flower Whole or Broken	2	29	311	2964	Hibiscus (Hibiscus sabdariffa) Flower Whole or Broken	99.3
Holy Basil (Ocimum tenuiflorum) Leaf Cut	4	63	683	3109	Holy Basil (Ocimum tenuiflorum) Leaf Cut	95.5
Horse Chestnut (Aesculus hippocastanum) Nut Cut	3	46	685	2883	Horse Chestnut (Aesculus hippocastanum) Nut Cut	99.5
Horsetail (Equisetum arvense) Herb Cut	3	44	687	2954	Horsetail (Equisetum arvense) Herb Cut	93.6
Hydrangea (Hydrangea arborescens) Root Cut	4	60	318	3110	Hydrangea (Hydrangea arborescens) Root Cut	98.4
Indonesian or Kointje Cinnamon, Padang or Batavia	1	217	507	3111	Indonesian or Kointje Cinnamon, Padang or	94.9

Cassia (Cinnamomum burmannii) Bark Stick (3-3.5")					Batavia Cassia (Cinnamomum burmannii) Bark Stick (3-3.5")	
Jamaican Dogwood (Piscidia sp.) Bark Cut	2	37	322	3004	Jamaican Dogwood (Piscidia sp.) Bark Cut	98.2
Jasmine (Jasminum officinale) Flower Bud Whole	2	31	324	3112	Jasmine (Jasminum officinale) Flower Bud Whole	99.9
Job's Tears (Coix lacryma-jobi) Seed Whole	1	12	695	2944	Job's Tears (Coix lacryma-jobi) Seed Whole	99.5
Juniper (Juniperus communis) Berry Whole	3	83	325	2968	Juniper (Juniperus communis) Berry Whole	99.6
Kava Kava (Piper methysticum) Root Cut	4	54	326	3000	Kava Kava (Piper methysticum) Root Cut	91.8
Kola (Cola nitida) Nut Cut	3	46	698	2945	Kola (Cola nitida) Nut Cut	94.7
Kudzu (Pueraria montana var. lobata) Root Cut	2	32	699	3008	Kudzu (Pueraria montana var. lobata) Root Cut	98.8
Lavender (Lavandula angustifolia) Flower Whole	1	11	876	9094	Lavender (Lavandula angustifolia) Flower Whole	98.5
Lemon (Citrus limon) Peel Cut	3	68	333	2941	Lemon (Citrus limon) Peel Cut	98.3
Lemon Balm (Melissa officinalis) Leaf Cut	8	113	332	3113	Lemon Balm (Melissa officinalis) Leaf Cut	79
Lemon Thyme (Thymus x citriodorus) Leaf Whole	2	67	704	3032	Lemon Thyme (Thymus x citriodorus) Leaf Whole	86.2
Lemongrass (Cymbopogon citratus) Leaf Cut	3	49	1251	3048	Lemongrass (Cymbopogon citratus) Leaf Cut	99
Licorice (Glycyrrhiza glabra) Root Cut	2	33	337	2962	Licorice (Glycyrrhiza glabra) Root Cut	98.5
Mace (Myristica fragrans) Aril Whole or Broken	3	122	1013	2978	Mace (Myristica fragrans) Aril Whole or Broken	99.9
Madder (Rubia tinctoria) Root Cut	2	38	716	3013	Madder (Rubia tinctoria) Root Cut	99.6
Maqui (Aristotelia chilensis) Berry Whole	1	13	718	2913	Maqui (Aristotelia chilensis) Berry Whole	99.4
Marjoram or Sweet Marjoram (Origanum majorana) Leaf Broken or Cut	10	131	348	3065	Marjoram or Sweet Marjoram (Origanum majorana) Leaf Broken or Cut	89.8
Marshmallow (Althaea officinalis) Root Cut	2	29	349	3095	Marshmallow (Althaea officinalis) Root Cut	98.2
Mayapple (Podophyllum peltatum) Root Cut	2	40	722	3006	Mayapple (Podophyllum peltatum) Root Cut	96.5
Mexican Oregano (Lippia graveolens) Leaf or Leaf and Flower Cut	8	179	495	2971	Mexican Oregano (Lippia graveolens) Leaf or Leaf and Flower Cut	79.3
Milk Thistle (Silybum marianum) Seed Whole	1	18	723	3028	Milk Thistle (Silybum marianum) Seed Whole	94
Mint (Mentha sp.) Leaf Cut [SPICATA]	12	282	809	3114	Mint (Mentha sp.) Leaf Cut	80.3
Moringa (Moringa oleifera) Leaf Cut	2	34	1278	3119	Moringa (Moringa oleifera) Leaf Cut	98.3
Mugwort (Artemisia vulgaris) Herb Cut	3	60	729	3116	Mugwort (Artemisia vulgaris) Herb Cut	85.7

Muira Puama (Croton echioides) Bark Cut	6	97	127 9	3117	Muira Puama (Croton echioides) Bark Cut	97.3
Mullein (Verbascum sp.) Leaf Cut	5	77	123 0	3118	Mullein (Verbascum sp.) Leaf Cut	97.4
Myrrh (Commiphora myrrha) Gum Resin Chunks	2	12	680	3120	Myrrh (Commiphora myrrha) Gum Resin Cut	98.6
Neem (Azadirachta indica) Leaf Cut	2	35	735	3088	Neem (Azadirachta indica) Leaf Cut	83.3
Nettle (Urtica dioca) Root Cut	3	55	128 6	3121	Nettle (Urtica dioca) Root Cut	81.5
Notoginseng (Panax notoginseng) Root Whole	1	94	889	2984	Notoginseng (Panax notoginseng) Root Whole	65.2
Nutmeg (Myristica fragrans) Seed Whole	3	59	364	2979	Nutmeg (Myristica fragrans) Seed Whole	100
Oat (Avena sativa) Straw Cut	2	32	123 2	3122	Oat (Avena sativa) Straw Cut	65
Oat (Avena sativa) Tops Whole	2	39	365	2918	Oat (Avena sativa) Tops Whole	99.7
Olive (Olea europaea) Leaf Cut	2	31	126 7	3123	Olive (Olea europaea) Leaf Cut	94.2
Onion (Allium cepa) Bulb Minced	2	62	500	2885	Onion (Allium cepa) Bulb Minced	98.9
Orange (Citrus sinensis) Peel Cut or Minced	4	150	121 1	3124	Orange (Citrus sinensis) Peel Cut or Minced	98.2
Oregano or Greek Oregano (Origanum vulgare) Leaf Cut	3	68	109	2995	Oregano or Greek Oregano (Origanum vulgare) Leaf Cut	93.7
Oregano (Origanum sp.) Leaf Cut with Possible Unknown Adulterant	3	61	145	3197	Oregano (Origanum sp.) Leaf Cut with Possible Unknown Adulterant	97.2
Oregon Grape (Mahonia aquifolium) Root Cut	3	44	371	2973	Oregon Grape (Mahonia aquifolium) Root Cut	99.5
Orris (Iris sp.) Root Cut	4	83	785	2967	Orris (Iris sp.) Root Cut	98.6
Parsley (Petroselinum crispum) Leaf Cut or Flakes	11	103	487	2990	Parsley (Petroselinum crispum) Leaf Cut or Flakes	73.7
Pau D'Arco (Tabebuia impetiginosa) Bark Cut	2	27	383	3031	Pau D'Arco (Tabebuia impetiginosa) Bark Cut	98.3
Pink Peppercorn (Schinus terebinthifolius) Fruit Whole	1	26	768	3023	Pink Peppercorn (Schinus terebinthifolius) Fruit Whole	99.7
Poke (Phytolacca americana) Root Cut	3	46	393	2993	Poke (Phytolacca americana) Root Cut	98
Poppy (Papaver somniferum) Seed Whole	2	47	772	2987	Poppy (Papaver somniferum) Seed Whole	98.9
Prickly Ash (Zanthoxylum clava-herculis) Bark Cut	2	35	125 3	3125	Prickly Ash (Zanthoxylum clava-herculis) Bark Cut	98.7
Pseudoginseng (Panax pseudoginseng) Root Whole	1	33	863	2985	Pseudoginseng (Panax pseudoginseng) Root Whole	97.7
Psyllium (Plantago ovata) Seed Whole	2	23	773	3005	Psyllium (Plantago ovata) Seed Whole	98

Red Clover (Trifolium pratense) Aerial Parts Cut	4	57	397	3050	Red Clover (Trifolium pratense) Aerial Parts Cut	99.6
Red Pepper (Capsicum annuum) Fruit Crushed or Flakes	4	115	894	2928	Red Pepper (Capsicum annuum) Fruit Crushed or Flakes	91.5
Red Root (Ceanothus americanus) Root Cut	3	52	121 2	3127	Red Root (Ceanothus americanus) Root Cut	73.1
Red Szechuan Peppercorn (Zanthoxylum bungeanum) Fruit Whole	2	80	824	3037	Red Szechuan Peppercorn (Zanthoxylum bungeanum) Fruit Whole	98.8
Red Winter Wheat (Triticum aestivum) Grain Whole	1	28	875	3049	Red Winter Wheat (Triticum aestivum) Grain Whole	98.9
Rose (Rosa sp.) Flower Buds or Petals Cut or Whole	3	41	402	3011	Rose (Rosa sp.) Flower Buds or Petals Cut or Whole	98.7
Rose (Rosa sp.) Hips Cut	2	50	126 6	3128	Rose (Rosa sp.) Hips Cut	99.9
Rosemary (Rosmarinus officinalis) Leaf Cracked or Whole	8	78	893	3012	Rosemary (Rosmarinus officinalis) Leaf Cracked or Whole	92.3
Rye (Secale cereale) Grain Whole	2	30	873	3025	Rye (Secale cereale) Grain Whole	99
Saffron (Crocus sativus) Stigma Whole or Broken	3	50	493	2948	Saffron (Crocus sativus) Stigma Whole or Broken	98.8
Sage (Salvia officinalis) Leaf Cut or Rubbed	7	141	408	3016	Sage (Salvia officinalis) Leaf Cut or Rubbed	82.1
Sarsaparilla or Jamaican Sarsaparilla (Smilax ornata) Root Bark Cut	5	79	694	3096	Sarsaparilla or Jamaican Sarsaparilla (Smilax ornata) Root Bark Cut	96.1
Sassafras (Sassafras albidum) Bark Cut	3	43	790	3020	Sassafras (Sassafras albidum) Bark Cut	87.2
Savory or Summer Savory (Satureja hortensis) Leaf Cut	5	152	882	3021	Savory or Summer Savory (Satureja hortensis) Leaf Cut	94.1
Saw Palmetto (Serenoa repens) Berry Cut	2	37	124 8	3129	Saw Palmetto (Serenoa repens) Berry Cut	98.7
Saw Palmetto (Serenoa repens) Berry Whole	1	17	101 7	3053	Saw Palmetto (Serenoa repens) Berry Whole	99.6
Schisandra (Schisandra chinensis) Berry Whole	3	52	125 5	3130	Schisandra (Schisandra chinensis) Berry Whole	98.8
Senna (Senna alexandrina) Leaf Whole or Cut	3	55	126 5	3131	Senna (Senna alexandrina) Leaf Whole or Cut	99.8
Sesame (Sesamum indicum) Seed Whole	1	33	799	3027	Sesame (Sesamum indicum) Seed Whole	99.7
Skullcap (Scutellaria lateriflora) Aerial Parts Cut	5	91	805	3024	Skullcap (Scutellaria lateriflora) Aerial Parts Cut	94.3
Slippery Elm (Ulmus rubra) Bark Shredded or Cut	3	44	807	3097	Slippery Elm (Ulmus rubra) Bark Shredded or Cut	95.9
Spelt (Triticum spelta) Grain Whole	1	13	874	3047	Spelt (Triticum spelta) Grain Whole	97.1

St. John's Wort (Hypericum perforatum) Aerial Parts Cut	6	49	121 7	3136	St. John's Wort (Hypericum perforatum) Aerial Parts Cut	96.8
Star Anise (Illicium verum) Seed Pod Whole	3	117	816	2966	Star Anise (Illicium verum) Seed Pod Whole	100
Stevia (Stevia rebaudiana) Leaf Cut	2	112	819	3029	Stevia (Stevia rebaudiana) Leaf Cut	96
Suma (Pfaffia paniculata) Root Cut	2	32	821	2991	Suma (Pfaffia paniculata) Root Cut	87.4
Tarragon (Artemisia dracunculus) Leaf Cut	7	107	892	2915	Tarragon (Artemisia dracunculus) Leaf Cut	93.2
Thyme (Thymus vulgaris) Leaf Cut	12	92	123 3	3135	Thyme (Thymus vulgaris) Leaf Cut	97.4
Tribulus (Tribulus terrestris) Fruit Whole	2	33	438	3048	Tribulus (Tribulus terrestris) Fruit Whole	98.2
Turkish Oregano (Origanum onites) Leaf Whole/Cut	8	193	106	2994	Turkish or Mediterranean Oregano (Origanum onites) Leaf Whole	99.6
Turmeric (Curcuma longa) Rhizome Cut	1	45	830	3056	Turmeric (Curcuma longa) Rhizome Cut	97.3
Uva Ursi (Arctostaphylos uva-ursi) Leaf Whole	2	21	444	3098	Uva Ursi (Arctostaphylos uva-ursi) Leaf Whole	98.9
Valerian (Valeriana officinalis) Root Cut	5	76	127 1	3134	Valerian (Valeriana officinalis) Root Cut	94.7
Vanilla (Vanilla planifolia) Bean Whole	9	192	868	3034	Vanilla (Vanilla planifolia) Bean Whole	98.3
Vitex (Vitex agnus-castus) Berry Whole	2	21	446	3035	Vitex (Vitex agnus-castus) Berry Whole	99.5
White Oak (Quercus alba) Bark Cut or Shredded	4	68	837	3009	White Oak (Quercus alba) Bark Cut or Shredded	98.2
White Peony (Paeonia lactiflora) Root Cut	3	53	384	3052	White Peony (Paeonia lactiflora) Root Cut	99.4
White Peppercorn (Piper nigrum) Fruit Whole	2	60	388	3003	White Peppercorn (Piper nigrum) Fruit Whole	98.3
White Poppy (Papaver somniferum) Seed Whole	1	31	512	2988	White Poppy (Papaver somniferum) Seed Whole	98.4
White Willow (Salix alba) Bark Cut	3	59	448	3014	White Willow (Salix alba) Bark Cut	98.4
Wild Cherry (Prunus serotina) Bark Cut	2	57	236	3007	Wild Cherry (Prunus serotina) Bark Cut	99.1
Wild Indigo (Baptisia sp.) Root Cut	3	46	844	3051	Wild Indigo (Baptisia sp.) Root Cut	99.8
Wild Yam (Dioscorea villosa) Root Cut	4	67	845	3099	Wild Yam (Dioscorea villosa) Root Cut	99.3
Winter Savory (Satureja montana) Leaf Cut	1	43	846	3022	Winter Savory (Satureja montana) Leaf Cut	90.6
Yellow or White Mustard (Sinapis alba) Seed Whole	2	71	359	2924	Yellow or White Mustard (Sinapis alba) Seed Whole	99.8
Yerba Santa (Eriodictyon californicum) Leaf Cut	3	55	126 4	3137	Yerba Santa (Eriodictyon californicum) Leaf Cut	92.6

Appendix 2: The results from the Classification Model test, including the category of each specimen, whether it is a Target or Non-Target Sample, the specimen ID, Scan number, Result Category and probability.

	Target (T) or Non-Target (NT)				
Category	**NT**	**Specim ent ID #**	**Sca n #**	**Result**	**Prob abilit y**
Jujube (Ziziphus jujuba) Fruit Whole	NT	1213	817	Star Anise (Illicium verum) Seed Pod Whole	56.7
Lemon Verbena (Aloysia citrodora) Leaf Cut	NT	1207	771	Mismatch	43.5
American Spikenard (Aralia racemosa) Root Cut	NT	1268	776	Mismatch	24.2
Wormwood (Artemisia absinthium) Leaf Cut	NT	1274	866	Marrubium vulgare	97.1
Shatavari (Asparagus racemosus) Root Cut	NT	1280	825	Hydrangea (Hydrangea arborescens) Root Cut	99.9
Blessed Thistle (Cnicus benedictus) Herb Cut	NT	1147	782	Mugwort (Artemisia vulgaris) Herb Cut	39.9
Papaya (Carica papaya) Leaf Cut	NT	1222	784	Mismatch	37
Motherwort (Leonurus cardiaca) Aerial Parts Cut	NT	1203	735	Lemon Balm (Melissa officinalis) Aerial Parts Cut	26.9
Rhodiola (Rhodiola rosea) Root Cut	NT	1275	743	Cardamom (Elettaria cardamomum) Seed Whole	97.9
Yellowdock (Rumex crispus) Root Cut	NT	1256	805	Mismatch	75.9
Shepherd's Purse (Capsella bursa-pastoris) Aerial Parts Cut	NT	1238	811	Mismatch	35.3
Tansy (Tanacetum vulgare) Aerial Cut	NT	1250	901	Mismatch	25.5
White Ash (Fraxinus americana) Bark Cut	NT	1259	759	Mismatch	85.7
Cat's Claw (Uncaria tomentosa) Bark Cut	NT	1181	815	Mismatch	50.7
Angelica (Angelica archangelica) Root Cut	NT	1150	774	Angelica (Angelica archangelica) Root Cut	33.4
Birch (Betula sp.) Bark Cut	NT	1156	780	Wild Cherry (Prunus serotina) Bark Cut	52.3
Boneset (Eupatorium sp.) Herb Cut	NT	1183	726	Boneset (Eupatorium sp.) Herb Cut	38
Calamus (Acorus calamus) Root Cut	NT	1135	874	Calamus (Acorus calamus) Root Cut	95.3
Cedar (Juniperus monosperma) Berry Whole	NT	1242	845	Cedar (Juniperus monosperma) Berry Whole	99.7
Comfrey (Symphytum officinale) Root Cut	NT	1152	813	Comfrey (Symphytum officinale) Root Cut	63
Cramp Bark (Viburnum opulus) Bark Cut	NT	1167	863	Cramp Bark (Viburnum opulus) Bark Cut	94.4
Eleuthero (Eleutherococcus senticosus) Root Cut	NT	1155	897	Eleuthero (Eleutherococcus senticosus) Root Cut	75.3

Gotu Kola (Centella asiatica) Herb Cut	NT	1191	829	Gotu Kola (Centella asiatica) Herb Cut	99.7
Holy Basil (Ocimum tenuiflorum) Leaf or Aerial Parts Cut	T	1243	739	Holy Basil (Ocimum tenuiflorum) Leaf or Aerial Parts Cut	77.9
Hydrangea (Hydrangea arborescens) Root Cut	T	1219	793	Hydrangea (Hydrangea arborescens) Root Cut	98.4
Cinnamon (Cinnamomum sp.) Bark Chips or Chunks	T	1137	843	Cinnamon (Cinnamomum sp.) Bark Chips or Chunks	99.2
Jasmine (Jasminum officinale) Flower Bud Whole	T	1196	797	Jasmine (Jasminum officinale) Flower Bud Whole	99.9
Lemon Balm (Melissa officinalis) Leaf Cut	T	1273	896	Lemon Balm (Melissa officinalis) Leaf Cut	83.7
Lemongrass (Cymbopogon citratus) Leaf Cut	T	1251	835	Lemongrass (Cymbopogon citratus) Leaf Cut	78.4
Marshmallow (Althaea officinalis) Root Cut	T	1202	875	Mismatch	22.7
Mint (Mentha sp.) Leaf Cut	T	1272	899	Mint (Mentha sp.) Leaf Cut	80.4
Moringa (Moringa oleifera) Leaf Cut	T	1278	849	Moringa (Moringa oleifera) Leaf Cut	90.3
Mugwort (Artemisia vulgaris) Herb Cut	T	1277	902	Mugwort (Artemisia vulgaris) Herb Cut	98.4
Muira Puama (Croton echioides) Bark Cut	T	1279	833	Muira Puama (Croton echioides) Bark Cut	98.4
Mullein (Verbascum sp.) Leaf Cut	T	1230	763	Mullein (Verbascum sp.) Leaf Cut	72.3
Neem (Azadirachta indica) Leaf Cut	T	1223	827	Lemon Verbena (Aloysia citriodora) Leaf Cut	67.2
Nettle (Urtica dioca) Root Cut	T	1236	904	Nettle (Urtica dioca) Root Cut	62.1
Oat (Avena sativa) Straw Cut	T	1232	709	Oat (Avena sativa) Straw Cut	61.6
Olive (Olea europaea) Leaf Cut	T	1267	907	Olive (Olea europaea) Leaf Cut	76.5
Orange (Citrus sinensis) Peel Cut or Minced	T	1211	788	Horse Chestnut (Aesculus hippocastanum) Nut Cut	94.4
Prickly Ash (Zanthoxylum clava-herculis) Bark Cut	T	1253	769	Prickly Ash (Zanthoxylum clava-herculis) Bark Cut	99.4
Red Root (Ceanothus americanus) Root Cut	T	1212	711	Red Root (Ceanothus americanus) Root Cut	67.3
Rose (Rosa sp.) Hips Cut	T	1266	855	Rose (Rosa sp.) Hips Cut	96.8
Sassafras (Sassafras albidum) Bark Cut	T	1246	747	Sassafras (Sassafras albidum) Bark Cut	78.8
Saw Palmetto (Serenoa repens) Berry Cut	T	1248	751	Saw Palmetto (Serenoa repens) Berry Cut	94.4
Schisandra (Schisandra chinensis) Berry Whole	T	1255	809	Schisandra (Schisandra chinensis) Berry Whole	99.4
Senna (Senna alexandrina) Leaf Whole or Cut	T	1265	749	Senna (Senna alexandrina) Leaf Whole or Cut	94.8
Slippery Elm (Ulmus rubra) Bark Shredded or Cut	T	1254	819	Marshmallow (Althaea officinalis) Root Cut	77.8
St. John's Wort (Hypericum perforatum) Aerial Parts Cut	T	1217	731	St. John's Wort (Hypericum perforatum) Aerial Parts Cut	96.4
Thyme (Thymus vulgaris) Leaf Cut	T	1233	861	Thyme (Thymus vulgaris) Leaf Cut	91.1

Uva Ursi (Arctostaphylos uva-ursi) Leaf Whole	T	1270	778	Mismatch	73.2
Valerian (Valeriana officinalis) Root Cut	T	1271	821	Valerian (Valeriana officinalis) Root Cut	85.7
Wild Yam (Dioscorea villosa) Root Cut	T	1258	720	Orris (Iris sp.) Root Cut	58.6
Yerba Santa (Eriodictyon californicum) Leaf Cut	T	1264	838	Yerba Santa (Eriodictyon californicum) Leaf Cut	99.1

References

1. Hiremath PS, P Bannigidad and S Geeta. 2010. Automated identification and classification of white blood cells (leukocytes) in digital microscopic images. *IJCA Special Issue on "Recent Trends in Image Processing and Pattern Recognition" RTIPPR*: 59-63.

2. Wäldchen J and P Mäder. 2018. Plant species identification using computer vision techniques: a systematic literature review. *Arch Computat Methods Eng*, 2025: 507–543.

3. Kumar N, PN Belhumeur, A Biswas, et al. 2012. Leafsnap: a computer vision system for automatic plant species identification. In A. Fitzgibbon et al. (Eds.): ECCV 2012, Part II, Springer-Verlag Berlin Heidelberg. 502–516.

4. Brosnan T and D-W Sun. 2017. Improving quality inspection of food products by computer vision—a review. *Agric Rev*, 38: 94-102.

5. Chopde S, M Patil, A Shaikh, et al. 2017. Developments in computer vision system, focusing on its applications in quality inspection of fruits and vegetables-a review. *Agric Rev*, 38: 94-102.

6. Harbaugh DT, Baldwin BG. — 2007. Phylogeny and biogeography of the sandalwoods (*Santalum*, Santalaceae): Repeated dispersals throughout the Pacific. *American Journal of Botany*, 94: 1028–1040.