Publisher

Maharshi Charaka Ayurveda

www.maharshicharaka.in

E-ISSN:2456-3110

Journal of Ayurveda and Integrated Medical Sciences

2025 Volume 10 Number 1 JANUARY

Research Article

Roots and Stems

Check for updates

Identification of Dry Ayurvedic Herbs (Roots and Stems) Through Ai/Computer Vision Technology

Bains R[1*], Arora R[2], Kumar S[3], Gupta R[4]

DOI:10.21760/jaims.10.1.12

[1*] Renuka Bains, Post Graduate Scholar, Dept of Dravyaguna Vigyan, Institute for Ayurveda Studies and Research, Kurukshetra, Haryana, India.

[2] Ravinder Arora, Associate Professor, Dept of Dravyaguna Vigyan, Institute for Ayurveda Studies and Research, Kurukshetra, Haryana, India.

[3] Satyender Kumar, Assistant Professor, Dept of Dravyaguna Vigyan, Institute for Ayurveda Studies and Research, Kurukshetra, Haryana, India.

[4] Rahul Gupta, Assistant Professor, Department of Electronics and Communication Engineering, University Institute of Engineering and Technology and Kurukshetra University, Haryana, Haryana India.

This study presents a novel AI-based approach for the identification and quality control of dry Ayurvedic herbs using Convolutional Neural Networks (CNNs). The model was developed to identify 43 types of dry Ayurvedic herbs, comprising 14 stems and 29 roots, based on visual features such as texture, colour, and shape. A dataset of 4,300 high-resolution images was curated using smartphones, and preprocessing techniques like normalization and augmentation were applied to enhance model robustness. The CNN model, with four convolutional layers (32 to 256 filters) and dropout layers, was designed to efficiently extract hierarchical features while preventing overfitting. The model achieved high training accuracy (94%) but encountered challenges in validation accuracy (85%), indicating difficulties in generalization. A confusion matrix revealed strong performance for distinct herb species but highlighted misclassifications among visually similar herbs. This study demonstrates the potential of AI and computer vision technologies to automate herb identification and quality control, reducing human dependency and errors. The system is deployable on mobile devices or servers, offering practical applications in the pharmaceutical and Ayurvedic industries, with significant benefits for consumer confidence and the authenticity of herbal products. Future work will focus on expanding the dataset, refining preprocessing methods, and utilizing enhanced computational resources, such as GPUs and cloud computing, to improve model scalability, efficiency, and generalization.

Keywords: Ayurvedic herbs, Convolutional Neural Networks, Image recognition, Dataset, Validation accuracy, Dry herb identification, AI, Machine learning

Corresponding Author	How to Cite this Article	To Browse
Renuka Bains, Post Graduate Scholar, Dept of Dravyaguna Vigyan, Institute for Ayurveda Studies and Research, Kurukshetra, Haryana, India. Email: renukabains2@gmail.com	Bains R, Arora R, Kumar S, Gupta R, Identification of Dry Ayurvedic Herbs (Roots and Stems) Through Ai/Computer Vision Technology. J Ayu Int Med Sci. 2025;10(1):95-105. Available From https://jaims.in/jaims/article/view/3937/	

Manuscript Received	Review Round 1	Review Round 2	Review Round 3	Accepted
2024-12-06	2024-12-16	2025-01-27	2025-01-07	2024-12-22
Conflict of Interest	**Funding**	**Ethical Approval**	**Plagiarism X-checker**	**Note**
None	Nil	Yes	11.96	

Introduction

Medicinal plants are integral to *Ayurveda*, India's ancient system of medicine, but misidentification and adulteration of herbs remain significant challenges in both local and global markets. Visually similar yet therapeutically distinct species are often mislabelled, compromising efficacy and safety.

Examples include *Polyalthia longifolia* being sold as *Saraca asoca* while *Carica papaya* seeds may be misleadingly marketed as *Piper longum*. The issue extends to internationally traded products as well. A recent data from FSSAI**[1]** India shows that *Cinnamomum cassia*, distinct from true *Cinnamomum zeylanicum*, is frequently substituted or mislabelled.

Traditional identification methods like macroscopic, microscopic, and chemical analyses**[2,3]**, while reliable, are labour-intensive. Modern technologies such as computer vision and AI offer efficient alternatives but remain underexplored for dry *Ayurvedic* herb identification.

Our research addresses this gap by developing an AI-based system using Convolutional Neural Networks (CNNs) to identify 43 types of dry Ayurvedic herbs - 14 stems or barks and 29 roots - based on visual features like texture, colour, and shape.

The methodology involved dataset creation, preprocessing (normalization and resizing), and building a CNN model with multiple convolutional, pooling, and dropout layers. Hyperparameters were optimized to maximize accuracy while preventing overfitting.

The model demonstrated strong performance during training and validation, accurately identifying unseen herb images. A detailed confusion matrix analysis highlighted strengths and areas for improvement.

The system can be deployed on mobile devices for real-time identification or on servers for large-scale analysis, offering a practical solution to herb adulteration.

By ensuring the authenticity of Ayurvedic herbs, this technology enhances consumer confidence, improves quality control in herbal markets, and supports global trade. Ultimately, it preserves the efficacy and reputation of Ayurvedic medicine worldwide.

Table 1: List Of Dry Herbs Used In This Project

SN	Name	Botanical Name	Family	Part Used
1.	Daru Haridra	Berberis aristata	Berberidaceae	Stem
2.	Palasha	Butea monosperma	Fabaceae	Stem bark
3.	Chavya	Piper retrofractum	Piperaceae	Stem
4.	Dalchini	Cinnamomum zeylanicum	Lauraceae	Stem bark
5.	Agaru	Aqualaria agallocha	Thymelaeaceae	Heart wood
6.	Arjuna	Terminalia arjuna	Combretaceae	Stem bark
7.	Kutaja	Holarrhena antidysenterica	Apocynaceae	Stem bark
8.	Shirisha	Albizia lebbeck	Mimodoideae	Stem bark
9.	Varuna	Crataeva nurvula	Capparidaceae	Stem bark
10.	Lodhra	Symplocos racemosa	Symplocaceae	Stem bark
11.	Katphala	Myrica nagi	Myricaceae	Stem bark
12.	Nimba	Azadirachta indica	Meliaceae	Stem bark
13.	Guduchi	Tinospora cordifolia	Menispermiaceae	Stem
14.	Ashoka	Saraca asoca	Fabaceae	Stembark
15.	Vidarikanda	Pueraria tuberosa	Fabaceae	Tuber
16.	Pippali	Piper longum	Piperaceae	Root
17.	Pushkarmula	Inula racemosa	Asteraceae	Root
18.	Yashtimadhu	Glycyrrhiza glabra	Fabaceae	Root
19.	Anantamula	Decalepis hamiltonii	Apocynaceae	Root
20.	Shati	Hedychium spicatum	Zingiberaceae	Root
21.	Vatsanabha	Aconitum chasmanthum	Ranunculaceae	Root
22.	Ashwagandha	Withania somnifera	Solanaceae	Root
23.	Eranda	Ricinus communis	Euphorbiaceae	Root
24.	Ativisha	Aconitum heterophyllum	Ranunculaceae	Root
25.	Bala	Sida cordifolia	Malvaceae	Root
26.	Vriddhadaru	Argyreia speciosa	Convolvulaceae	Root
27.	Musta	Cyperus rotundus	Cyperaceae	Rhizome
28.	Vacha	Acorus calamus	Araceae	Rhizome
29.	Kutaki	Picrorhiza kurroa	Scrophulariaceae	Root
30.	Kantkari	Solanum xanthocarpum	Solanaceae	Root
31.	Shatavari	Asparagus racemosus	Liliaceae	Root
32.	Jatamansi	Nardostachys jatamansi	Valerianaceae	Root
33.	Punarnava	Boerhavia diffusa	Nyctaginaceae	Root
34.	Sarpgandha	Rauvolfia serpentina	Apocynaceae	Root
35.	Haridra	Curcuma longa	Zingiberaceae	Rhizome
36.	Nishotha	Operculina turpethum	Convolvulacaceae	Root
37.	Pashanbheda	Berginia ligulata	Saxifragaceae	Root
38.	Chitraka	Plumbago zeylanica	Plumbaginaceae	Root
39.	Manjishtha	Rubia cordifolia	Rubiaceae	Root
40.	Nagarmotha	Cyperus scariosus	Cyperaceae	Root
41.	Safed Musli	Chlorophytum borivilianum	Liliaceae	Root
42.	Agnimantha	Clerodendrum phlomidis	Verbenaceae	Root
43.	Shunthi	Zingiber officinalis	Zingiberaceae	Rhizome

Materials and Methods

Materials

A comprehensive review of Ayurvedic texts, journals, and modern resources bridged traditional herb knowledge with image recognition technology. A dataset of 4,300 high-resolution images of 43 dry herbs was created using smartphones (iPhone 14, Redmi Note 10) in collaboration with Ayurvedic institutions. Image analysis and model training were supported by Intel Xeon processors, NVIDIA Tesla V100 GPUs, and software like Photoshop, TensorFlow, Keras, and Python libraries (OpenCV, scikit-learn).

Methodology

1. Data Collection: Images were collected to cover diverse views for improved model accuracy.
2. Pre-processing and Augmentation: Images were resized, normalized, and augmented for better model generalization.
3. CNN Model Development: A convolutional neural network (CNN) was trained for herb identification.
4. API and UI Development: An API and user-friendly interface enabled image uploads and classification results.
5. System Integration and Testing: The system was integrated, tested, and deployed, followed by real-world evaluations.

Model Development - Tensor Flow Code Explanation

1. Importing Libraries

- tensorflow for machine learning.
- models and layers from tensorflow.keras for neural networks.
- pyplot for plotting graphs.

2. Hyperparameters:

- Image size: 150x150 pixels
- Batch size: 50
- Channels: 3 (RGB)
- Epochs: 15

3. Loading and Pre-processing Image Data:

- The dataset is loaded using image_dataset_from_directory.
- Images are resized to 150x150 pixels and rescaled to a [0, 1] range.

4. Data Augmentation

- Random flipping and rotation enhance the model's robustness.

5. Model Architecture

Convolutional Layers: Three Conv2D layers (32, 64, and 128 filters) extract features at increasing complexity, using ReLU activation for non-linearity.
Pooling Layers: MaxPooling2D layers follow each convolutional layer to downsample feature maps, reducing spatial dimensions while preserving important information.
Fully Connected Layers:A Flatten layer converts 2D feature maps to a 1D vector, followed by a Dense layer with 512 units (ReLU activation) and a final Dense layer for class probabilities (Softmax activation).
Total Parameters: The model comprises 17,179,762 parameters, including 17,179,250 trainable and 512 non-trainable parameters.

6. Training and Evaluation

- The model uses Adam optimizer and categorical cross-entropy loss.
- It's trained over 15 epochs and evaluated using a confusion matrix.

7. Deployment

- The trained model is deployed via a Streamlit app, allowing users to upload images for herb classification.

Results

1) Model Summary

- **Convolutional Layers:** Four layers (filters 32 to 256) for feature extraction, each followed by MaxPooling.
- **Pooling Layers:** MaxPooling2D layers down-sample feature maps, reducing dimensions and complexity.
- **Dropout Layers:** Applied after MaxPooling and Dense layers to prevent overfitting.
- **Flatten Layer:** Converts 2D output to 1D for dense layer input.
- **Dense Layers:** Two layers, first with 256 units and dropout; final with 50 units for class prediction.
- **Total Parameters:** 17,179,762 (17,179,250 trainable, 512 non-trainable).

```
C:\Users\umaks\AppData\Roaming\Python\Python311\site-packages\keras\src\layers\convolutional\base_conv.
  super().__init__(activity_regularizer=activity_regularizer, **kwargs)
```

Model: "sequential_1"

Layer (type)	Output Shape	Param #
conv2d_4 (Conv2D)	(None, 256, 256, 32)	896
max_pooling2d_4 (MaxPooling2D)	(None, 128, 128, 32)	0
conv2d_5 (Conv2D)	(None, 128, 128, 64)	18,496
max_pooling2d_5 (MaxPooling2D)	(None, 64, 64, 64)	0
conv2d_6 (Conv2D)	(None, 64, 64, 128)	73,856
max_pooling2d_6 (MaxPooling2D)	(None, 32, 32, 128)	0
dropout_3 (Dropout)	(None, 32, 32, 128)	0
conv2d_7 (Conv2D)	(None, 32, 32, 256)	295,168
max_pooling2d_7 (MaxPooling2D)	(None, 16, 16, 256)	0
dropout_4 (Dropout)	(None, 16, 16, 256)	0
flatten_1 (Flatten)	(None, 65536)	0
dense_2 (Dense)	(None, 256)	16,777,472
batch_normalization_1 (BatchNormalization)	(None, 256)	1,024
dropout_5 (Dropout)	(None, 256)	0
dense_3 (Dense)	(None, 50)	12,850

Total params: 17,179,762 (65.54 MB)

Trainable params: 17,179,250 (65.53 MB)

Non-trainable params: 512 (2.00 KB)

Figure: 1

```
Epoch 1/15
85/85 ━━━━━━━━━━━━━━━━━━━━ 259s 3s/step - accuracy: 0.4020 - loss: 2.4040 - val_accuracy: 0.5860 - val_loss: 1.4686
Epoch 2/15
85/85 ━━━━━━━━━━━━━━━━━━━━ 254s 3s/step - accuracy: 0.7433 - loss: 0.9144 - val_accuracy: 0.7431 - val_loss: 0.9180
Epoch 3/15
85/85 ━━━━━━━━━━━━━━━━━━━━ 252s 3s/step - accuracy: 0.7958 - loss: 0.6984 - val_accuracy: 0.7834 - val_loss: 0.7563
Epoch 4/15
85/85 ━━━━━━━━━━━━━━━━━━━━ 258s 3s/step - accuracy: 0.8290 - loss: 0.5914 - val_accuracy: 0.7665 - val_loss: 0.7514
Epoch 5/15
85/85 ━━━━━━━━━━━━━━━━━━━━ 262s 3s/step - accuracy: 0.8532 - loss: 0.4930 - val_accuracy: 0.7919 - val_loss: 0.7321
Epoch 6/15
85/85 ━━━━━━━━━━━━━━━━━━━━ 263s 3s/step - accuracy: 0.8746 - loss: 0.4240 - val_accuracy: 0.7941 - val_loss: 0.7075
Epoch 7/15
85/85 ━━━━━━━━━━━━━━━━━━━━ 264s 3s/step - accuracy: 0.8629 - loss: 0.4297 - val_accuracy: 0.7983 - val_loss: 0.6272
Epoch 8/15
85/85 ━━━━━━━━━━━━━━━━━━━━ 262s 3s/step - accuracy: 0.8901 - loss: 0.3568 - val_accuracy: 0.8047 - val_loss: 0.6483
Epoch 9/15
85/85 ━━━━━━━━━━━━━━━━━━━━ 261s 3s/step - accuracy: 0.8910 - loss: 0.3418 - val_accuracy: 0.8195 - val_loss: 0.6326
Epoch 10/15
85/85 ━━━━━━━━━━━━━━━━━━━━ 262s 3s/step - accuracy: 0.9034 - loss: 0.3261 - val_accuracy: 0.8238 - val_loss: 0.5826
Epoch 11/15
85/85 ━━━━━━━━━━━━━━━━━━━━ 261s 3s/step - accuracy: 0.9020 - loss: 0.3211 - val_accuracy: 0.8344 - val_loss: 0.5545
Epoch 12/15
85/85 ━━━━━━━━━━━━━━━━━━━━ 260s 3s/step - accuracy: 0.9092 - loss: 0.2931 - val_accuracy: 0.8386 - val_loss: 0.5313
Epoch 13/15
...
Epoch 14/15
85/85 ━━━━━━━━━━━━━━━━━━━━ 262s 3s/step - accuracy: 0.9036 - loss: 0.2817 - val_accuracy: 0.8089 - val_loss: 0.6217
Epoch 15/15
85/85 ━━━━━━━━━━━━━━━━━━━━ 264s 3s/step - accuracy: 0.9186 - loss: 0.2750 - val_accuracy: 0.8174 - val_loss: 0.6467
Output is truncated. View as a scrollable element or open in a text editor. Adjust cell output settings...
```

Figure: 2

2) Training

The model training shows rapid improvement in early epochs, stabilizes in the mid-phase with consistent accuracy gains, and reaches peak performance by the final epochs (91.86% training, 83.86% validation).

Loss values steadily decrease, indicating strong generalization without overfitting, though a small gap suggests potential for improved generalization on unseen data.

3) Accuracy over Epochs:

- **Training Accuracy**: Steady improvement from 88% to 94% shows strong learning without signs of overfitting.
- **Validation Accuracy**: Fluctuates with a peak around 85%, indicating less stable generalization.
- **Epoch Variability**: Significant oscillations in validation accuracy during mid-to-late epochs suggest challenges in generalizing to new data.

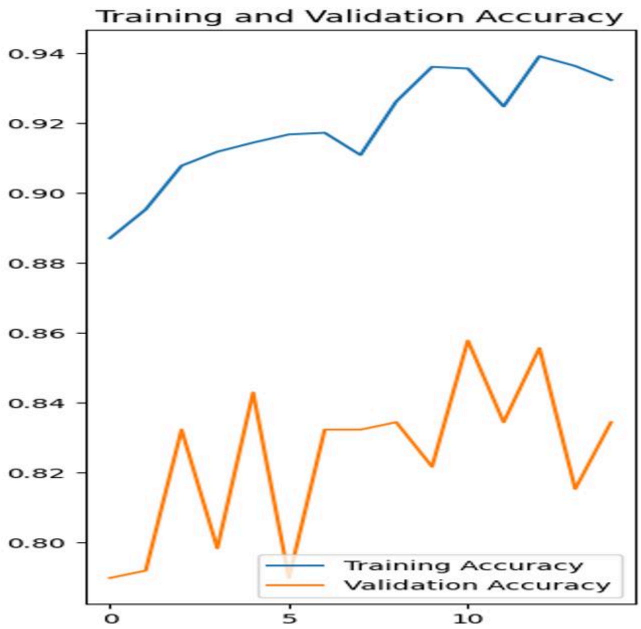

Figure: 3

4) Loss Over Epoch

- **Training Loss:** Sharp initial decrease, reaching minimal error around epochs 12-14, indicating effective learning without overfitting.
- **Validation Loss:** Starts high, fluctuates heavily, with instability persisting through epochs 3-14, showing inconsistent generalization.
- **Train-Validation Loss Discrepancy:** Early closeness diverges as training loss decreases steadily, while validation loss remains erratic, signalling challenges in generalizing to new data.

Figure: 4

5) Confusion Matrix Overview

- **Diagonal Dominance:** Strong accuracy, especially for distinct classes like *Zingiber Officinale Root* and *Berberis Aristata Stem*, with minimal misclassifications.
- **Off-Diagonal Errors:** Misclassifications primarily occur between visually similar species, like *Aconitum* varieties, indicating difficulty in differentiating closely related classes.
- **Overall Performance:** High model accuracy with room for refinement to better distinguish similar species, suggesting potential for further optimization.

6) Prediction of Model

- **Dataset Shuffling:** Images are shuffled before training, enhancing the model's learning by providing random data order.
- **Image Prediction**: The model correctly predicts the herb *Asparagus Racemosus Root* for the first image, matching its true label.
- **End of Sequence Warning:** An "OUT_OF_RANGE" warning appears, indicating that all available data has been processed.

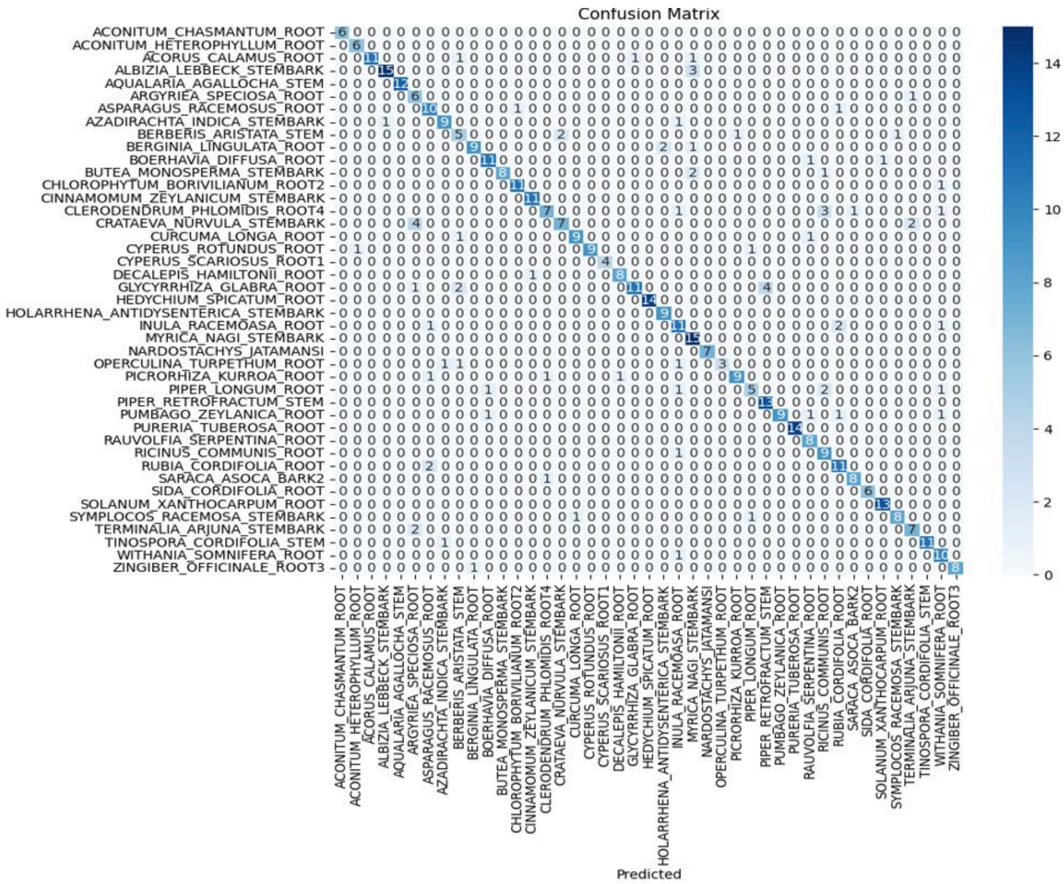

Figure: 5

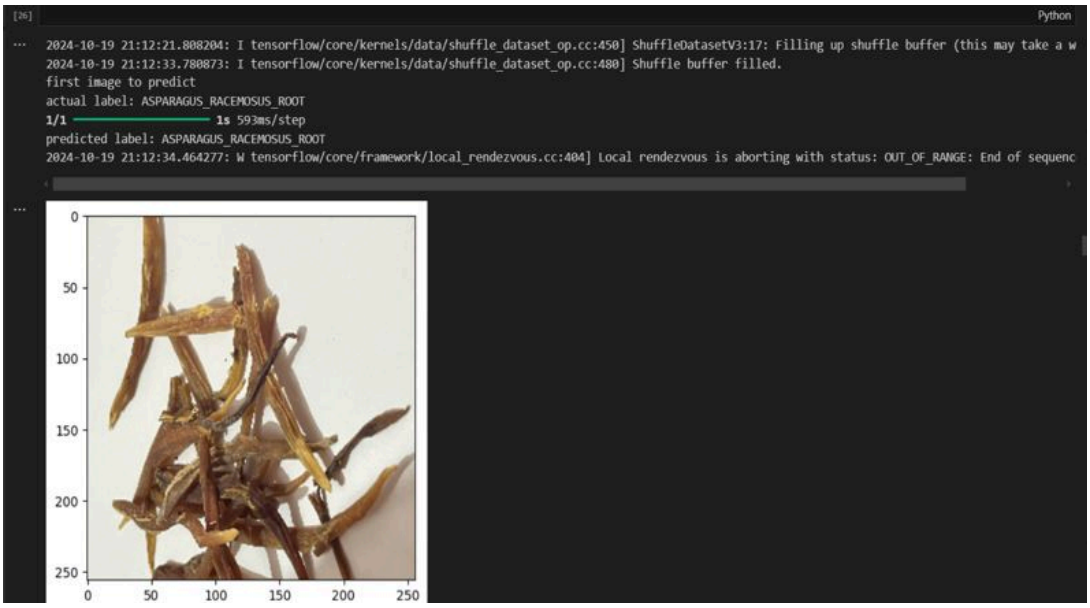

Figure: 6

Visual Predictions and Model Confidence Analysis for Dry Herbs:


ACONITUM_CHASMANTUM_ROOT
Predicted: ACONITUM_CHASMANTUM_ROOT
Confidence: 99.49%
ACONITUM_HETEROPHYLLUM_ROOT
Predicted: ACONITUM_HETEROPHYLLUM_ROOT
Confidence: 99.06%
ACORUS_CALAMUS_ROOT
Predicted: ACORUS_CALAMUS_ROOT
Confidence: 75.3%
ALBIZIA_LEBBECK_STEMBARK
Predicted: ALBIZIA_LEBBECK_STEMBARK
Confidence: 87.11%
AQUALARIA_AGALLOCHA_STEM
Predicted: AQUALARIA_AGALLOCHA_STEM
Confidence: 100.0%
ARGYRIEA_SPECIOSA_ROOT
Predicted: ARGYRIEA_SPECIOSA_ROOT
Confidence: 98.92%
ASPARAGUS_RACEMOSUS_ROOT
Predicted: ASPARAGUS_RACEMOSUS_ROOT
Confidence: 99.95%
AZADIRACHTA_INDICA_STEMBARK
Predicted: AZADIRACHTA_INDICA_STEMBARK
Confidence: 99.74%
BERBERIS_ARISTATA_STEM
Predicted: SYMPLOCOS_RACEMOSA_STEMBARK
Confidence: 44.27%

Figure: 7


BERGINIA_LINGULATA_ROOT
Predicted: HOLARRHENA_ANTIDYSENTERICA_STEMBARK
Confidence: 36.23%
BOERHAVIA_DIFFUSA_ROOT
Predicted: BOERHAVIA_DIFFUSA_ROOT
Confidence: 99.77%
BUTEA_MONOSPERMA_STEMBARK
Predicted: BUTEA_MONOSPERMA_STEMBARK
Confidence: 99.96%


CHLOROPHYTUM_BORIVILIANUM_ROOT2
Predicted: CHLOROPHYTUM_BORIVILIANUM_ROOT2
Confidence: 98.16%
CINNAMOMUM_ZEYLANICUM_STEMBARK
Predicted: CINNAMOMUM_ZEYLANICUM_STEMBARK
Confidence: 100.0%
CLERODENDRUM_PHLOMIDIS_ROOT4
Predicted: CLERODENDRUM_PHLOMIDIS_ROOT4
Confidence: 51.31%
CRATAEVA_NURVULA_STEMBARK
Predicted: ARGYRIEA_SPECIOSA_ROOT
Confidence: 88.88%
CURCUMA_LONGA_ROOT
Predicted: CURCUMA_LONGA_ROOT
Confidence: 98.08%
CYPERUS_ROTUNDUS_ROOT
Predicted: CYPERUS_ROTUNDUS_ROOT
Confidence: 99.34%

Figure: 8


CYPERUS_SCARIOSUS_ROOT1
Predicted: CYPERUS_SCARIOSUS_ROOT1
Confidence: 99.99%
DECALEPIS_HAMILTONII_ROOT
Predicted: CINNAMOMUM_ZEYLANICUM_STEMBARK
Confidence: 69.05%
GLYCYRRHIZA_GLABRA_ROOT
Predicted: PIPER_RETROFRACTUM_STEM
Confidence: 98.58%
HEDYCHIUM_SPICATUM_ROOT
Predicted: HEDYCHIUM_SPICATUM_ROOT
Confidence: 99.61%
HOLARRHENA_ANTIDYSENTERICA_STEMBARK
Predicted: HOLARRHENA_ANTIDYSENTERICA_STEMBARK
Confidence: 99.87%
INULA_RACEMOASA_ROOT
Predicted: INULA_RACEMOASA_ROOT
Confidence: 98.92%

Figure: 9

Figure: 10

Figure: 11

Discussion

The CNN model for dry *Ayurvedic* herb identification, designed with four convolutional layers (increasing filters from 32 to 256), effectively captures hierarchical visual features like texture and patterns, while MaxPooling layers reduce dimensionality and Dropout layers mitigate overfitting. With over 17 million parameters, the model balances complexity and efficiency. Performance metrics showed training accuracy improving steadily from 88% to 94%, Demonstrating effective learning, though validation accuracy stabilized around 85%, indicating challenges in generalization. The training loss steadily decreased, but validation loss exhibited fluctuations, suggesting potential overfitting or insufficient dataset diversity. A detailed confusion matrix highlighted high accuracy for certain herbs and occasional misclassifications, particularly among visually similar species, while achieving perfect classification for some herbs, underscoring the model's reliability. The CNN's architecture efficiently captures subtle visual distinctions, though additional regularization could further enhance generalization.

This model has significant practical implications, offering automation in herb identification and quality control, thereby reducing dependence on expert evaluation.

Conclusion

The study successfully developed a CNN-based model for identifying and ensuring the quality control of dry Ayurvedic herbs using visual features such as texture, colour, and shape, demonstrating strong potential with high classification accuracy and effective feature extraction through four convolutional layers. The model achieved significant training accuracy improvements and steady learning trends, though validation accuracy and loss trends highlighted challenges in generalization. The confusion matrix revealed high accuracy for distinct herbs but misclassifications among morphologically similar species, underscoring the need for advanced features and dataset refinement. Practical applications include real-time herb identification, benefiting the pharmaceutical and Ayurvedic industries by reducing human reliance and errors. Future studies should focus on expanding the dataset to include diverse visual variations, utilizing advanced preprocessing techniques, and upgrading computational infrastructure with GPUs, TPUs, or cloud platforms to enhance scalability, efficiency, and robustness, thereby modernizing Ayurvedic practices and advancing medicinal plant research.

References

1. Food Safety and Standards Authority of India (FSSAI). Guidance note on cinnamon [Internet]. CCASIA; 2014 Mar 10 [cited 2024 Oct 19]. Available from: https://www.*fssai.gov.in/upload/advisories/2018/04/5ac47df86d178Guidance_Note_Cinnamon_CCASIA_10_03_2014.pdf [Crossref][PubMed][Google Scholar]*

2. World Health Organization. Quality Control Methods for Medicinal Plant Materials. Geneva: WHO; 1992. . *[Crossref][PubMed][Google Scholar]*

3. Government of India, Ministry of Health and Welfare, Department of AYUSH. The Ayurvedic Pharmacopeia of India. New Delhi: Government of India. . *[Crossref][PubMed][Google Scholar]*