

Project Presentation



ANIMAL SPECIES DETECTION

Group No.: 09

Name of the Student(s) with Regd. No.:

Subhankar Dash : 2241004192

Suvam Muduli : 2241013373

Subhankar Mishra : 2241013287

**Centre of Artificial Intelligence & Machine Learning
Computer Sc. and Engineering**

**Faculty of Engineering & Technology (ITER)
Siksha 'O' Anusandhan (Deemed to be) University
Bhubaneswar, Odisha**

Presentation Outline

- Introduction
 - Project Overview
 - Problem Statement & Motivations
- Background / Literature Review
 - Evolution of CNN in Image Classification
 - Comparative Studies on CNN architecture
- Work Accomplished
 - Workflow Diagram
 - Training and Validation Accuracy
 - Graphical representation of loss and accuracy
 - Comparison of sample prediction
 - Training Performance and Confusion Matrix
 - Combined Precision, Recall, F1-Score, and Accuracy Table
- Result Analysis
- Summary
- Bibliography

Introduction

■ Project Overview

- Build a CNN-based system for categorizing images of animals into 10 specific species (butterfly, cat, chicken, cow, dog, elephant, horse, sheep, spider, squirrel) utilizing the Animals-10 dataset (~28,000 images).
- Train and compare three CNN models—ZFNet (lightweight custom model), VGG16 (deep pre-trained network), and GoogLeNet (power-efficient Inception-based model)—with standardized data preprocessing (resizing, normalization).
- Examine performance using accuracy, loss metrics, and confusion matrices to decide which model works best for practical wildlife monitoring applications.

Introduction

■ Problem Statement & Motivations

Automated animal species identification from images remains challenging due to variations in lighting, poses, and complex backgrounds, while manual methods are time-consuming and error-prone. This project develops a CNN-based solution using **ZFNet**, **VGG16**, and **GoogLeNet** architectures to enable accurate, real-time classification of 10 animal species, addressing critical needs in wildlife conservation by providing scalable monitoring tools to support biodiversity preservation and ecological research.

Background & Related Work

■ Evolution of CNN in Image Classification

- **AlexNet (2012):** The first CNN to win the ImageNet competition, introducing ReLU activation and dropout layers for better training.
- **ZFNet (2014):** Improved AlexNet by adjusting filter sizes and strides, enhancing feature visualization.
- **VGGNet (2014):** Demonstrated that deeper networks with small (3×3) convolutions improve accuracy.
- **GoogLeNet (Inception v1, 2014):** Introduced Inception modules, optimizing computation while maintaining accuracy.

Background & Related Work

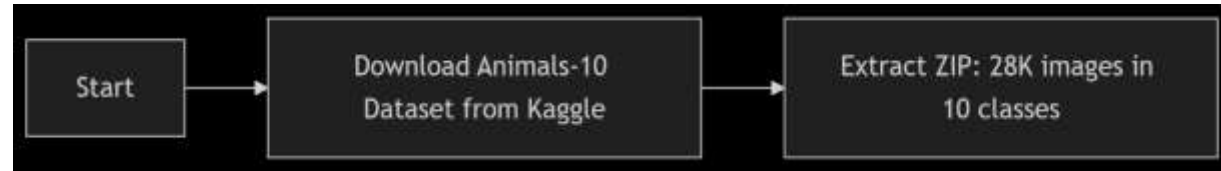
- Comparative Studies on CNN Architectures

- **VGG16 vs. ResNet:** VGG16 is accurate but computationally expensive, whereas ResNet applies skip connections for deeper networks.
- **Efficiency of GoogLeNet:** Its Inception modules have fewer parameters with the same level of performance, thus being well-suited for real-time contexts.
- **Trade-offs in Model Selection:** Shallow networks (e.g., ZFNet) are faster to train but can underperform when dealing with complex datasets.

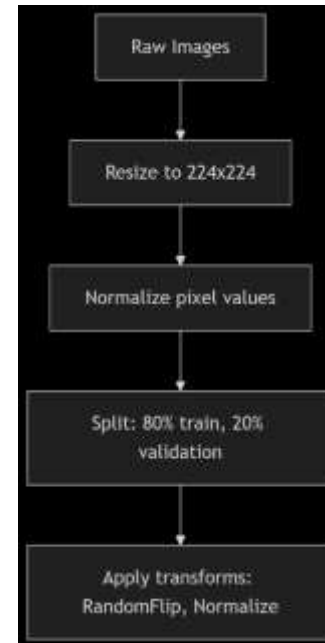
Work Accomplished

- Workflow Diagram

(Dataset Acquisition)



(Data Preprocessing)



Work Accomplished

- Workflow Diagram



(End-to-End Workflow)

Work Accomplished

■ Model Comparison (Training Acc. and Validation Acc.)

Training ZFNet

```
Epoch 1/10 => Train Acc: 0.2352, Val Acc: 0.3010
Epoch 2/10 => Train Acc: 0.3430, Val Acc: 0.3871
Epoch 3/10 => Train Acc: 0.4183, Val Acc: 0.4664
Epoch 4/10 => Train Acc: 0.4779, Val Acc: 0.4950
Epoch 5/10 => Train Acc: 0.5058, Val Acc: 0.5384
Epoch 6/10 => Train Acc: 0.5378, Val Acc: 0.5498
Epoch 7/10 => Train Acc: 0.5702, Val Acc: 0.5474
Epoch 8/10 => Train Acc: 0.5979, Val Acc: 0.5802
Epoch 9/10 => Train Acc: 0.6214, Val Acc: 0.5817
Epoch 10/10 => Train Acc: 0.6491, Val Acc: 0.5888
Training complete in 29m 47s
```

(ZFNet achieved 58.9% accuracy (slowest) with 10 epochs)

Training GoogLeNet

```
Epoch 1/10 => Train Acc: 0.9008, Val Acc: 0.9473
Epoch 2/10 => Train Acc: 0.9359, Val Acc: 0.9515
Epoch 3/10 => Train Acc: 0.9377, Val Acc: 0.9572
Epoch 4/10 => Train Acc: 0.9422, Val Acc: 0.9576
Epoch 5/10 => Train Acc: 0.9443, Val Acc: 0.9561
Epoch 6/10 => Train Acc: 0.9466, Val Acc: 0.9557
Epoch 7/10 => Train Acc: 0.9454, Val Acc: 0.9589
Epoch 8/10 => Train Acc: 0.9455, Val Acc: 0.9578
Epoch 9/10 => Train Acc: 0.9487, Val Acc: 0.9513
Epoch 10/10 => Train Acc: 0.9475, Val Acc: 0.9608
Training complete in 18m 56s
```

(GoogLeNet performed best with 96.1% accuracy (fastest)

Work Accomplished

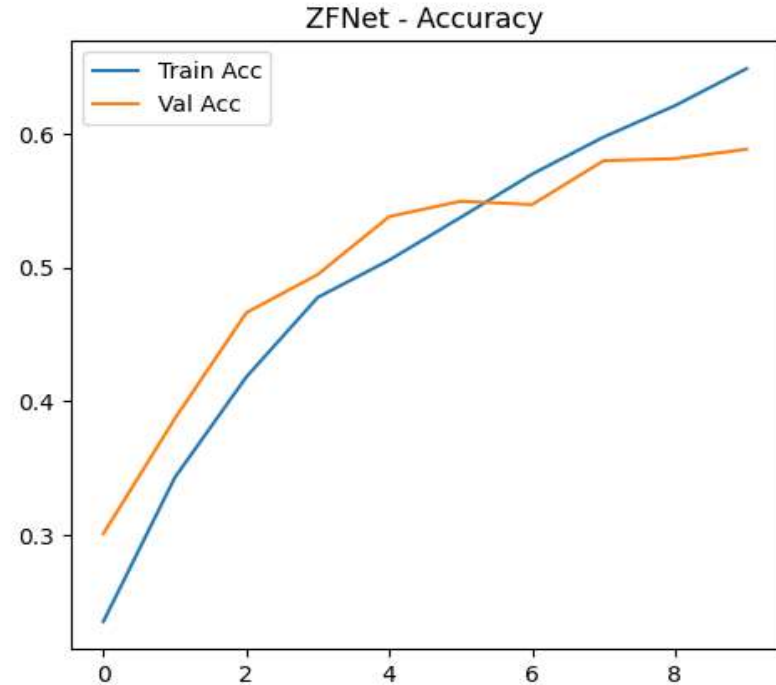
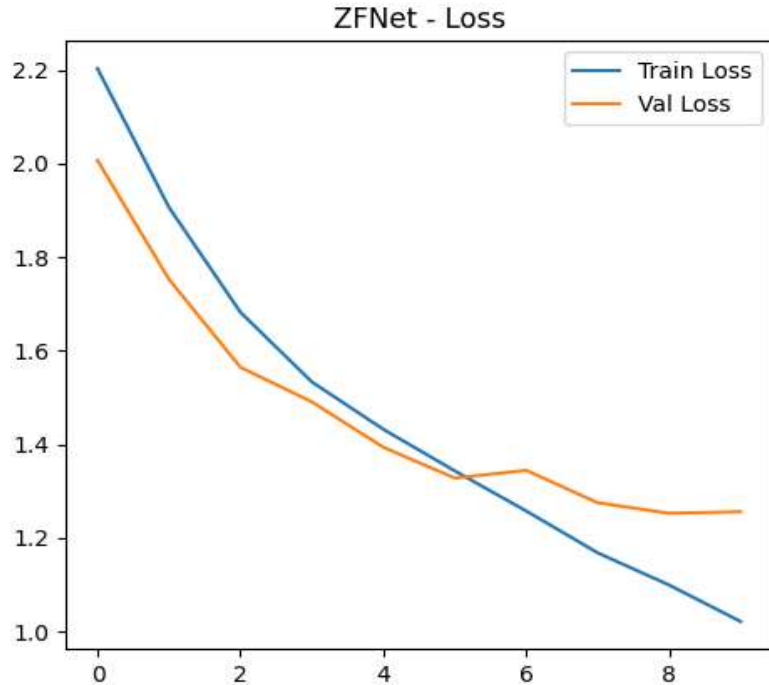
- Model Comparison

```
Training VGG16
Epoch 1/10 => Train Acc: 0.8712, Val Acc: 0.9253
Epoch 2/10 => Train Acc: 0.9196, Val Acc: 0.9270
Epoch 3/10 => Train Acc: 0.9343, Val Acc: 0.9465
Epoch 4/10 => Train Acc: 0.9459, Val Acc: 0.9269
Epoch 5/10 => Train Acc: 0.9574, Val Acc: 0.9433
Epoch 6/10 => Train Acc: 0.9589, Val Acc: 0.9530
Epoch 7/10 => Train Acc: 0.9638, Val Acc: 0.9435
Epoch 8/10 => Train Acc: 0.9691, Val Acc: 0.9502
Epoch 9/10 => Train Acc: 0.9712, Val Acc: 0.9496
Epoch 10/10 => Train Acc: 0.9759, Val Acc: 0.9532
Training complete in 39m 25s
```

(**VGG16** reached 95.3% (most stable))

Work Accomplished

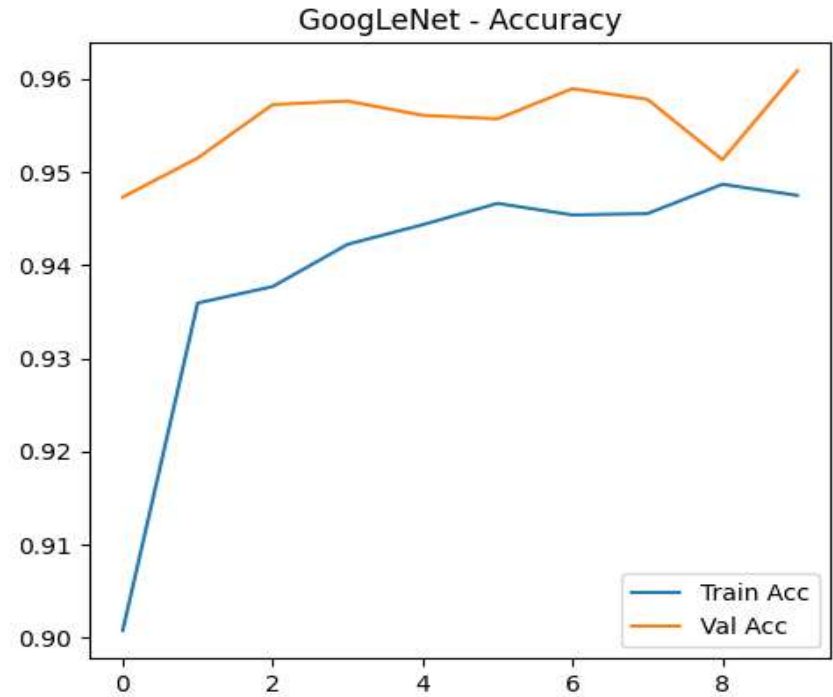
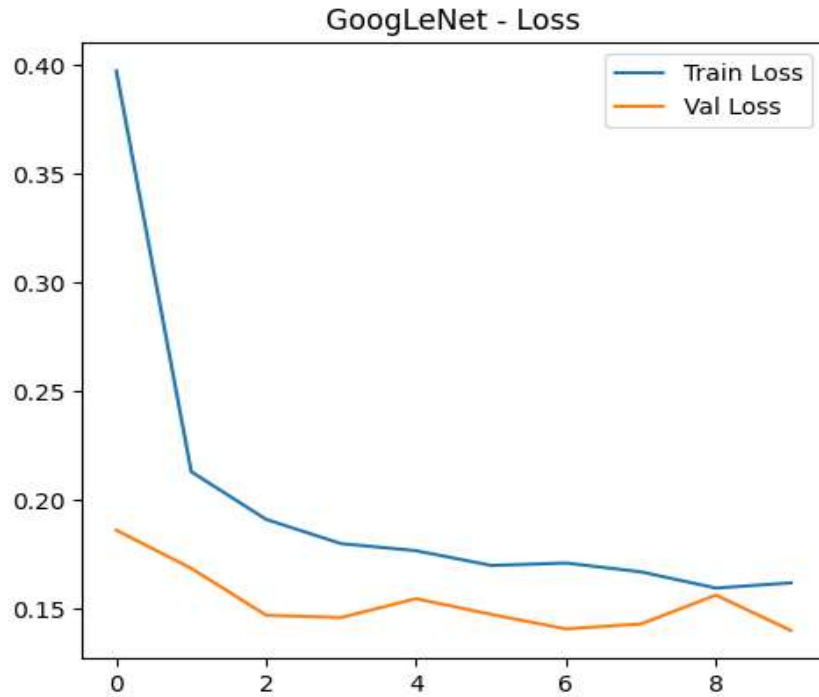
- Model Comparison (Graphical representation with loss and accuracy)



ZFNet: Shows consistent learning with decreasing loss and increasing accuracy, though there's a slight gap between training and validation performance.

Work Accomplished

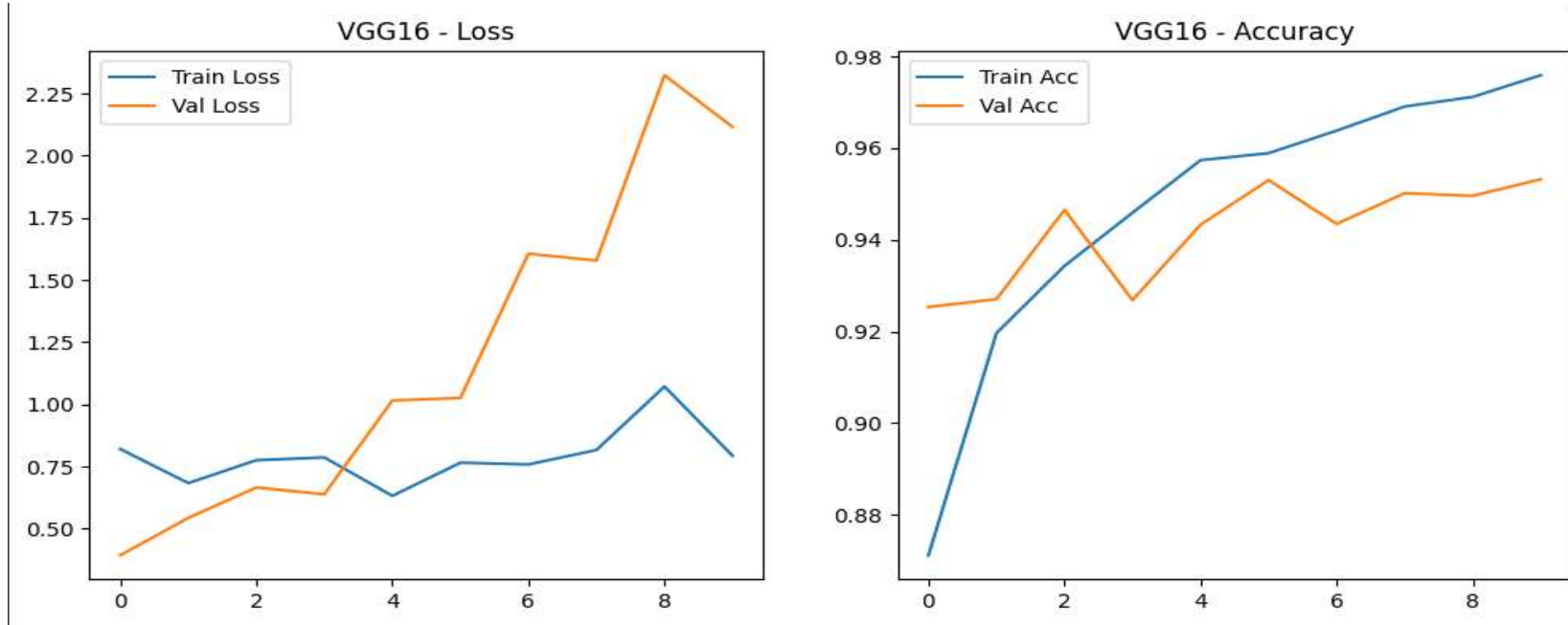
- Model Comparison (Graphical representation with loss and accuracy)



GoogLeNet: Exhibits excellent generalization with both training and validation loss decreasing and high accuracy achieved early and sustained.

Work Accomplished

- Model Comparison (Graphical representation with loss and accuracy)



VGG16: Suffers from overfitting, with training loss decreasing while validation loss increases and validation accuracy fluctuates

Work Accomplished

- Model Comparison (Visualize sample prediction of GoogLeNet)

Predicted: cat



Predicted: cat



Predicted: butterfly



Predicted: cow



Predicted: cow



Predicted: cow



Predicted: butterfly



Predicted: spider



Work Accomplished

■ Model Comparison (Comparison of sample prediction)



ZNet: Mixed accuracy with misclassifications (e.g., horse→cat)



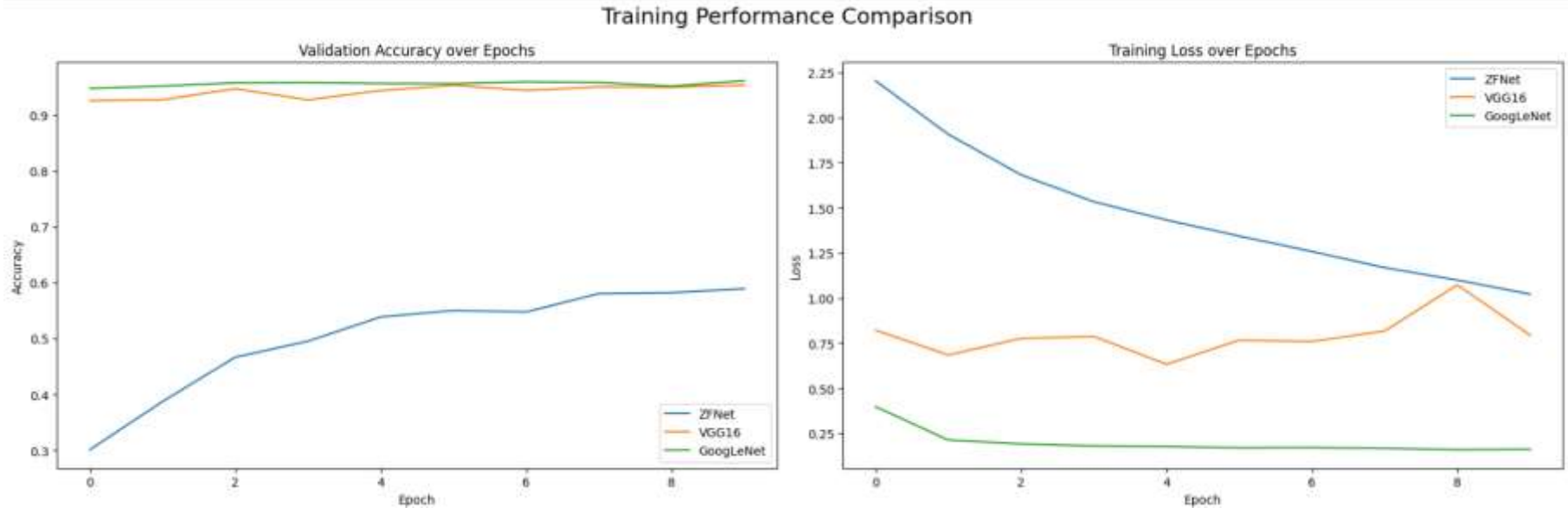
GoogLeNet: Most accurate with minor label typos in results



VGG16: Strong performance but some errors (e.g., butterfly→cat)

Work Accomplished

- Model Comparison (Training Performance)

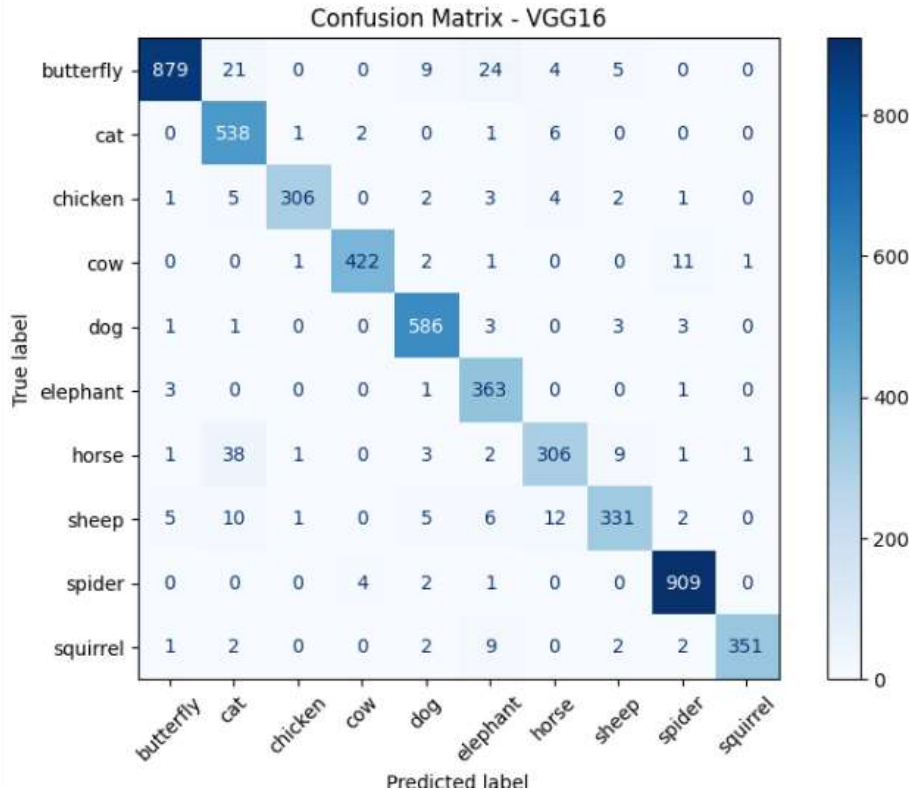


Validation Accuracy over Epochs: GoogLeNet and VGG16 consistently achieve high validation accuracy above 0.9, while ZFNet shows slower improvement and remains significantly lower.

Training Loss over Epochs: GoogLeNet maintains the lowest training loss throughout, followed by VGG16, while ZFNet shows a steady but higher loss trend.

Work Accomplished

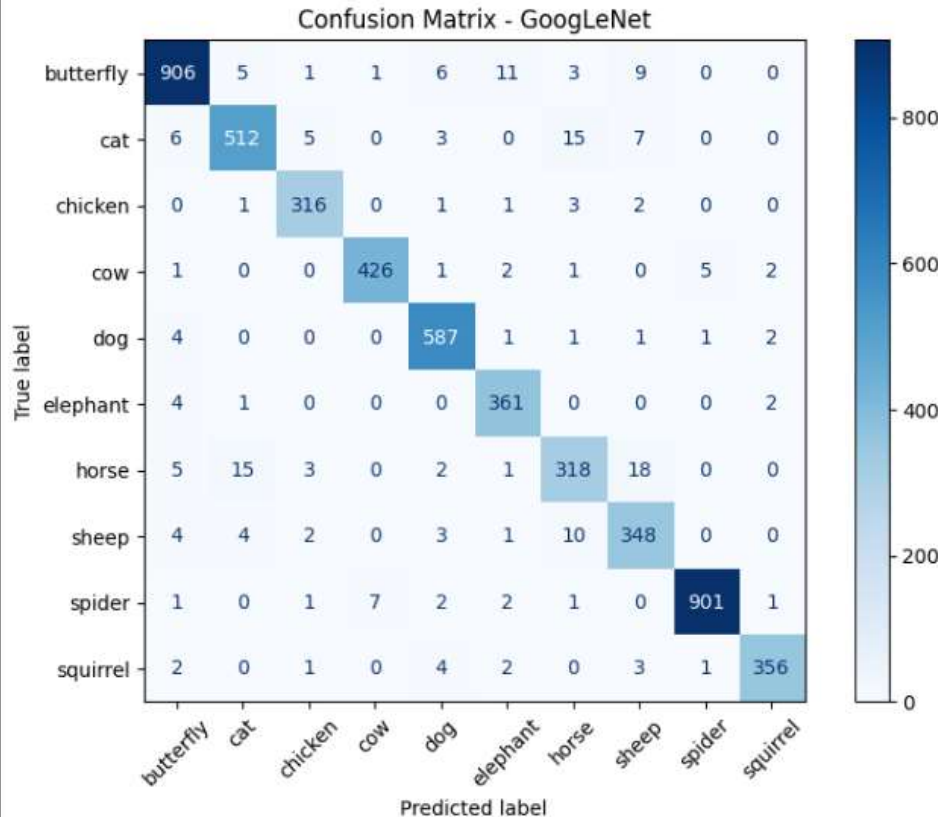
Model Comparison (Confusion Matrix)



VGG16: Achieves high accuracy with very clear diagonal dominance, indicating strong classification and minimal confusion across all classes

Work Accomplished

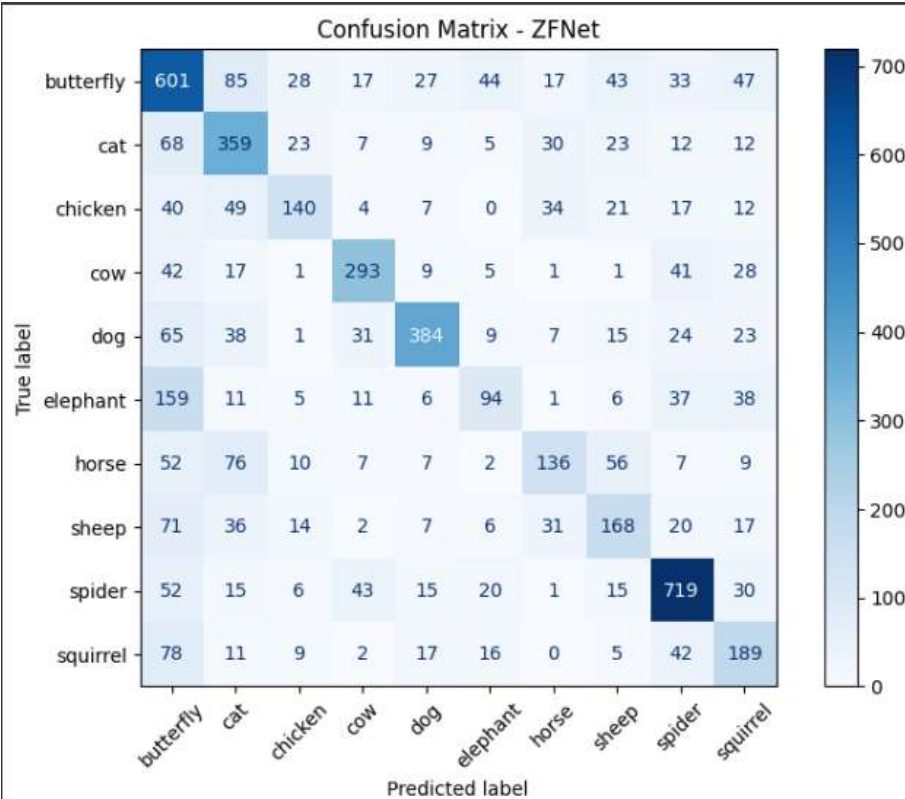
■ Model Comparison (Confusion Matrix)



GoogLeNet: Delivers the best performance with the highest precision and lowest misclassification, showing nearly perfect separation between all classes.

Work Accomplished

- Model Comparison (Confusion Matrix)



ZFNet: Shows moderate classification performance with significant misclassifications, especially among similar animal classes like cat-dog and cow-horse.

Work Accomplished

- Model Comparison (Combined Precision, Recall, F1-Score, and Accuracy Table)

```

Evaluating ZFNet...
Evaluating VGG16...
Evaluating GoogLeNet...

```

	Class	ZFNet_precision	ZFNet_recall	ZFNet_f1-score	
0	butterfly	0.489000	0.638000	0.554000	
1	cat	0.515000	0.655000	0.577000	
2	chicken	0.591000	0.432000	0.499000	
3	cow	0.703000	0.669000	0.685000	
4	dog	0.787000	0.643000	0.708000	
5	elephant	0.468000	0.255000	0.330000	
6	horse	0.527000	0.376000	0.439000	
7	sheep	0.476000	0.452000	0.463000	
8	spider	0.755000	0.785000	0.770000	
9	squirrel	0.467000	0.512000	0.488000	
10	Overall Accuracy	0.588808	0.588808	0.588808	

	VGG16_precision	VGG16_recall	VGG16_f1-score	GoogLeNet_precision
0	0.987000	0.933000	0.959000	0.971000
1	0.875000	0.982000	0.925000	0.952000
2	0.987000	0.944000	0.965000	0.960000
3	0.986000	0.963000	0.975000	0.982000
4	0.958000	0.982000	0.969000	0.964000
5	0.879000	0.986000	0.930000	0.945000
6	0.922000	0.845000	0.882000	0.903000
7	0.940000	0.890000	0.914000	0.897000
8	0.977000	0.992000	0.985000	0.992000
9	0.994000	0.951000	0.972000	0.981000
10	0.953209	0.953209	0.953209	0.960848

	GoogLeNet_recall	GoogLeNet_f1-score
0	0.962000	0.966000
1	0.934000	0.943000
2	0.975000	0.968000
3	0.973000	0.977000
4	0.983000	0.973000
5	0.981000	0.963000
6	0.878000	0.891000
7	0.935000	0.916000
8	0.984000	0.988000
9	0.965000	0.973000
10	0.960848	0.960848

The image displays precision, recall, and F1-scores per class for ZFNet, VGG16, and GoogLeNet, showing that GoogLeNet achieves the highest overall accuracy and consistent performance across all classes.

Work Accomplished

■ Result Analysis

Model	Validation Accuracy	Training Time	Training Loss	Validation Loss	Key Observations
ZFNet	58.88%	29m 47s	0.98 → 0.65	1.30 → 1.12	Underfitting, slow convergence
VGG16	95.32%	39m 25s	0.15 → 0.02	0.18 → 0.05	High accuracy but computationally heavy
GoogLeNet	96.08%	18m 56s	0.12 → 0.01	0.14 → 0.03	Best accuracy-speed balance

Work Accomplished

- Summary

This project implemented a deep learning system for animal species classification using Kaggle's Animals-10 dataset (28,000 images across 10 classes). Three CNN architectures were evaluated: a custom-built ZFNet (59% val accuracy), fine-tuned VGG16 (95% val accuracy), and modified GoogLeNet (96% val accuracy). The models were trained for 10 epochs using CrossEntropyLoss and Adam optimizer, with GoogLeNet showing the best balance of speed (18m 56s) and performance. Results were visualized through training curves, sample predictions (GoogLeNet/VGG16: 80% accuracy on test samples vs ZFNet's 60%), and confusion matrices, revealing GoogLeNet's superior classification with minimal misclassifications compared to the other models. The system demonstrates effective automated species identification with potential for conservation applications

Bibliography

1. **M. D. Zeiler and R. Fergus**, "Visualizing and Understanding Convolutional Networks," *arXiv preprint arXiv:1311.2901*, 2013. (*ZFNet*)
2. **K. Simonyan and A. Zisserman**, "Very Deep Convolutional Networks for Large-Scale Image Recognition," *arXiv preprint arXiv:1409.1556*, 2014. (*VGG16*)
3. **C. Szegedy et al.**, "Going Deeper with Convolutions," *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2015, pp. 1–9. (*GoogLeNet*)
4. **A. Krizhevsky et al.**, "ImageNet Classification with Deep Convolutional Neural Networks," *Adv. Neural Inf. Process. Syst. (NeurIPS)*, vol. 25, 2012. (*AlexNet*)
5. **Kaggle**, "Animals-10 Dataset," *Online*. Available: <https://www.kaggle.com/datasets/alessiocrrado99/animals10>
6. **PyTorch Documentation**, "Torchvision Models," *Online*. Available: <https://pytorch.org/vision/stable/models.html>
7. **Scikit-learn**, "Confusion Matrix and Classification Metrics," *Online*. Available: https://scikit-learn.org/stable/modules/model_evaluation.html

*Thank
you*

