MSDSM Batch - 2

Submitted by,
A Visshal - 2204107001
Kondagunta Srividya - 2204107028

Computer Vision - Group Assignment

Animal classification using facial images with score-level fusion

Problem Statement

Animal classification is a critical task in various domains, including wildlife conservation, research, and veterinary science. Traditional methods for animal identification often rely on physical tags or tracking devices. However, these methods can be invasive, labor-intensive, and may not work well with certain species. This project aims to develop a non-invasive and efficient system for animal classification using facial images, which can be particularly useful for wildlife monitoring and research.

Dataset or Database:

The project will utilize the LHI-Animal-Faces dataset comprising 44 distinct classes of animal head , plants and human head images, resulting in a total of 2,200 images. In Figure 1, we provide an illustrative display of five sample images from each category.

Our exclusive focus is on animal faces where we have considered 4 classes - Cat, Monkey, Panda and Zebra with a total of ~100 images. This focus accentuates the dataset's unique characteristics, as the images exhibit significant intra-class similarities attributed to evolutionary relationships, leading to some animal face categories resembling one another. Simultaneously, the dataset presents considerable inter-class variations, encompassing factors such as rotation, posture differences, and subtypes within the classes.

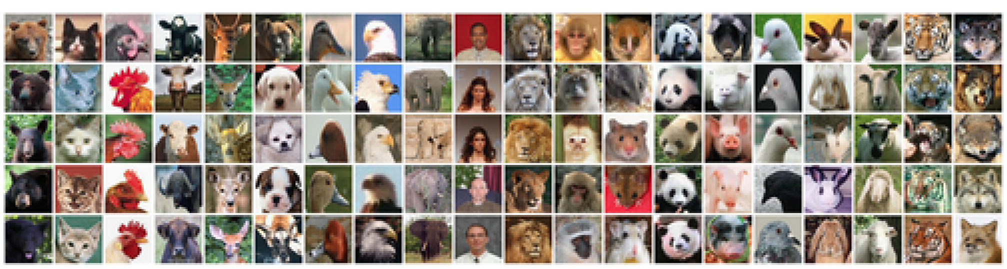
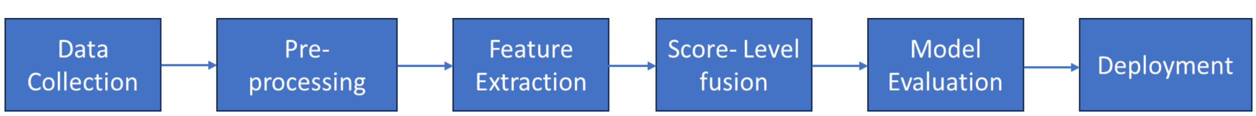


Figure 1: LHI-Animal-Faces dataset. Five images are shown for each category
Source: https://vcla.stat.ucla.edu/people/zhangzhang-si/HiT/exp5.html

Methodology

The methodology for animal classification using facial images with score-level fusion will involve the following key steps:



Submitted by,
A Visshal - 2204107001
Kondagunta Srividya - 2204107028

- Data Collection: The LHI-Animal-Faces dataset, consisting of 44 distinct classes of animal head images, is used. The project emphasizes four classes: Cat, Monkey, Panda, and Zebra. Download the LHI-Animal-Faces dataset from https://vcla.stat.ucla.edu/people/zhangzhang-si/HiT/exp5.html .

- Preprocessing: The dataset undergoes preprocessing, including resizing, normalization, and augmentation, ensuring consistency and enhancing model performance.

- Feature Extraction:Deep learning techniques, specifically Convolutional Neural Networks (CNNs), are employed to extract relevant features from facial images. A CNN model is trained on the dataset.

- Score-Level Fusion: Classification scores from multiple models or sources are combined using score-level fusion techniques. This step aims to improve classification accuracy, employing methods such as weighted averaging or voting.

- Model Evaluation: The performance of the animal classification model is assessed using metrics like accuracy, precision, recall, and F1 score. Fine-tuning is performed to optimize the model.

- Deployment: The final model is implemented in real-world settings, such as wildlife monitoring systems, enabling accurate animal classification based on facial images.

Database:
The LHI-Animal-Faces dataset is utilized, showcasing significant intra-class similarities and considerable inter-class variations. This dataset comprises 2,200 images across 44 classes.

Expected Outcomes:
The expected outcomes of the project include:
- A robust animal classification system capable of accurately identifying animals from facial images.
- Score-level fusion should cause an improvement in accurate classification compared to using CNNs by themselves.
- Improved non-invasive wildlife monitoring and research capabilities.
- Potential applications in species conservation, wildlife tracking, and ecological research.
- Research papers and documentation outlining the methodology and results.

The project addresses a practical need for non-invasive animal classification using modern computer vision techniques, benefiting the scientific community and wildlife conservation efforts.

Results and Discussion:

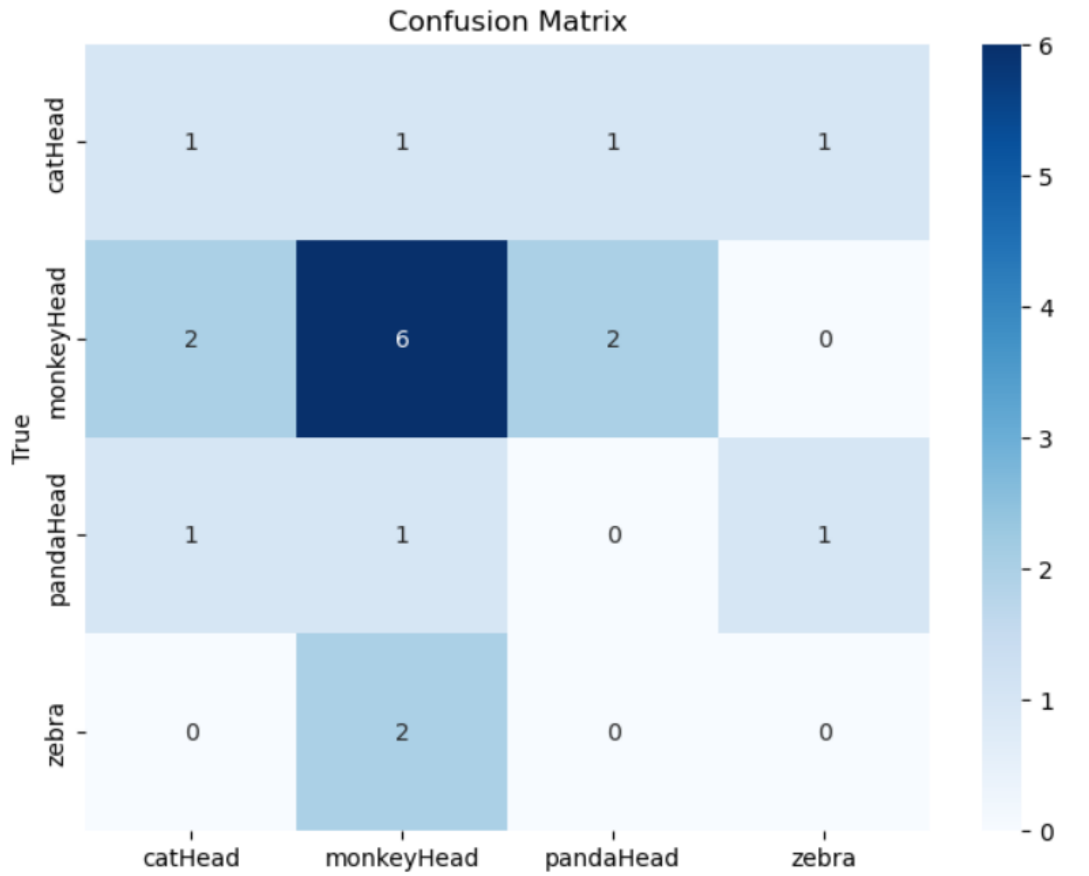
Results from the model evaluation demonstrate the accuracy and effectiveness of the proposed animal classification system. Score-level fusion contributes to improved classification compared to using CNNs alone. The project's practical applications include non-invasive wildlife monitoring, species conservation, and ecological research.

Outcome from the project code:

<u>CNN Model:</u>

```
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 conv2d (Conv2D)             (None, 222, 222, 32)      896

 max_pooling2d (MaxPooling2  (None, 111, 111, 32)      0
 D)

 conv2d_1 (Conv2D)           (None, 109, 109, 64)      18496

 max_pooling2d_1 (MaxPoolin  (None, 54, 54, 64)        0
 g2D)

 conv2d_2 (Conv2D)           (None, 52, 52, 128)       73856

 max_pooling2d_2 (MaxPoolin  (None, 26, 26, 128)       0
 g2D)

 flatten_2 (Flatten)         (None, 86528)             0

 dense_4 (Dense)             (None, 128)               11075712

 dense_5 (Dense)             (None, 4)                 516

=================================================================
Total params: 11169476 (42.61 MB)
Trainable params: 11169476 (42.61 MB)
Non-trainable params: 0 (0.00 Byte)
_________________________________________________________________
```

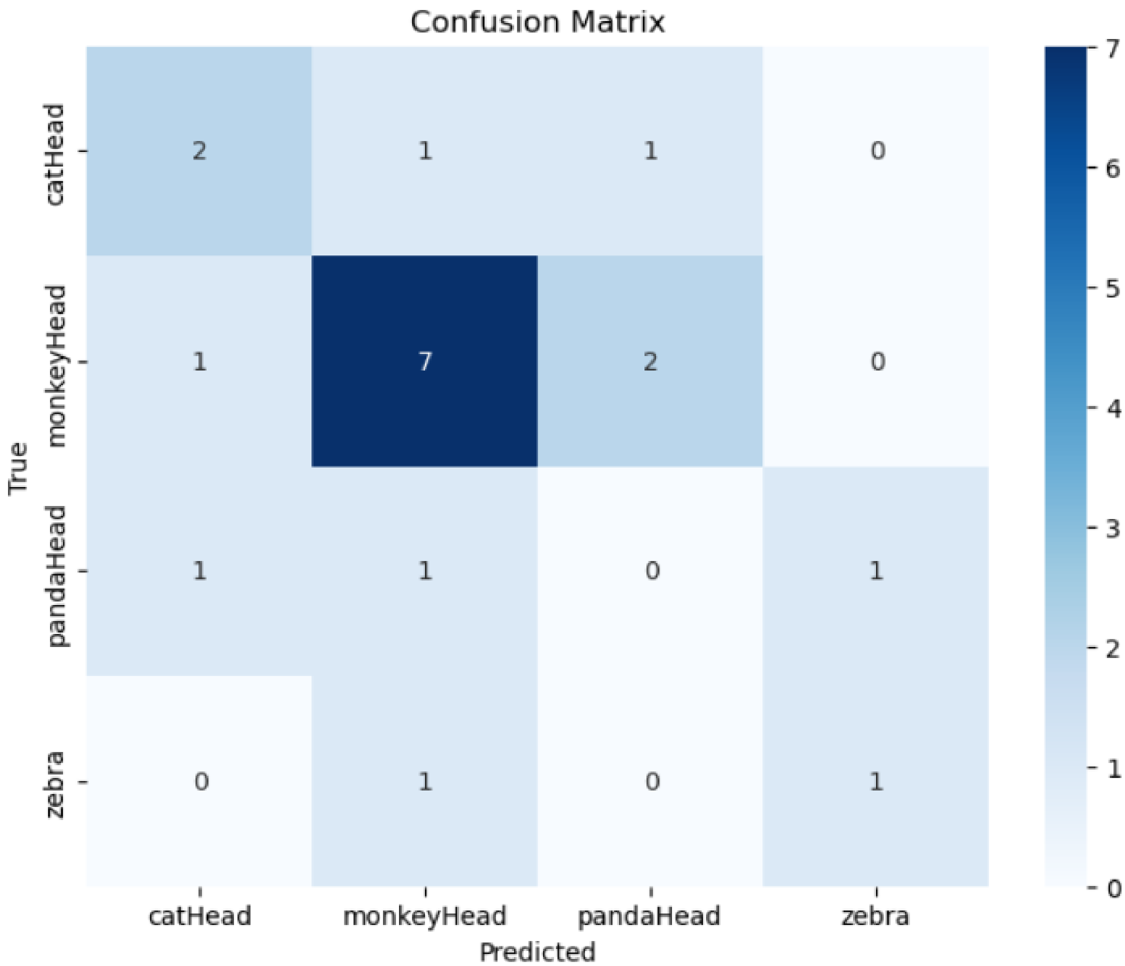
Accuracy: 0.3684210526315789



<u>VGG16 model:</u>

```
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 vgg16 (Functional)          (None, 7, 7, 512)         14714688

 flatten_1 (Flatten)         (None, 25088)             0

 dense_2 (Dense)             (None, 128)               3211392

 dense_3 (Dense)             (None, 4)                 516

=================================================================
Total params: 17926596 (68.38 MB)
Trainable params: 3211908 (12.25 MB)
Non-trainable params: 14714688 (56.13 MB)
_________________________________________________________________
```

```
Accuracy: 0.5263157894736842
```

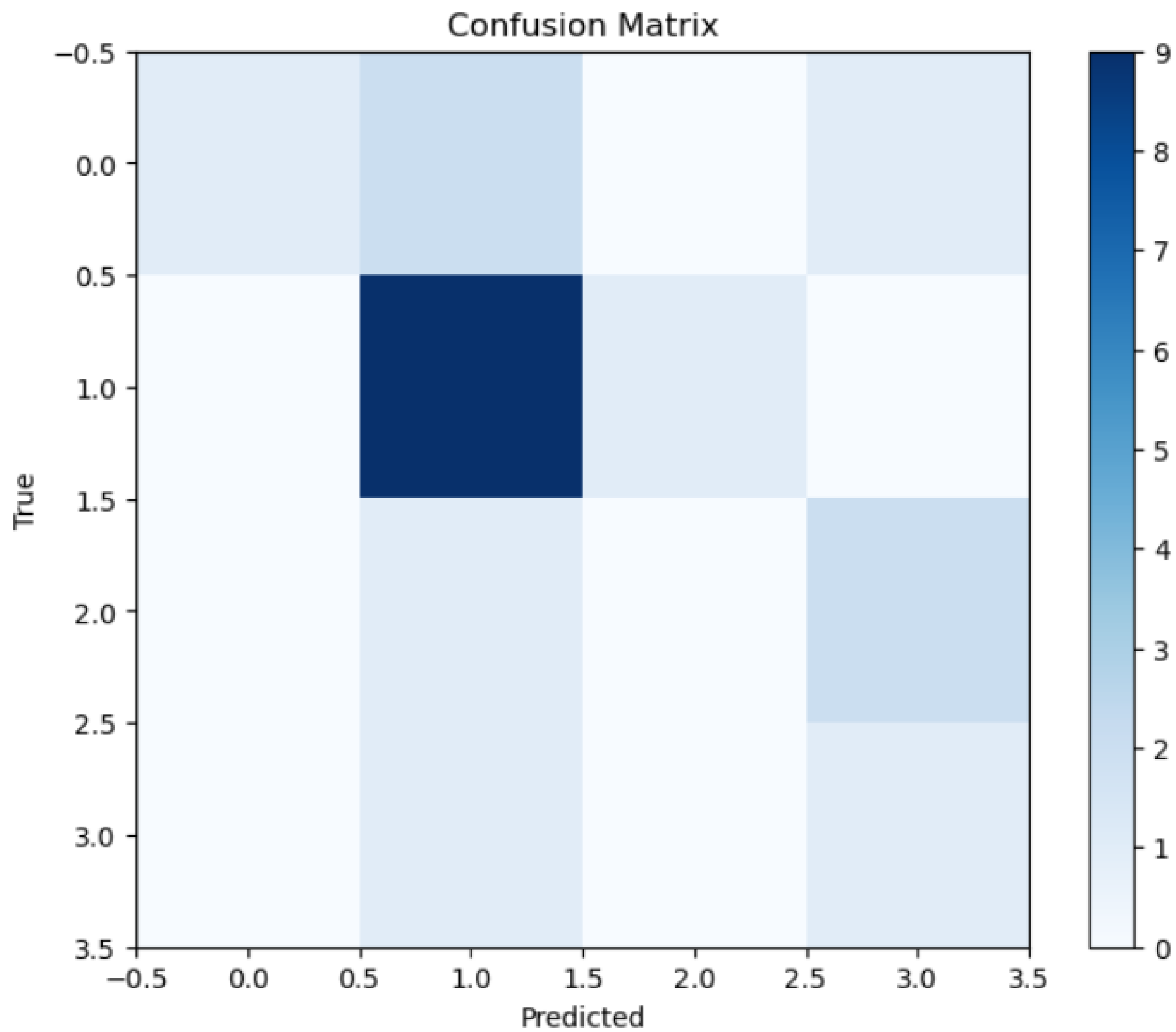


<u>Score level fusion</u>:

```python
# Combine the predictions using score-level fusion (weighted averaging)
combined_predictions = 0.5 * (predictions_CNN + predictions_VGG16)

# Convert the combined predictions to class labels
final_predictions = np.argmax(combined_predictions, axis=1)

# Step 6: Evaluate the Model
accuracy = accuracy_score(test_generator.classes, final_predictions)
conf_matrix = confusion_matrix(test_generator.classes, final_predictions)

print("Accuracy:", accuracy)

# Visualize confusion matrix as a heatmap
plt.figure(figsize=(8, 6))
plt.imshow(conf_matrix, cmap='Blues', interpolation='nearest')
plt.title('Confusion Matrix')
plt.colorbar()
plt.xlabel('Predicted')
plt.ylabel('True')
plt.show()
```

Submitted by,
A Visshal - 2204107001
Kondagunta Srividya - 2204107028

```
Accuracy: 0.5789473684210527
```



Conclusion:
The project addresses the need for non-invasive animal classification using modern computer vision techniques. By focusing on facial images and employing score-level fusion, a higher accuracy is obtained than individual models and the system offers a valuable contribution to wildlife conservation efforts and scientific research.

Reference:
S. Taheri and Ö. Toygar, "Animal classification using facial images with score-level fusion," in IET Computer Vision, vol. 12, no. 5, pp. 679-685, 8 2018, doi: 10.1049/iet-cvi.2017.0079.

Dataset: LHI Animal Face dataset: https://vcla.stat.ucla.edu/people/zhangzhang-si/HiT/exp1.html