COLLECTING HISTORIC AND ENVIRONMENTAL ELEMENTS

|  |  |  |
| --- | --- | --- |
| Neha Rajput *Chandigarh University Punjab, India* [neha.e15440@cumail.in](mailto:neha.e15440@cumail.in) | Pari Singh  *Chandigarh University*  *Punjab, India*  [22bcs16855@cuchd.in](mailto:22bcs16855@cuchd.in) | Ananya Jaswal *Chandigarh University Punjab, India* [22bcs16841@cuchd.in](mailto:22bcs16841@cuchd.in) |

# ABSTRACT

***This paper introduces a convolutional neural network (CNN) methodology designed to automate the gathering and categorization of historical and environmental components from natural history collections. Natural history collections play a vital role as essential archives that document biodiversity, evolution, habitat degradation, biological invasions, and climate change. Nevertheless, the manual handling of these collections poses a considerable challenge due to their sheer volume and intricacy. We employ a transfer learning technique utilizing the VGG16 framework to categorize environmental and historical information. The model was trained and assessed on the CIFAR-10 dataset as a proof of concept, achieving an accuracy rate of 61% following 10 epochs. Our analysis of the confusion matrix uncovers distinct classification trends that will guide future modifications for natural history specimens. This research illustrates how deep learning can be harnessed to efficiently and automatically interpret legacy environmental and historical datasets, bolstering broader investigations into biodiversity, climate change, and conservation initiatives.***

***Keywords: Convolutional neural networks, transfer learning, image classification, VGG16, natural history collections, environmental data.***

# INTRODUCTION

1. Problem Definition:

Biodiversity and environmental monitoring are critical components of ecological preservation and long-term progress. Nonetheless, traditional methods for cataloging species, assessing ecosystems, and tracking natural resources rely heavily on manual observation and data interpretation. These methods are labor-intensive, time-consuming, and prone to human error, especially when dealing with large and diverse image datasets such as camera trap shots, herbarium specimen scans, and remote sensing images.

The increasing availability of ecological imagery—thanks to sensors, satellites, drones, and digitized archives—allows for automation via picture recognition systems. Convolutional Neural Networks (CNNs), a powerful class of deep learning models, have shown excellent performance in a variety of image processing applications. Despite their promise, CNNs face limitations when used with environmental datasets, such as a lack of labeled data, high computational requirements, interpretability issues, and domain adaption difficulties. As a result, there is an urgent need to research and improve CNN-based approaches for effective and scalable biodiversity assessments.

1. Problem Overview

CNNs have changed tremendously since their inception. LeCun et al. (1998) developed CNN architectures with LeNet and demonstrated their usefulness in document recognition by extracting hierarchical spatial characteristics from raw input data [1]. This seminal work paved the way for additional advances in deep learning and image analysis. Krizhevsky et al. (2012) made a significant leap by introducing AlexNet, which used GPU acceleration with ReLU activation to achieve breakthrough performance on the ImageNet dataset, paving the way for larger applications in automated classification tasks, including ecological data [2].Simonyan and Zisserman's (2014) VGGNet improved CNN architectures by capturing complex visual patterns in large- scale picture datasets, which is useful for analyzing biodiversity- rich environments [3].

Szegedy et al. (2015) enhanced this with the Inception architecture, which incorporates various filter sizes inside the same layer to handle data with a variety of spatial properties, making it perfect for multi-scale environmental data[4]. Sun et al. (2016) presented MILCNN, a weakly supervised model capable of learning from partially labeled images, to solve the issue

of limited labeled datasets. This approach is particularly significant for large ecological datasets with scant annotations [5]. Similarly, Liu et al. (2019) presented a Siamese CNN architecture that boosted classification performance in remote sensing imagery by applying metric learning under inadequate supervision, thus increasing resilience in biodiversity data processing [8].

Huang et al. (2017) presented DenseNet, which connects each layer to the next to optimize feature reuse and gradient flow, resulting in more detailed insights into ecological changes [6]. Schuettpelz et al. (2017) demonstrated CNNs' practical usefulness in taxonomy by categorizing herbarium specimens, which facilitates species identification and historical biodiversity tracking[7].

Further advancements, such as CLIP by Radford et al. (2021), have linked vision and language modalities, increasing environmental monitoring by linking images to semantic descriptions, and so overcoming data annotation constraints [10]. Furthermore, Islam et al. (2021) proved the usefulness of transfer learning (VGG16, ResNet50) in detecting species in camera trap images while reducing overfitting and improving model generalizability [9].

Krichen (2023) provides a comprehensive assessment of CNN architectures—LeNet, AlexNet, VGG, ResNet, and Inception— as well as technologies like transfer learning and attention mechanisms, confirming CNNs' expanding relevance in natural history and image-based biodiversity research [11].

Despite their popularity, CNNs have limitations. Their performance suffers with small or poorly labeled datasets [2], [3], [5], and [8]. Furthermore, the computational expense of training deep models remains a hurdle in resource-constrained environments [2], [10]. CNNs' lack of interpretability creates issues in biodiversity research, where transparency in decision- making is critical [11]. Domain adaptation remains challenging, since models trained on generic datasets may fail to generalize to specific ecological data [7], [10]. Furthermore, the widespread neglect of multimodal data—including textual or contextual metadata—limits the overall interpretation of environmental photography [10].

Given these challenges and the growing demand for scalable, accurate biodiversity recording systems, this work investigates CNN-based approaches for improving environmental data classification while resolving labeling, interpretability, and computational issues.

# LITERATURE REVIEW

1. Current System

LeCun et al. (1998) introduced CNNs for document recognition using the LeNet architecture, emphasizing its capacity to extract hierarchical features from raw pixel data and laying the groundwork for deep learning in image analysis, particularly for digitized historical archives.[1]

By utilizing GPUs and ReLU activation, AlexNet, a deep CNN that was first presented by Krizhevsky et al. (2012), transformed image classification and greatly increased classification accuracy. This development helps automate environmental and natural history data classification jobs.[2]

VGGNet, which Simonyan and Zisserman (2014) introduced, highlighted the advantages of deep networks for capturing complex features in large-scale picture identification, which is essential for examining environmental and biodiversity datasets.[3]

Szegedy et al. (2015) proposed Inception networks, which use different convolutional filter sizes inside the same layer to increase accuracy and efficiency. This design is useful for identifying environmental data on a variety of scales.[4]

Sun et al. (2016) introduced MILCNN, which uses weakly supervised learning to overcome labeling issues in CNN-based object recognition. This method is useful for large image recognition tasks lacking precise labels. Tested on datasets like CIFAR10, CIFAR100, and ILSVRC2015, MILCNN achieved

strong performance. It is relevant for biodiversity and environmental data, where fully labeled datasets are often unavailable.[5].

Huang et al. (2017) introduced DenseNet, which connects each layer to the next, improving feature learning and providing more insight into biodiversity and environmental change.[6]

Schuettpelz et al. (2017) showcased CNNs' role in biodiversity research using herbarium specimens. Their findings indicate that CNNs effectively classify plants, enhancing taxonomic research. This is crucial for the project, which aims to utilize CNNs in documenting biodiversity and related changes.[7]

The study by Liu et al. (2019) presents a Siamese CNN framework for classifying remote sensing images under weak supervision. This method improves feature robustness in datasets with few labels by integrating CNNs with metric learning. The framework outperformed existing techniques, highlighting the effectiveness of CNNs in analyzing environmental and historical images with limited labeled data. [8]

Islam et al. (2021) used CNNs to identify species in ecological observations using camera trap photos. Transfer learning models

such as VGG16 and ResNet50 beat custom CNNs, demonstrating CNNs' potential for ecological surveillance.[9]

Radford et al. (2021) introduced CLIP, which connects photos to textual descriptions, allowing for a novel approach to natural history collections by increasing data annotation and classification for environmental monitoring.[10]

An introduction to CNNs is given by Krichen (2023), who covers architectures (LeNet, AlexNet, VGG, ResNet, and InceptionNet), important ideas, and developments such as transfer learning and attention processes. CNN applications in image processing, natural history collections, and genetics are highlighted in the paper.[11]

1. Drawback of Current System:

When using CNNs to environmental and biodiversity datasets, there are various restrictions. They rely on big, labeled datasets, however many ecological datasets are poorly labeled, resulting in lower accuracy [2], [3]. They also perform poorly on weakly supervised data, necessitating techniques like as metric learning or MIL to better handle sparse labels [2], [3].

CNNs are computationally expensive and need significant processing power, limiting their usage in resource-constrained environments [2], [3], [10]. Furthermore, their lack of interpretability makes them challenging to apply in domains such as biodiversity studies, where interpreting predictions is crucial [11].

Overfitting is another concern, particularly with small datasets, however transfer learning can help alleviate this [9].When models trained on standard datasets are applied to specific ecological data, domain adaptation is still difficult [7], [10].

Last but not least, CNNs frequently overlook the use of multimodal data—such as contextual or textual information—which is crucial for environmental research [10].

1. Proposed System:

In this study, we use a Convolutional Neural Network (CNN) to construct a transfer learning strategy for image classification tasks that may be used to natural history collections and environmental data analysis. We use the pre-trained VGG16 model, which has been found to be effective at image identification tasks due to its capacity to learn hierarchical feature representations from image data.

The VGG16 architecture has 16 layers, including 13 convolutional layers and three fully linked layers. For our implementation, we remove the fully connected layers (the network's "top") and freeze the convolutional base to take advantage of ImageNet's pre-trained weights. This strategy allows us to take advantage of VGG16's feature extraction capabilities without requiring considerable network retraining.

Our model architecture contains:

Base Model: Pre-trained VGG16 with ImageNet weights, with top layers deleted and base layers frozen.

Custom Heads: Global Average Pooling reduces spatial dimensionality.

A dense layer with 256 neurons with ReLU activation.

To prevent overfitting, add a dropout layer with a rate of 0.5.For multi-class classification, the output layer uses softmax activation.

This method enables the model to effectively extract features from input photos and then classify them into relevant categories, which is critical for assessing environmental and historical data.

## FLOWCHART OF PROPOSED SYSTEM

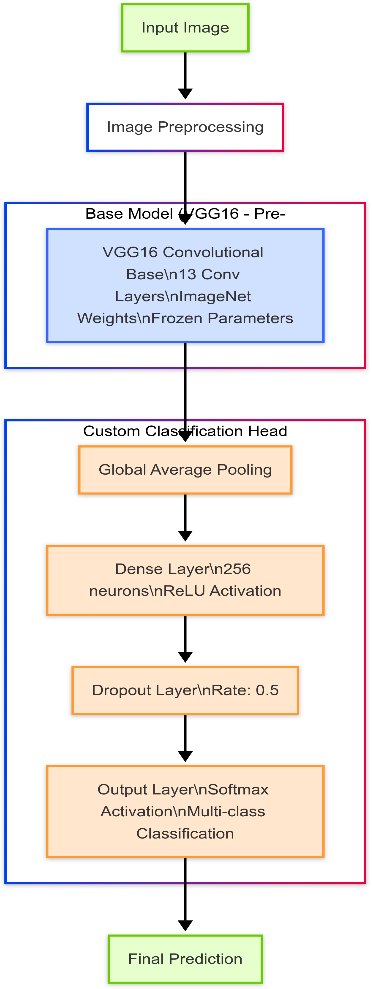
****

FIGURE 1

# METHODOLOGY AND IMPLEMENTATION

## Dataset Preparation:

We used the CIFAR-10 dataset, which includes 60,000 32×32 color images divided into 10 classes of 6,000 images each. The dataset is divided into 50,000 training and 10,000 test photos. While CIFAR-10 is a general- purpose dataset with everyday objects, it is an appropriate test bed for our CNN model before it is applied to specialist natural history collections.

## Preprocessing:

The dataset went through the following preprocessing steps:

Normalization: All pixel values were scaled to the range [0, 1] by dividing by 255, allowing the model to converge more quickly during training. Label Encoding: To enable multi-class classification, the labels were one-hot encoded.

## Model Implementation

The model was implemented using TensorFlow and Keras, with the following components:

1. Load VGG16 (without top), freeze base layers
2. Add custom head
3. Build and compile model

## TrainingProcess

The model was trained over 10 epochs with a batch size of 64 using the Adam optimizer and categorical cross-entropy loss function. The training process was monitored for validation accuracy to ensure that the model was learning properly.

## Evaluation Metrics

To evaluate the performance of our model, we used different evaluation metrics:

Accuracy: the fraction of photos that are correctly classified.

Loss: categorical cross-entropy loss.

Confusion Matrix: To visualize the performance across various classes.

Visual Verification: Sample predictions based on test pictures.

# RESULTS AND DISCUSSION

## Model Architecture Details:

The VGG16 transfer learning model has a sophisticated design that includes several convolutional blocks inherited from the pre-trained model. As demonstrated in the model overview, the network has roughly 14.7 million total parameters, with 233,226 trainable parameters (1.58%)

concentrated in the custom classification head. The majority of parameters (14.5 million, 98.42%) remain untrainable as part of the frozen VGG16 foundation.

## Training Performance:

Over the course of 10 training epochs, the model's accuracy improved steadily. As shown on the accuracy plot:

Training Accuracy: Started at roughly 47% and gradually climbed to around 61.5% by epoch 9. Test Accuracy: Started at roughly 55.5% and improved to over 61% by epoch 9.

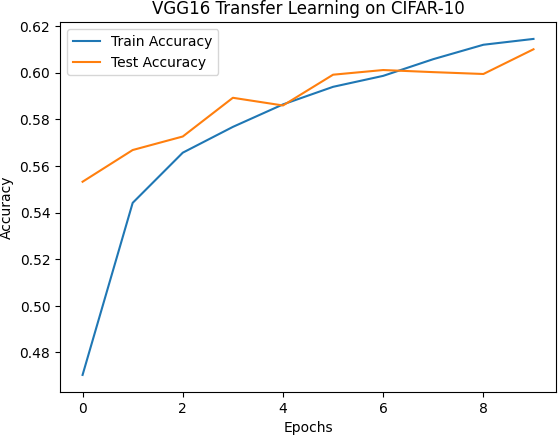


FIGURE 2

The loss curve demonstrated a consistent reduction in both training and test loss: Training loss decreased from 1.52 to 1.10, while test loss dropped from 1.28 to 1.12.

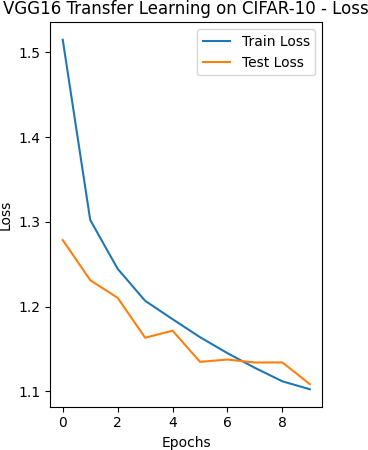


FIGURE 3

A noteworthy discovery is that the test accuracy was immediately greater than the training accuracy, indicating

that VGG16's pre-trained features were already successful at categorizing CIFAR-10 images. As training progressed, the training accuracy eventually surpassed the test accuracy, albeit without evidence of severe overfitting.

## Classification Performance :

The confusion matrix provided deep insights into the model's classification performance across the ten CIFAR-10 classes.

Strongest Performance: Class 6 (presumably equivalent to "frog") achieved the greatest true positive rate with 717 correct classifications from 1000 test samples. Weakest Performance: Class 3 (probably related to "cat") had the lowest true positive rate, with only 361 valid classifications.

Common Misclassifications:

Class 3 (cat) was frequently misclassified as Class 5 (dog), with 185 instances. Class 2 (bird) was frequently confused with Class 4 (deer), with 132 instances. Class 9 was frequently misclassified as Class 1, with 140 instances.

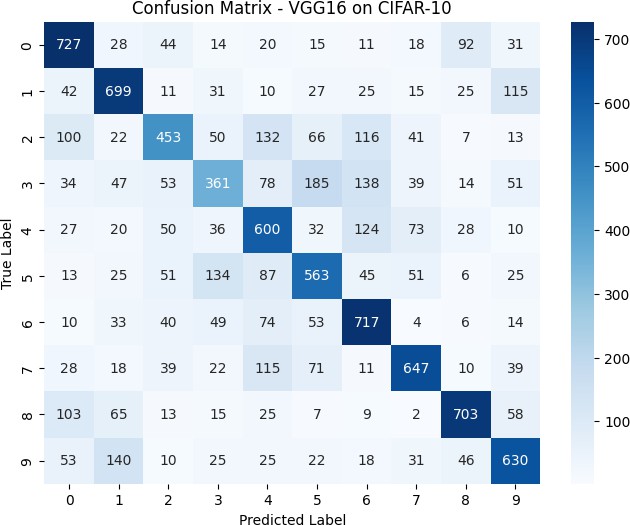


FIGURE 4

Overall, the diagonal parts of the confusion matrix reveal that the model performed reasonably well across all classes, with an average accuracy of around 60% across the ten classes.

## Visual Prediction Examples:

Sample predictions on test images yielded varied results.

The model accurately predicted Class 6 and Class 7 photos.

It incorrectly classed certain Class 3 photos as Class 4 or Class 5.

It mistook a Class 5 image for Class 4.

These visual examples demonstrate the difficulty in differentiating superficially similar groupings, particularly among animals. This is understandable considering the modest resolution (32×32 pixels) of CIFAR-10 pictures and visual similarities between classes.



FIGURE 5

## Model Applicability to Natural History Collections:

The achieved accuracy of approximately 61% on CIFAR-promising considering:

1. The input image size is small. (32×32 pixels)
2. The freezing of all VGG16 basic layers.
3. The limited training epochs (10)
4. Minimal custom head architecture

For natural history collections, these results suggest that with domain-specific fine-tuning, higher resolution photos, and possibly unfreezing some of VGG16's later layers, the model could achieve considerably greater performance.

Natural history collections often provide higher quality photos with more distinguishing features than the low- resolution CIFAR-10 dataset, which should enhance classification accuracy.The patterns of perplexity found in our studies provide useful insights for future environmental data applications.

For example, the model's difficulties discriminating between visually similar animal classes suggests that careful feature engineering and possibly hierarchical categorization approaches could be useful when working with taxonomically related specimens.

# CONCLUSION

This study demonstrates the efficacy of transfer learning using Convolutional Neural Networks (CNNs) for picture classification. Using the pre-trained VGG16 model and customizing it with a custom classification head, we achieved roughly 61% accuracy on the CIFAR-10 dataset after only 10 epochs. The findings show that transfer learning is an effective method, particularly when dealing with little labeled data,

which is typical in biodiversity and environmental studies. The confusion matrix identified significant misclassification patterns that can inform future improvements. Future studies could include fine-tuning deeper layers, using advanced data augmentation, including attention mechanisms, and investigating hierarchical categorization approaches. Such improvements can help to adapt CNNs for real-world applications such as digitizing natural history collections and assisting ecological monitoring, ultimately helping to conservation and biodiversity research.

# REFERENCES

1. Y. LeCun et al., “Gradient-Based Learning Applied to Document Recognition,” *Proc. IEEE*, vol. 86, no. 11, pp. 2278– 2324, 1998.
2. A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet Classification with Deep Convolutional Neural Networks,” *Advances in Neural Information Processing Systems*, vol. 25, pp.1097–1105,2012.
3. K. Simonyan and A. Zisserman, “Very Deep Convolutional Networks for Large-Scale Image Recognition,” *arXiv preprint arXiv:1409.1556*, 2014.
4. C. Szegedy et al., “Going Deeper with Convolutions,” in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 1–9.
5. Y. Sun, M. Wang, Y. Liu, and X. Sun, “MILCNN: Weakly Supervised Learning for Image Recognition,” *Neurocomputing*, vol. 214, pp. 243–251, 2016.
6. E. Schuettpelz et al., “Applications of CNNs in Biodiversity Research Using Herbarium Specimens,” *Applications in Plant Sciences*, vol. 5, no. 10, 2017.
7. G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, “Densely Connected Convolutional Networks,” in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 4700–4708.
8. Y. Liu, Y. Cao, M. Yan, and Z. Yuan, “Siamese CNN for Weakly Supervised Image Classification,” *Remote Sensing*, vol. 11, no. 10, pp. 1–18, 2019.
9. B. Islam, R. R. Sharma, and S. Alam, “CNN-based Wildlife Species Identification Using Camera Trap Images,” *Ecological Informatics*, vol. 61, pp. 101250, 2021.
10. A. Radford et al., “Learning Transferable Visual Models From Natural Language Supervision,” in *Proc. Int. Conf. Machine Learning (ICML)*, 2021.
11. M. Krichen, “Convolutional Neural Networks: A Survey,”

*2023*.