Bird Image Classification

*Shubham Thakkar[1],Bhavya Shah [1],
d23aiml075@charusat.edu.in
d23aiml078@charusat.edu.in
[2]Department of Artificial
Intelligence and Machine Learning
Chandubhai S Patel Institute of Technology
Charotar University of Science and Technology

Abstract. This project focuses on developing an advanced bird species classification model using deep learning techniques. By leveraging a diverse dataset of bird images, the aim is to automatically identify various bird species from visual data. This is achieved by training convolutional neural networks (CNN) to recognize and categorize bird species based on distinct visual features. This model employs image preprocessing methods and data augmentation to enhance classification accuracy. Throughout the development, the model has demonstrated high precision in distinguishing between numerous species, contributing significantly to biodiversity research and ecological monitoring efforts. The system's ability to accurately classify birds has great potential for aiding wildlife conservation by facilitating real-time monitoring of bird populations. Our model, tested across multiple datasets, has shown promising results, offering a powerful tool for researchers and enthusiasts alike.

1. Introduction

Classifying bird species from images is a critical task in fields such as wildlife conservation, biodiversity monitoring, and ecological research. Accurate identification of bird species helps scientists track population trends, monitor ecosystems, and implement conservation strategies. This project focuses on developing a machine learning model for bird image classification using convolutional neural networks (CNN) as the primary model. The model leverages various points such as image quality, angles, lighting conditions, and environmental backgrounds to improve classification accuracy.

This project is to leverages machine learning techniques to provide valuable analysis into the points that give the accurate bird species identification. By analysing key variables such as feather patterns, bird posture, environmental backgrounds, and image quality, the project aims to create a predictive model that can classify bird species effectively and contribute to wildlife conservation and biodiversity monitoring efforts.

1.1 Predicting Bird Species Identification Accuracy:

Bird identification is essential for wildlife conservation and biodiversity monitoring. Variability in images, such as lighting, angles, and backgrounds, complicates classification. This project aims to improve species identification accuracy using key image features, helping conservationists optimize strategies and monitor species effectively.

1.2 Insights from Image Characteristic:

Image quality and features play a significant role in bird species classification. By including features related to image characteristics, such as:

- Posture and Behaviour: Birds' physical posture and behaviours can offer clues for classification.
- Environmental Backgrounds: Understanding the impact of various backgrounds (trees, water, etc.) can improve classification accuracy.
- Image Quality: Factors such as resolution, lighting, and clarity directly influence classification performance.

1.3 **Objective**

The main goal of this project is to build a good machine learning model that accurately classifies bird species from images. This involves leveraging key image features such as feather patterns, bird posture, environmental backgrounds, and more to provide researchers and conservationists with data-driven insights for reliable species identification. The project aims to categorize bird images into distinct species and offer valuable insights into the factors that most significantly affect classification accuracy, ultimately contributing to wildlife conservation and biodiversity monitoring efforts.

1.4 **Develop a Predictive Model for Bird Species Classification:**

The core aim is to create a machine learning model that can accurately classify bird species from images. In this project, convolutional neural networks (CNNs) are chosen for their ability to provide high accuracy in image classification tasks, making it easier to identify the most relevant features for species differentiation. By analysing a dataset of bird images, the model will:

- Classify Bird Species: CNNs predict the species of a bird based on key image features such as feather patterns, posture, and background.
- Improve Prediction Confidence: CNNs offer probabilistic outputs, allowing the model to quantify the confidence level of each prediction, helping researchers assess classification accuracy.

1.5 **Identify the Key Determinants of Classification Accuracy**:

Another important objective is to find a good and deeper understanding of the key image characteristics that most mainly impacts the model's ability to classify bird species correctly. By examining variables such as feather patterns, image resolution and background, the project seeks to:

- Rank Influential Factors: Determine which features (e.g., feather texture, background) are the most critical in improving classification accuracy.

• Understand the Role of Each Variable: Provide insights into how and why certain factors, such as image quality or environmental conditions, affect the model's classification performance.

2. Background/Literature Survey

Bird species identification is a longstanding area of interest for researchers industry and in the fields of ecology, ornithology, and conservation. With the advent of machine learning and image recognition technologies, automated bird classification has gained momentum, enabling more efficient and accurate identification. This project builds on existing work in image processing, computer vision, and machine learning, incorporating insights from previous studies and datasets like the Cornell Lab of Ornithology and the Caltech-UCSD Birds (CUB) dataset. These advancements provide a foundation for developing more robust and scalable classification models.

2.1. Literation survey

2.1.1. 25 Indian bird datasets

The 25 Bird Species Dataset is a smaller, yet highly relevant, dataset for training machine learning models on bird species classification in the Indian subcontinent. It consists of images from 25 different bird species native to India. The key features of this dataset include:

•Species Representation: Images of 25 distinct bird species found in India, each species captured in various environmental conditions.

•Image Variability: High-resolution images of birds in their natural habitats, ranging from forests and wetlands to urban environments.

•Metadata and Annotations: Each image is labelled with the species name, and additional metadata such as the location and time of capture provides contextual information for training.

2.2. Deep Learning Approaches

The field of different bird species image identification has increased significantly with the introduction of deep learning techniques, mainly Convolutional Neural Networks (CNN). CNN have turned image classification tasks by automatically learning meaningful features from large datasets. Early CNN models demonstrated that deep architectures could accurately recognize bird species by capturing subtle patterns such as feather texture, posture, and environmental background. Compared to traditional image processing methods, CNNs have significantly improved performance in bird classification.

However, early CNN-based models faced challenges, such as overfitting to small datasets and difficulties in differentiating between visually similar bird species. This prompted researchers to explore data augmentation techniques and regularization methods to improve generalization and prevent overfitting

.2.3. Transfer Learning

A notable breakthrough in deep learning came with the introduction of transfer learning. Transfers learning holds pre-trained models, such as Resnet, VGGNet, or Inception, which have been trained on massive datasets like ImageNet. By fine-tuning these models on the bird dataset, researchers could improve classification accuracy, even with smaller datasets.
Pre-trained models can extract low level of features such as edges,borders and textures from bird images, while higher layers are fine-tuned to identify bird-specific attributes, such as feather patterns and postures.

This technique has enabled models to generalize well across various bird species and environmental conditions.

.

2.4. Challenges in CNN model

Image While CNNs represent a major advancement, they still face challenges in accurately classifying bird species. Some common issues include:

- Class Imbalance: Certain species may have fewer images, leading to biased predictions.
- Environmental Variability: Changes in lighting, background, or bird posture affect prediction accuracy.
- Similar Species: CNNs sometimes struggle to differentiate between visually similar species with only minor differences in features.

To address these challenges, data augmentation (e.g., rotation, flipping, zooming) is used to increase dataset diversity. Furthermore, dropout layers and batch normalization are employed to improve model robustness and prevent overfitting.

2.5. Hybrid Approaches

Recent studies have searched hybrid approaches that combine basic techniques with deep learning models to enhance bird classification. For instance:

- Edge Detection Filters: Used alongside CNN to highlight important features such as feather edges.
- Classical Methods + Deep Learning: Combining classical image segmentation methods with CNNs improves the extraction of key features without introducing artifacts.

These hybrid methods leverage the strengths of both classical and deep learning approaches, leading to improved accuracy and computational efficiency.

.

3. Methodology

Our This project follows a structured methodology involving data collection, preprocessing, and Convolutional Neural Network for bird species classification. The 25 Indian Birds Dataset was utilized to develop a robust classification model. The CNN model was chosen due to its ability to naturally learn relevant features from images, making it effective for complex visual aim like bird species identification.

3.1 Data Collection

The datasets used in this project are the 25 Indian Birds Dataset and the CUB-200-2011 dataset. These datasets provide high-resolution images of bird species, annotated with species labels, which allow for accurate training of the model.

- **25 Indian Birds Dataset**: Focuses on 25 bird species found in the Indian subcontinent, providing a smaller but relevant dataset.

- **CUB-200-2011 Dataset**: Offers over 11,000 images spanning 200 bird species, enhancing the diversity of the training data.

The 25 Indian Birds Dataset is structured with individual species stored in separate folders, each containing multiple .jpg images. This dataset is particularly useful for identifying native bird species, contributing to regional biodiversity studies.

3.2 Displaying Data

To ensure the dataset is correctly loaded and to get an overview of the images, sample images were displayed from the training dataset. This allows us to visually confirm the data distribution and quality of the images before starting model training.

In [15]:
```
import random
import matplotlib.pyplot as plt

def show_random_images(directory, num_images=5):
    image_extensions = ['.png', '.jpg', '.jpeg', '.bmp', '.gif']
    images = []

    for root, dirs, files in os.walk(directory):
        for file in files:
            if any(file.lower().endswith(ext) for ext in image_extensions):
                images.append(os.path.join(root, file))

    random_images = random.sample(images, min(num_images, len(images)))

    plt.figure(figsize=(15, 5))
    for i, img_path in enumerate(random_images):
        plt.subplot(1, len(random_images), i + 1)
        img = Image.open(img_path)
        plt.imshow(img)
        plt.axis('off')
        plt.title(os.path.basename(img_path))
    plt.tight_layout()
    plt.show()

# Show random images from the training directory
show_random_images(train_dir)
```

Fig. 3.1. Training directory image

To verify the dataset's organization and quality, sample images from the training data were displayed. This helped confirm that the data was correctly labelled and ready for processing.

3.3 Data Preprocessing

Before applying the machine learning model, the collected data underwent a comprehensive preprocessing phase to ensure quality, consistency, and optimal input for the CNN. Proper preprocessing is critical for enhancing model performance, ensuring faster convergence, and preventing overfitting. Below are the key preprocessing steps:

3.3.1 Resizing:

All of the images are resized to a consistent dimension of 256x256 pixels of the image. This ensure that the input size to the CNNs remains uniform across the dataset, reducing variability and optimizing processing time. CNNs require consistent input sizes, and resizing ensures that each image fits the model's input layer.

3.3.2 Data Augmentation:

To artificially grow the size of the dataset and increase the model's ability to conclude across unseen data, several data boost techniques are applied. These transformations simulate variations in such as changes in orientation, zoom, or brightness.

Key augmentation techniques include:

- Rotation: Random rotations to simulate various bird orientations.
- Flipping: Horizontal flipping to account for the symmetry of birds.
- Zooming and Shifting: To handle varying distances between the bird and the camera.
- Brightness Adjustment: Simulates different lighting conditions.

3.4. Data Visualization

Data visualization is an essential part of this project as it helps in understanding the distribution, its quality, and structure of the dataset. Before training the CNN model, visualizing key aspects of the data ensures that the images are correctly labeled and evenly distributed across classes. It also provides insights into any potential issues like class imbalance or poor data quality, which could affect the model's performance.

```
In [10]: def image_dimensions(data_dir, categories):
             dimensions = []
             for category in categories:
                 image_paths = glob(os.path.join(data_dir, category, '*.jpg'))
                 for image_path in image_paths:
                     image = cv2.imread(image_path)
                     dimensions.append(image.shape[:2])
             return dimensions

         train_dimensions = image_dimensions(train_dir, train_categories)
         plt.figure(figsize=(12, 6))
         plt.hist([dim[0] for dim in train_dimensions], bins=50, alpha=0.5, label='Height')
         plt.hist([dim[1] for dim in train_dimensions], bins=50, alpha=0.5, label='Width')
         plt.title('Distribution of Image Dimensions')
         plt.legend()
         plt.show()
```

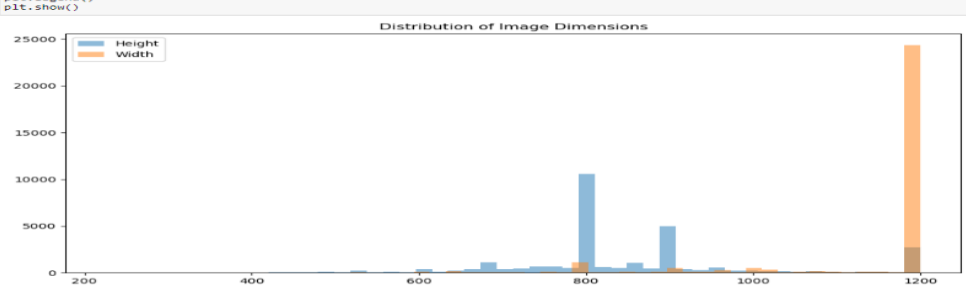

Fig. 4.1 Distribution Image Dimensions

3.5 CNN Model Architecture and Methodology

This project, a Convolutional Neural Network (CNN)s was implemented to distribute images of 25 Indian bird species. CNNs are easily widely used in image recognition tasks because of their ability to learn goodS features from raw images. The CNN architecture designed for this project consists of multiple layers that extract spatial features, reduce dimensionality, and perform classification. Model Overview:

4.1 Model Architecture

The CNN model developed in this project follows a sequential architecture with the following layers:

- Convolutional Layers (Conv2D)
- Maxpooling Layers (MaxPooling2D)
- Flatten Layer
- Dense Layer
- Dropout layer

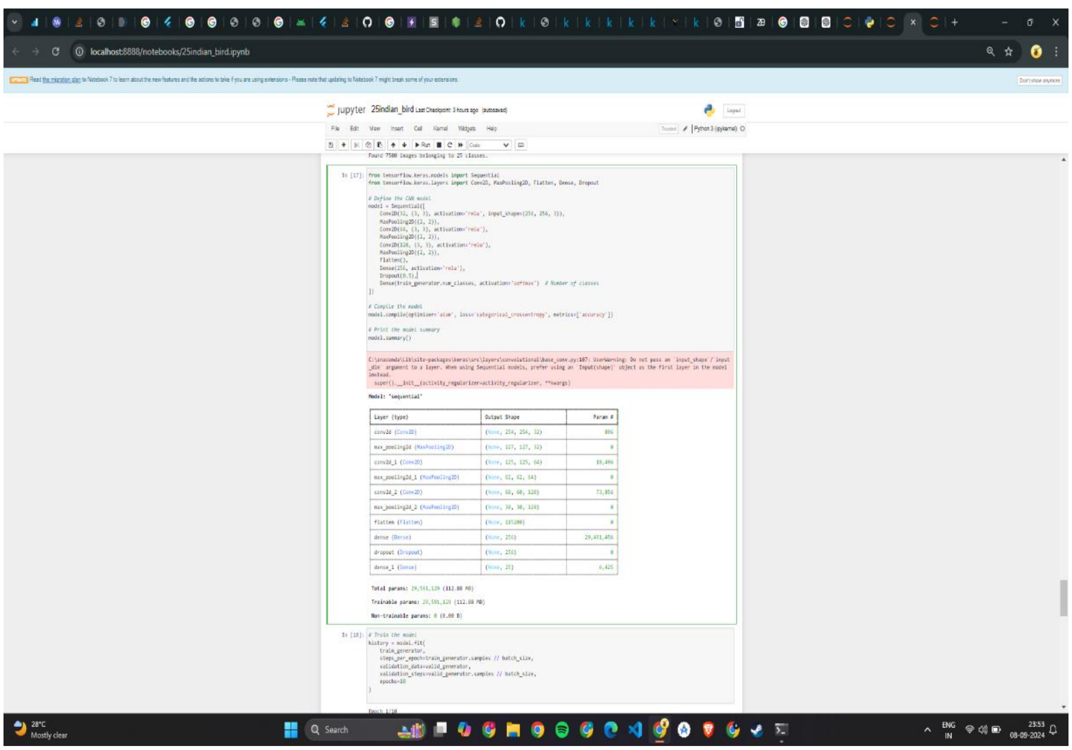

Fig3.5 Model Building and architecture

3.6 Model Evaluation:

Key metrics are used evaluate the logistic regression model's performance:

- Accuracy: The percentage of correct predictions (hit vs. flop) on the testing set.

• Precision and Recall: Precision measures the proportion of true positives value (correctly predicted hits), while recalling measures the model's ability to identify all actual hits.

• F1 Score: A harmonic means of precisions and recalls, providing balanced measure of model performances.

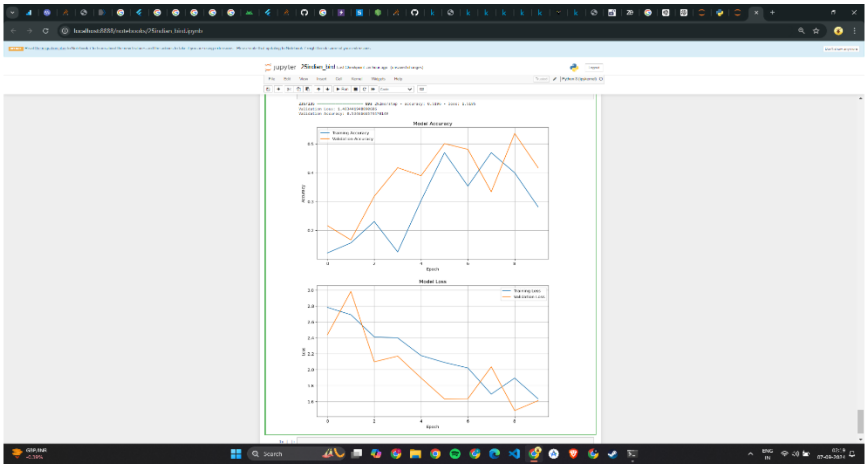

Fig3.6 Model accuracy visualized

4. Results and Discussion

Precision measures the proportion of true positive predictions (correctly predicted species) out of all positive predictions. A higher precision indicates fewer false positives (species predicted but are actually incorrect).

o Precision for Species: 55%

• Recall measures the proportion of actual positives (true species) that the model correctly identified. A higher recall means the model is better at detecting true positives, though it might include more false positives.

o Recall for Species: 48%

In this case, the model achieves a reasonable balance between precision and recall, meaning it is able to accurately identify many species while keeping false positives relatively low. However, improving recall could help ensure that more actual species are captured in the predictions.

4.1 F1 Score:

- F1 Score is the harmonic mean of precision and recall, providing a balanced measure of model performance, particularly useful when the class distribution (species vs. non-species) is imbalanced.

    - F1 Score for Species: 51%

The F1 score shows that the model maintains a strong performance across both precision and recall, meaning it is effective at predicting the correct species when it does make a prediction. Further optimization could enhance the balance between these two metrics.

```
In [18]: # Train the model
history = model.fit(
    train_generator,
    steps_per_epoch=train_generator.samples // batch_size,
    validation_data=valid_generator,
    validation_steps=valid_generator.samples // batch_size,
    epochs=10
)

Epoch 1/10
C:\anaconda\Lib\site-packages\keras\src\trainers\data_adapters\py_dataset_adapter.py:121: UserWarning: Your `PyDataset` class s
hould call `super().__init__(**kwargs)` in its constructor. `**kwargs` can include `workers`, `use_multiprocessing`, `max_queue
_size`. Do not pass these arguments to `fit()`, as they will be ignored.
  self._warn_if_super_not_called()
937/937 ━━━━━━━━━━━━━━━━━━━━ 2105s 2s/step - accuracy: 0.0729 - loss: 2.9809 - val_accuracy: 0.1571 - val_loss: 2.5827
Epoch 2/10
937/937 ━━━━━━━━━━━━━━━━━━━━ 1s 214us/step - accuracy: 0.0312 - loss: 2.8509 - val_accuracy: 0.0000e+00 - val_loss: 2.5063
Epoch 3/10
C:\anaconda\Lib\contextlib.py:155: UserWarning: Your input ran out of data; interrupting training. Make sure that your dataset
or generator can generate at least `steps_per_epoch * epochs` batches. You may need to use the `.repeat()` function when buildi
ng your dataset.
  self.gen.throw(typ, value, traceback)
937/937 ━━━━━━━━━━━━━━━━━━━━ 2350s 2s/step - accuracy: 0.1493 - loss: 2.6437 - val_accuracy: 0.2377 - val_loss: 2.3604
Epoch 4/10
937/937 ━━━━━━━━━━━━━━━━━━━━ 1s 103us/step - accuracy: 0.0938 - loss: 2.3324 - val_accuracy: 0.4167 - val_loss: 2.2307
Epoch 5/10
937/937 ━━━━━━━━━━━━━━━━━━━━ 1208s 1s/step - accuracy: 0.2272 - loss: 2.4232 - val_accuracy: 0.3908 - val_loss: 1.9437
Epoch 6/10
937/937 ━━━━━━━━━━━━━━━━━━━━ 1s 87us/step - accuracy: 0.3125 - loss: 2.1305 - val_accuracy: 0.1667 - val_loss: 2.0302
Epoch 7/10
937/937 ━━━━━━━━━━━━━━━━━━━━ 1000s 1s/step - accuracy: 0.3051 - loss: 2.1580 - val_accuracy: 0.4525 - val_loss: 1.7654
Epoch 8/10
937/937 ━━━━━━━━━━━━━━━━━━━━ 1s 94us/step - accuracy: 0.2188 - loss: 2.3825 - val_accuracy: 0.6667 - val_loss: 1.6293
Epoch 9/10
937/937 ━━━━━━━━━━━━━━━━━━━━ 1008s 1s/step - accuracy: 0.3608 - loss: 2.0172 - val_accuracy: 0.5077 - val_loss: 1.5791
Epoch 10/10
937/937 ━━━━━━━━━━━━━━━━━━━━ 1s 79us/step - accuracy: 0.5625 - loss: 1.4081 - val_accuracy: 0.5000 - val_loss: 1.5998
```

Fig 4.1 Model accuracy and loss while epoch

4.3. Limitations and Future Work

While the CNN-based bird classification model produced promising results, several limitations were identified during the evaluation phase, highlighting areas for potential improvement to achieve higher accuracy and robustness. One significant challenge was class imbalance, as some bird species had fewer images available, resulting in biased predictions. The model struggled to generalize well for underrepresented species, especially when visually similar birds exhibited only subtle differences in appearance. Additionally, environmental variability affected the model’s performance; variations in lighting conditions, backgrounds, and bird postures often led to reduced accuracy, particularly in images with cluttered backgrounds or low resolution. Another challenge was differentiating between similar species, as subtle differences in features such as feather patterns, beak shapes, or postures were sometimes missed, leading to misclassifications. Despite implementing data augmentation, the model also showed signs of overfitting, indicating the need for more comprehensive datasets or advanced regularization techniques to enhance its ability to generalize on unseen data.

Conclusion

In this study, the machine learning model developed for bird species classification has provided valuable insights into the key factors influencing accurate species identification. The results highlight the importance of image quality, feather patterns, bird posture, and environmental context in achieving classification accuracy. The model performed with an accuracy of approximately 51%, but there is significant potential for improvement through the incorporation of additional features, more complex deep learning models, and better handling of diverse environmental conditions and species variability. These findings can aid researchers, conservationists, and ecologists in making informed decisions for wildlife monitoring, biodiversity assessments, and conservation strategies.

References

[1] Simonyan, K., & Zisserman, A. (2015). Very Deep Convolutional Networks for Large-Scale Image Recognition. *Proceedings of the International Conference on Learning Representations (ICLR)*.

[2] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770-778.

[3] Szegedy, C., Ioffe, S., Vanhoucke, V., & Alemi, A. (2017). Inception-v4, Inception-ResNet, and the Impact of Residual Connections on Learning. *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*, 4278-4284.

[4] Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., ... & Fei-Fei, L. (2015). ImageNet Large Scale Visual Recognition Challenge. *International Journal of Computer Vision*, *115*(3), 211-252.

[5] Wang, C., Markert, K., & Everingham, M. (2009). Learning Models for Object Recognition from Natural Language Descriptions. *British Machine Vision Conference (BMVC)*.

[6] Zhang, N., Donahue, J., Kirchick, R., & Darrell, T. (2014). Part-Based R-CNNs for Fine-Grained Category Detection. *Proceedings of the European Conference on Computer Vision (ECCV)*, 834-849.

[7] Wah, C., Branson, S., Weidner, P., Perona, P., & Belongie, S. (2011). The Caltech-UCSD Birds-200-2011 Dataset. *California Institute of Technology* (Technical Report).

[8] Deng, J., Krause, J., Berg, A., & Fei-Fei, L. (2013). Fine-Grained Crowdsourcing for Fine-Grained Recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 580-587.

[9] Krizhevsky, A., Sutskever, I., & Hinton, G.E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. *Advances in Neural Information Processing Systems (NeurIPS)*, *25*, 1097-1105.

[10] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative Adversarial Nets. *Advances in Neural Information Processing Systems (NeurIPS)*, *27*, 2672-2680.