Image Recognition using One-Shot Learning

Arindam Shrivastava[1] and Madhavendra Singh Rathore[1]

[1]Department of Computer Science, Indian Institute of Technology Varanasi

May 17, 2023

Abstract

This report explores the concept of one-shot learning, a machine learning technique that aims to classify objects or concepts with limited training data. It discusses the methodologies, and potential applications of one-shot learning in identifying the cattle, highlighting its significance in addressing the data scarcity problem.

1 Introduction

One-shot learning is a specialized machine learning approach that addresses the challenge of training models with limited labeled data. Unlike traditional machine learning methods that require a large number of labeled examples per class, one-shot learning focuses on learning to recognize new instances based on a small set of reference samples. This approach is particularly useful in scenarios where obtaining extensive labeled data for each individual is impractical or costly.

In the project, it was specifically decided to use one-shot learning due to its ability to handle the challenges associated with limited labeled data in cattle identification. Acquiring a large dataset of labeled muzzle images for each individual cow is often a time-consuming and resource-intensive task. By implementing one-shot learning, the project aimed to develop an efficient and accurate cattle identification system that minimizes the need for extensive training data collection.

Furthermore, one-shot learning allows to exploit the unique features present in cattle muzzles effectively. The muzzle region contains distinctive patterns such as fur markings, texture, and shape variations that are specific to individual animals. By training a model using one-shot learning, the model was enable to learn and recognize these unique features, enabling accurate identification even with minimal training data. [1]

This report will delve into the implementation details, performance evaluation, and implications of our one-shot learning-based cattle identification system, highlighting its advantages and contributions to the field.

2 Related Work

Deep learning is a data-driven method and has been researched for computer vision applications in animal production. Deep learning models can capture spatial and temporal dependencies of images/videos through the use of shared-weight filters and can be trained end-to-end without strenuous hand-crafted design of feature extractors , empowering the models to adaptively discover the underlying class-specific patterns and the most discriminative features automatically. The paper [2] reviewed the deep learning-based approaches for classification, object detection and segmentation, pose estimation, and tracking for different kinds of animals such as cattle, pigs, sheep, and poultry.

While previous studies have made significant contributions to cattle identification, the project aims to advance the field by implementing the one-shot learning technique specifically for cattle identification using muzzle datasets. By leveraging deep learning models and the distinctive features present in cattle muzzles, the report aims to overcome the limitations of traditional methods and develop an accurate and efficient cattle identification system that minimizes the need for extensive training data.

One-shot learning techniques have gained attention in the field of computer vision, offering potential solutions to the limited labeled data problem. The subsequent sections of this report will present methodology and results, demonstrating the effectiveness of one-shot learning-based approach for cattle identification using muzzle datasets.

3 Methodology

This section describes the methodology used to implement the Siamese Neural Network for the Formatted Muzzle Dataset.

3.1 Dataset

The Formatted muzzle dataset consists of two folders training and testing. The training folder consists of 264 folders, and the testing folder consists of 50 folders each of which represents a class. Each class has 10 Images. Here the training folder is used to train the model and the testing folder is used as validation/testing data.

3.2 Model Architecture

Figure 1 describes the basic architecture of a Siamese Neural Network. Two images are processed through convolutional networks with exactly same architecture and weights. The feature vectors are obtained as output of the Convolutional networks. These feature Vectors are used to calculate similarity score.

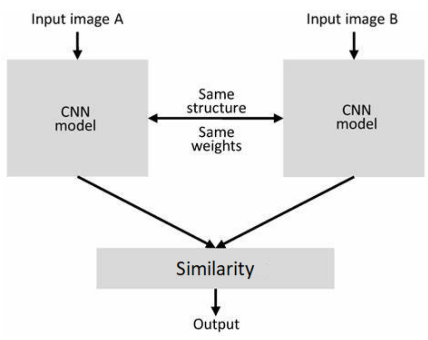

Figure 1: View of the Model.

3.2.1 Convolutional Network

The convolutional network used was inspired from AlexNet [3]. Figure 2 describes the Convolutional Network. The Network consists of:

1. A Conv2D layer with 96 filters of size 11 x 11 and strides of 4
2. A Max Pooling layer with size 3x3 and strides of 2
3. A Conv2D layer with 256 filters of size 5 x 5 and strides of 1
4. A Max Pooling layer with size 3x3 and strides of 2

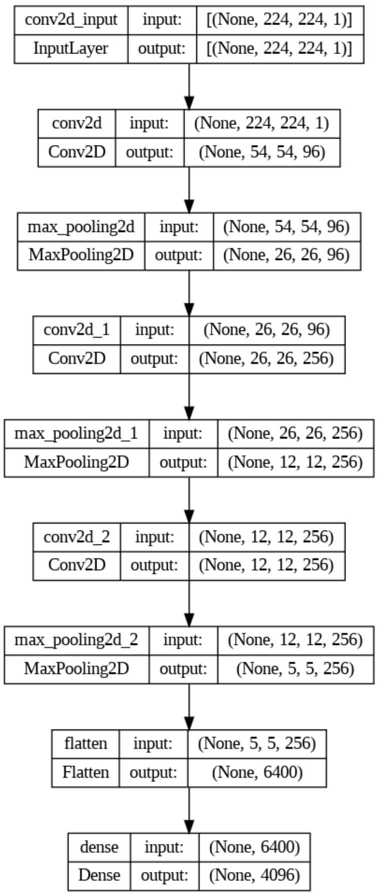

Figure 2: The Convolutional Network

5. A Conv2D layer with 256 filters of size 3 x 3 and strides of 1
6. A Max Pooling layer with size 3x3 and strides of 2
7. A Dense Layer with 4096 units and sigmoid activation

The final layer of the ConvNet outputs the feature vector of the input image.

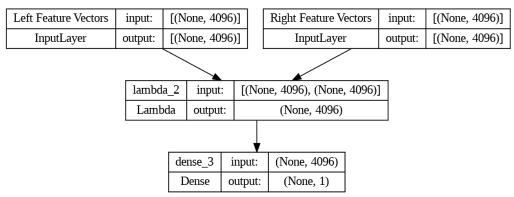

Figure 3: The Similarity Block

3.2.2 Similarity Score

The Similarity score is calculated by taking the absolute difference of the two feature vectors and passing it through a Dense layer with sigmoid activation. Figure 3 describes the architecture.

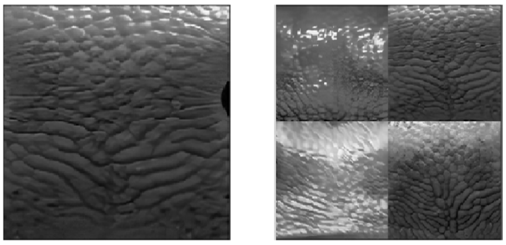

Figure 4: 4 Way One-Shot Task

The Absolute Difference is calculated using a Lambda layer.

3.3 Training

The Model is compiled by using **Binary Cross-entropy** loss function and **Adam** Optimizer with learning rate of 0.00006.

The batches for training are created as an array of pair of images and the labels 0 or 1 showing whether the images are from same class or not. Half of the image pairs are taken from same class while the other half in batch is take from different classes, i.e half the labels are set 1 and other half are set 0.

The model is trained for 250 epochs, validated every 25 epochs. The model with the best validation score is saved.

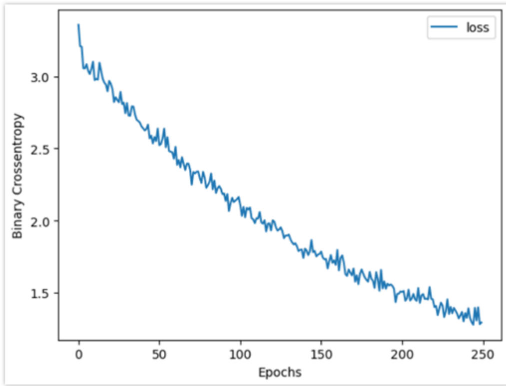

Figure 5: Loss v/s Epoch

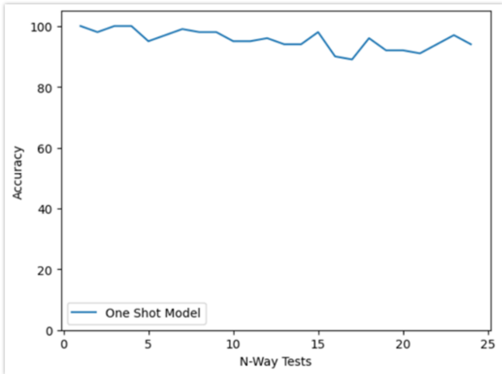

Figure 6: Accuracy v/s N way tests

4 Results

For Testing the model, a one-shot task wast created which consisted of a pair of a Test image and a support set. The support set consisted of *N* images out of which one image is of same class as test image and other *N-1* images are of different class.

The prediction is a correct prediction if the model gives the highest similarity value to the image in support set having same class as the test image. Figure 4 shows a 4 way one shot task where the support set has a size of 4.

Figure 5 shows the graph of loss to epochs and Figure 6 shows the graph of accuracy of the model with N Way Tests.

Conclusions

In conclusion, the project successfully implemented the one-shot learning technique for cattle identification using muzzle datasets. By leveraging the unique features present in cattle muzzles and training a siamese network model, we achieved accurate and efficient identification of individual animals. With one shot learning technique, the accuracy of the model remained almost constant with increasing number of samples and always lied above 90 percent.This approach eliminates the need for labor-intensive and error-prone traditional identification methods. The system offers a practical and reliable solution for livestock management, streamlining agricultural operations, improving animal welfare, and enhancing productivity.

Acknowledgements

We extend our heartfelt gratitude to our project guide and mentor, Niraj Sir, for his invaluable assistance, provision of essential resources, and unwavering support throughout the duration of this project.

We would also like to extend our gratitude to Prof. S. K. Singh for providing us with the opportunity to work on this project.

References

[1] Gregory Koch, Richard Zemel, Ruslan Salakhutdinov, et al. Siamese neural networks for one-shot image recognition. In

ICML deep learning workshop, volume 2. Lille, 2015.

[2] Md Ekramul Hossain, Ashad Kabir, Lihong Zheng, Dave Swain, Shawn McGrath, and Jonathan Medway. A systematic review of machine learning techniques for cattle identification: Datasets, methods and future directions. *Artificial Intelligence in Agriculture*, 2022.

[3] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6):84–90, 2017.