Exploring GAN Variants for Balancing Imbalanced Datasets

Joud Dababneh

Dr. Yousef Sanjlawe

Special Topics in AI

Department of Artificial Intelligence

King Abdullah II School of Information Technology

University of Jordan

2025

1 Introduction and Problem Statement

Most classification algorithms are only suitable for balanced dataset that is equivalently distributed across different classes [1]. However, in real-world, datasets are often imbalanced, and sometimes the minority-class is extremely important and need to be accurately classified [2]. For example, medical diagnosis researchers are more care about samples of patients other than healthy people. To mitigate the imbalanced problem, conventional approaches are oversampling, undersampling and cost-sensitive learning.

The main idea of undersampling is to discard most majority-class samples to achieve balance across classes, but this approach might lose lots of information. Cost-sensitive learning [3] method provides different weights through different types of samples, and it pay more attention to samples in minority-class. Nonetheless, using cost-sensitive learning, it is hard to obtain an accurate estimate of misclassification cost. Oversampling is also called data augmentation, which achieves balance mainly by adding minority-class samples. However, traditional data augmentation usually apply some geometric changes on the image classification dataset, e.g. rotation, scaling, translation or mirroring [4], and it might disrupt original relevant features.

The GAN model [5] proposed by Ian J. Goodfellow provides a new paradigm from the perspective of adversarial games, which implicitly learn data distributions by transforming a sample from a simplistic distribution (such as Gaussian) to the data distribution by optimizing a min-max objective function between a pair of function approximators called the generator and the discriminator. See Figure 1.

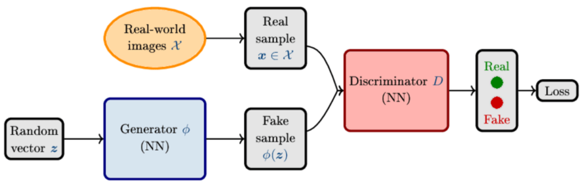

Figure 1: Original GAN Structure

The fundamental difficulty of imbalanced learning problem is that data imbalance significantly compromise the performance of most standard learning algorithms. Most standard algorithms assume a balanced class distribution or equal misclassification costs. When applying complex imbalanced datasets, these algorithms fail to represent the distribution of the data and hence have an unfavorable accuracy.

The GAN shows excellent performances in generating realistic samples. Recent works apply the GAN as an augmentation tool to deal with imbalanced dataset.

2 Description of Dataset & Imbalance Analysis

The dataset used for this report was the Animals-10 Dataset [6] which contains about 28K medium quality animal images belonging to 10 categories: dog, cat, horse, spider, butterfly, chicken, sheep, cow, squirrel, and elephant. The classes chosen in this report were dog, spider, and elephant which the latter being the under-represented class that we will use the GAN variants on. See Figure 2 below for a distribution of the dataset.

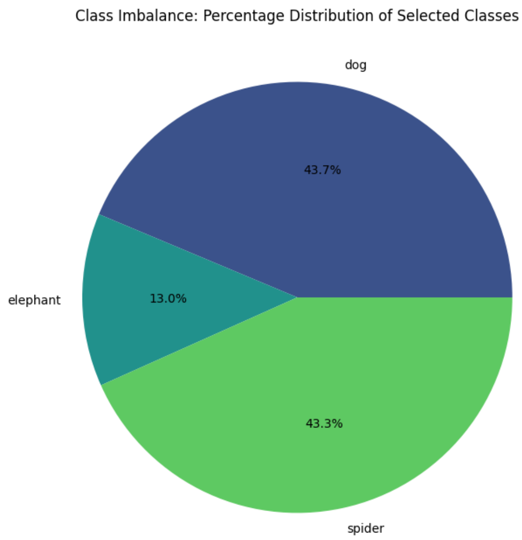

Figure 2: Pie Chart showcasing the data imbalance of the dataset

3 Details of GAN Architectures and Training

3.1 Vanilla GAN

A Vanilla Generative Adversarial Network (GAN) was implemented to synthesize images of elephants from the Animal-10 dataset. The GAN comprises a Generator and a Discriminator, trained adversarially to model the distribution of elephant images.

The Generator is a fully connected neural network that takes as input a 100-dimensional noise vector sampled from a normal distribution. It outputs a 784-dimensional vector corresponding to a flattened grayscale image, with hidden layers employing ReLU activations and the output layer using Tanh to normalize pixel values to the range [-1, 1].

The Discriminator is a binary classifier structured as a fully connected network, using LeakyReLU activations internally and a final Sigmoid activation to estimate the probability that a given image is real. It is trained to distinguish real elephant images from those produced by the Generator.

Training was conducted over 200 epochs using the Adam optimizer (learning rate = 0.0002, beta1 = 0.5) and binary cross-entropy loss. The model was trained exclusively on elephant class images extracted from the Animal-10 dataset. Periodic visualization of generated samples demonstrated progressive refinement of elephant image quality over training.

3.2 Conditional GAN

A Conditional Generative Adversarial Network (CGAN) was implemented to synthesize elephant images for dataset balancing. This CGAN integrates class labels into both its Generator and Discriminator, enabling class-conditional image generation (See Figure 3 for the Structure of CGAN).

The Generator features an improved architecture with effective upsampling and takes a 100-dimensional noise vector along with the class label as input, producing generated images. The Discriminator is designed to differentiate between real and fake images, conditioned on the class label, incorporating spectral normalization and gradient penalty for enhanced training stability. Training focused on stabilizing dynamics and improving image quality through proper conditioning.

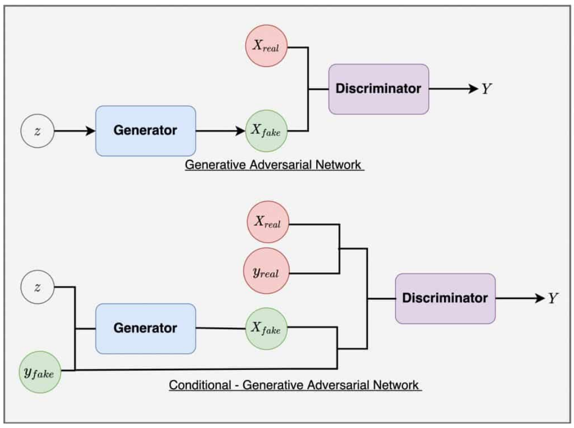

Figure 3: CGAN Structure

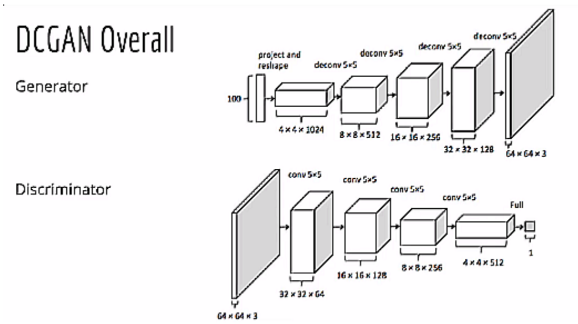

Figure 4: DCGAN Structure

3.3 Deep Convolutional GAN

A Deep Convolutional Generative Adversarial Network (DCGAN) is a popular architecture for generating realistic images using deep learning. It consists of two main components: a generator and a discriminator, both built with convolutional layers. The generator takes random noise as input and uses transposed convolution layers to upsample it into a synthetic image. The discriminator, built with convolutional and LeakyReLU layers, evaluates whether an image is real or generated. Through adversarial training, where the generator learns to fool the discriminator while the discriminator learns to distinguish fake images from real ones, DCGANs effectively produce highly realistic images.

4 Classifier Setup and Evaluation

The same fully connected neural network classifier was trained on the the original imbalanced dataset and the GAN generated balanced datasets to identify animal classes, with particular attention to the elephant class. The architecture consists of three linear layers with ReLU activations and a softmax output over the 3 animal classes.

The classifier was trained using the CrossEntropyLoss function and Adam optimizer (learning rate = 0.001) over 20 epochs with a batch size of 64.

4.1 Original Imbalanced Dataset

The CNN classifier trained on the original imbalanced Animal-10 dataset demonstrated strong performance on the majority classes—dog and spider—achieving high precision and recall due to their larger representation in the training data. In contrast, the elephant class, being significantly underrepresented, showed notably lower recall and precision, indicating the model's limited ability to generalize to this minority class.

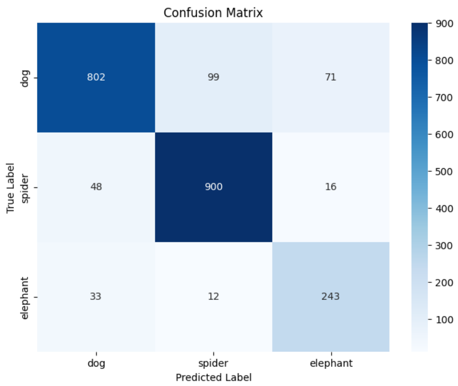

Figure 5: Confusion Matrix of Original Imbalanced Dataset

4.2 Vanilla GAN Balanced Dataset

The vanilla CNN classifier showed clear improvements in overall accuracy and stability after incorporating balanced training data, particularly through the inclusion of GAN-generated elephant images. This led to enhanced recall and F1-score for the elephant class, which had previously underperformed due to class imbalance. While the model maintained high performance on the dog and spider classes, the most significant gains were observed in the model's ability to correctly identify elephant images, reflecting the success of data augmentation in mitigating bias. However, challenges remained in ensuring the quality and diversity of generated samples, as synthetic data occasionally introduced artifacts that limited further performance gains. Despite these limitations, the vanilla classifier demonstrated greater robustness and class-wise fairness when trained on the augmented, balanced dataset.

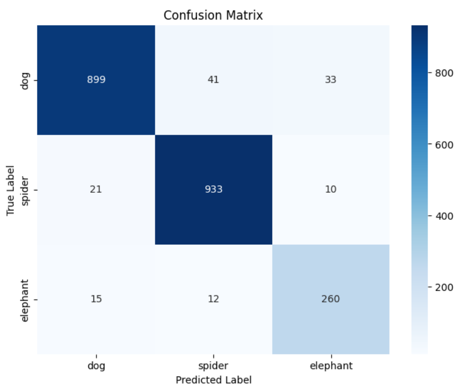

Figure 6: Confusion Matrix of Vanilla GAN balanced dataset

4.3 CGAN Balanced Dataset

Training involved the Adam optimizer (learning rate = 0.001) and `sparse_categorical_crossentropy` loss, conducted over 30 epochs with a batch size of 16. Early stopping and learning rate reduction on plateau were employed.

For Dataset Balancing, a maximum of 2000 CGAN-generated elephant images were added *only* to the training set. The validation and test sets exclusively contained real images to ensure unbiased evaluation.

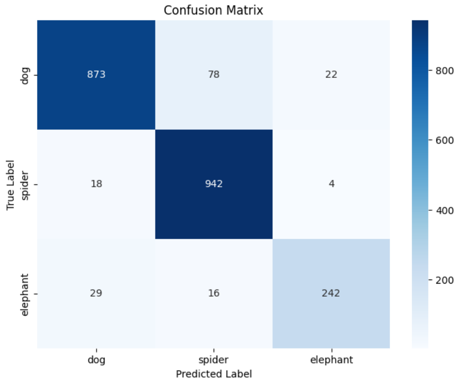

Figure 7: Confusion Matrix of CGAN balanced dataset

4.4 DCGAN Balanced Dataset

The results indicate that the DCGAN was able to generate elephant images with a reasonable level of visual quality, as reflected in the classifier's performance. While the generated images are not indistinguishable from real samples, the classifier was able to recognize them with promising accuracy, suggesting that the GAN captured key features of the elephant class.

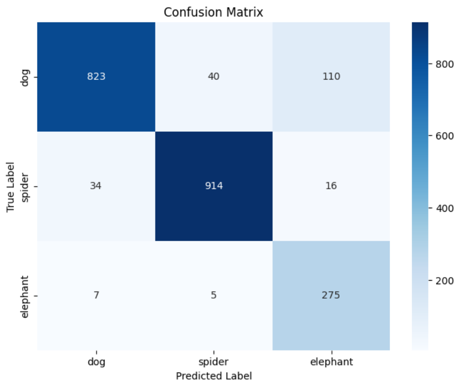

Figure 8: Enter Caption

5 Results & Comparisons

The classification performance on the original imbalanced dataset showed strong results for the majority classes (dog and spider), but significantly lower recall and precision for the minority class (elephant). This highlights the challenge posed by imbalanced datasets to standard classification algorithms.

Upon incorporating GAN-generated elephant images to balance the dataset, the Vanilla GAN, CGAN, and DCGAN variants all demonstrated improvements in the classifier's ability to identify the elephant class.

5.1 6. Observations and Conclusions

This study explored the application of various Generative Adversarial Network (GAN) variants—Vanilla GAN, Conditional GAN (CGAN), and Deep Convolutional GAN (DCGAN)—to address the problem of imbalanced datasets in image classification. Specifically, we focused on balancing the "elephant" class within the Animals-10 dataset.

Our observations confirm that data augmentation using GANs is a viable strategy for mitigating the challenges posed by class imbalance. By generating synthetic samples of the minority class, we were able to improve the performance of a standard CNN classifier on the underrepresented "elephant" class. This was evident in the enhanced recall and precision metrics for the elephant class across all GAN variants compared to the original imbalanced dataset.

While the generated images might not be perfectly indistinguishable from real samples, their utility in improving classifier performance on imbalanced datasets is clear. The models learned to recognize key features of the minority class from these synthetic samples, leading to more robust and fair classification across all categories. Future work could focus on further improving the quality and diversity of generated samples to achieve even greater performance gains and explore more advanced GAN architectures or training techniques to address potential artifacts introduced by synthetic data.

References

[1] G. Haixiang, L. Yijing, J. Shang, G. Mingyun, H. Yuanyue, and G. Bing, "Learning from class-imbalanced data: Review of methods and applications," *Expert systems with applications*, vol. 73, pp. 220–239, 2017.

[2] M. A. Mazurowski, P. A. Habas, J. M. Zurada, J. Y. Lo, J. A. Baker, and G. D. Tourassi, "Training neural network classifiers for medical decision making: The effects of imbalanced datasets on classification performance," *Neural networks*, vol. 21, no. 2-3, pp. 427–436, 2008.

[3] Y.-A. Chung, H.-T. Lin, and S.-W. Yang, "Cost-aware pre-training for multiclass cost-sensitive deep learning," *arXiv preprint arXiv:1511.09337*, 2015.

[4] M. I. T. A. T. Hassner and T. L. J. T. M. Gérard, "Do we really need to collect millions of faces for effective face recognition," *Computer Vision–ECCV*, vol. 2016, 2016.

[5] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks," 2014.

[6] A. Corrado, "Animals-10," 2019. Accessed: 2025-05-26.