Boosting Ensemble of CNN for Vision

By

Muddasar Yasin

(Registration No: 00000361648)

A thesis submitted to the National University of Sciences and Technology, Islamabad,

in partial fulfillment of the requirements for the degree of

Masters of Science in

Electrical Engineering

Supervisor: Dr. Ahmad Salman

School of Electrical Engineering & Computer Science

National University of Sciences and Technology (NUST)

Islamabad, Pakistan

(2024)

THESIS ACCEPTANCE CERTIFICATE

Certified that final copy of MS/MPhil thesis entitled "Boosting Ensemble of CNN for Vision" written by Muddasar Yasin, (Registration No 361648), of SEECS has been vetted by the undersigned, found complete in all respects as per NUST Statutes/Regulations, is free of plagiarism, errors and mistakes and is accepted as partial fulfillment for award of MS/M Phil degree. It is further certified that necessary amendments as pointed out by GEC members of the scholar have also been incorporated in the said thesis.

Signature: ______________________

Name of Advisor: Dr. Ahmad Salman

Date: 21-Feb-2024

HoD/Associate Dean: ______________________ HOD IAEC SEECS NUST Sector H-12 Islamabad for

Date: 28/03/24

Signature (Dean/Principal): ______________________ Dr. Muhammad Ajmal Principal, NUST School of Electrical Engg & Computer Science H-12, Islamabad

Date: 28 MAR 2024

FORM TH-4

National University of Sciences & Technology

MASTER THESIS WORK

We hereby recommend that the dissertation prepared under our supervision by: (Student Name & Reg. #) Muddasar Yasin [361648]

Titled: Boosting Ensemble of CNN for Vision

be accepted in partial fulfillment of the requirements for the award of Master of Science (Electrical Engineering) degree.

Examination Committee Members

1. Name: Salman Abdul Ghafoor — Signature: ______ 25-Mar-2024 10:02 AM

2. Name: Wajid Mumtaz — Signature: ______ 25-Mar-2024 10:02 AM

Supervisor's name: Ahmad Salman — Signature: ______ 25-Mar-2024 12:50 PM

for Lyin

HOD IAEC
SEECS NUST
Sector H-12 Islamabad

HoD/Associate Dean — Date: 28/03/24

COUNTERSINGED

Date: 28 MAR 2024

Dr. Muhammad Ajmal
Principal
NUST School of Electrical
Engg & Computer Science
H-12, Islamabad

Dean/Principal

Approval

It is certified that the contents and form of the thesis entitled "Boosting Ensemble of CNN for Vision" submitted by Muddasar Yasin have been found satisfactory for the requirement of the degree

Advisor : Dr. Ahmad Salman

Signature: ______________

Date: 21-Feb-2024

Committee Member 1:Dr. Salman Abdul Ghafoor

Signature:

22-Feb-2024

Committee Member 2:Dr. Wajid Mumtaz

Signature: ______________

Date: 22-Feb-2024

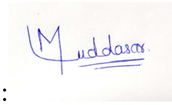
Signature: ______________

Date: ______________

AUTHOR'S DECLARATION

I **Muddasar Yasin** hereby state that my MS thesis titled "**Boosting Ensemble on CNN for Vision**" is my own work and has not been submitted previously by me for taking any degree from National University of Sciences and Technology, Islamabad or anywhere else in the country/ world.

At any time if my statement is found to be incorrect even after I graduate, the university has the right to withdraw my MS degree.

Student Signature: 

Name: Muddasar Yasin

Date: 28th March, 2024

Certificate for Plagiarism

It is certified that PhD/M.Phil/MS Thesis Titled "Boosting Ensemble of CNN for Vision" by Muddasar Yasin has been examined by us. We undertake the follows:

a. Thesis has significant new work/knowledge as compared already published or are under consideration to be published elsewhere. No sentence, equation, diagram, table, paragraph or section has been copied verbatim from previous work unless it is placed under quotation marks and duly referenced.

b. The work presented is original and own work of the author (i.e. there is no plagiarism). No ideas, processes, results or words of others have been presented as Author own work.

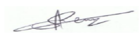
c. There is no fabrication of data or results which have been compiled/analyzed.

d. There is no falsification by manipulating research materials, equipment or processes, or changing or omitting data or results such that the research is not accurately represented in the research record.

e. The thesis has been checked using TURNITIN (copy of originality report attached) and found within limits as per HEC plagiarism Policy and instructions issued from time to time.

Name & Signature of Supervisor

Dr. Ahmad Salman

Signature : 

Dedication

This thesis is dedicated to my family whose boundless love, my esteemed supervisor, Dr. Ahmad Salman whose guidance and expertise have been the cornerstone of my academic journey. This achievement is as much yours as it is mine.

Acknowledgments

Praise be to Allah, the Creator and Sustainer of the Universe, whose power determines honor and humility as He wills. Without His guidance, no endeavor can succeed. From the commencement of my journey at NUST to the day of my departure, I acknowledge His blessings that paved my way and illuminated the path to success. There is no measure that can adequately repay His abundant favors throughout my research period, enabling me to complete it triumphantly.

Contents

List of Figures

List of Tables

Abstract

This study investigates methods to enhance the accuracy of Convolutional Neural Networks (CNNs) through the application of Boosting Ensemble methodologies. The research covers diverse image datasets, including CIFAR-10, CIFAR-100, and Fashion MNIST, aiming to utilize AdaBoost, a widely adopted boosting algorithm, to enhance CNN performance.

A crucial aspect of this research is the assessment of AdaBoost's effectiveness in addressing imbalanced datasets. Imbalanced datasets, marked by uneven class distributions, pose a common challenge in machine learning. Understanding how AdaBoost addresses these scenarios is a central focus of this study.

The empirical findings highlight AdaBoost as a valuable complementary strategy for improving CNN accuracy, especially in cases with imbalanced class distributions. An important observation is that the ensemble model, incorporating AdaBoost, achieves a significant 6% higher test accuracy compared to the baseline CNN. This improvement indicates a substantial enhancement in the model's generalization ability to unseen data.

Chapter 1

Introduction

In recent years, Convolutional Neural Networks (CNNs) have emerged as powerful tools for various computer vision tasks, showcasing exceptional capabilities in tasks such as image recognition, object detection, and semantic segmentation. However, with the continuous growth in the complexity and scale of datasets, there is a rising demand for methodologies that can enhance the robustness and generalization of CNNs.

Boosting, a well-established ensemble learning technique, presents a promising approach to tackle these challenges and enhance the performance of CNNs. This introduction establishes the groundwork for investigating the synergy between boosting ensemble techniques and CNNs. As we embark on this exploration, our aim is to uncover insights that not only contribute to advancing the current understanding of ensemble learning in the context of CNNs but also open avenues for innovative applications in the dynamic fields of computer vision and deep learning.

Ensemble Learning has gained recognition as a powerful strategy to enhance the performance of machine learning models[2]. Ensemble methods operate on the principle of making collective decisions[3]. This involves a group of individual classifiers working together to determine the most appropriate output. The decision-making process can be achieved through voting or averaging probabilities. In the case of voting, each classifier predicts a class, and the final class is determined through a voting mechanism among them. To avoid tie situations, it is recommended to use an odd number of classifiers.

Alternatively, individual classifiers can predict the probability for a class, and the final class is determined by averaging these probabilities. The former approach is termed as hard voting, while the latter is referred to as soft voting.[18] Ensemble methods enhance performance by reducing the variance in prediction errors made by the individual classifiers. We encounter ensemble learning in our daily lives, such as when deciding to watch a movie based on review ratings, which essentially represents a collective decision.

The foundation of ensemble learning lies in the concept of the wisdom of the crowd. This theory suggests that combining knowledge from multiple sources often leads to decisions that are superior to those made by a single entity. In 1990, Schapire[1] introduced a novel approach known as the Adaboost algorithm, which combines several weak learners to function collectively as a strong learner.

Since 2008, researchers have been utilizing ensemble learning approaches to address real-life challenges in various domains, including petrochemicals, bioinformatics, medicine, remote sensing, education, and software bug detection. An ensemble model involves the collaboration of multiple classifiers that train on the same dataset, and their outputs are combined using methods such as weighted averaging, simple averaging, voting, or probability. Ensemble methods leverage this concept in addressing machine learning (ML) problems, working towards predicting the most accurate output compared to relying on a single method.

1.1 Problem Statement

Convolutional Neural Networks (CNNs) have emerged as highly effective tools for computer vision tasks, excelling in image recognition and classification. By utilizing convolutional layers to automatically learn intricate features from raw pixel data, CNNs showcase a unique ability to capture spatial hierarchies. This characteristic makes them pivotal in various applications related to image processing and analysis.

The present forefront of computer vision heavily relies on the utilization of deep learning models, specifically Convolutional Neural Networks (CNNs), for tasks such as image classification, object detection, and semantic segmentation. While individual CNNs have showcased exceptional performance, there is a recognized need to enhance accuracy and robustness. This research aims to overcome the limitations of standalone CNN models by introducing a boosting ensemble approach. The key challenges encompass efficiency, integration of weak classifiers and many more.

The challenge of handling class imbalance in classification scenarios, where certain classes are underrepresented, has prompted exploration into Ensemble Learning techniques such as AdaBoost[8]. AdaBoost, or Adaptive Boosting, sequentially trains weak learners, assigning higher weights to misclassified instances to iteratively improve accuracy. This study investigates the synergy between CNNs, Ensemble Learning, and AdaBoost, seeking to leverage the strengths of both approaches to address the complexities associated with imbalanced datasets.

1.2 Contribution

This research significantly contributes to the realm of computer vision through the exploration of ensemble models for enhancing the accuracy of Convolutional Neural Networks (CNNs). The incorporation of ensemble techniques serves as a distinctive approach to tackle the dynamic challenges presented by intricate and extensive visual datasets.

In merging ensemble models with CNNs, the research endeavors to enhance the predictive capacities of the network, overcoming the constraints inherent in standalone CNN architectures. The inventive use of ensemble learning is designed to harness the varied representations captured by individual models, with the ultimate goal of fostering a CNN that is both more robust and accurate.

Chapter 2

Literature Review

In this section, we examine and evaluate the literature and research papers that have utilised ensemble models in Deep Learning. As the pursuit of enhanced performance continues, researchers have turned their attention toward leveraging the strength of ensemble learning techniques to further boost the capabilities of convolution neural networks.

Gowthami S and Harikumar R[17] focus on the performance analysis of boosting-based transfer learning in deep CNN for image classification, addressing the challenges of imbalanced datasets and improving classifier performance. The experiments conducted on benign and malignant melanoma images from the ISIC database demonstrate the effectiveness of the approach, achieving an accuracy of 99.19 % and a sensitivity of 98.46%. Neelesh Mungoli [19] has proposed an Adaptive Ensemble Learning framework that combines ensemble learning strategies with deep learning architectures to enhance the performance of deep neural networks. By leveraging intelligent feature fusion methods, the framework generates more discriminative and effective feature representations, leading to improved model performance and generalization capabilities.

Tsehay Admassu Assegie [14] proposes a breast cancer prediction model using decision tree and adaptive boosting (Adaboost) algorithms. The model is evaluated using an extensive experimental analysis on a breast cancer dataset from the Kaggle data repository. The dataset consists of 569 observations, with 37.25% being benign and 62.74% being malignant. The class distribution of the dataset is highly imbalanced, leading to poor performance of the decision tree algorithm in predicting malignant observations. To address this, the adaptive boosting algorithm is employed to improve the performance of the decision tree on malignant observations. The analysis shows that the adaptive boosting algorithm achieves an accuracy of 92.53%, while the decision tree algorithm achieves an accuracy of 88.80%. Haoyu Zhang, Yushi Chen, and Xin He [15] proposed a method called Boosting-CNN that combines deep convolutional neural networks (CNNs) and ensemble learning for hyperspectral image (HSI) classification. It uses multiple well-designed CNNs and adaptive boosting to improve classification accuracy. The final classification result is obtained through weighted voting of the CNNs. To address the issue of imbalanced training samples in HSI classification, the paper introduces a soft class balanced loss to mitigate the influence of imbalance. Experimental results on two popular hyperspectral datasets (Salinas and Pavia University) demonstrate that the proposed method achieves better classification accuracy compared to other methods.

Aboozar Taherkhani [13] proposed AdaBoost-CNN, an Adaptive Boosting algorithm that enhances the classification performance of traditional CNN models for multi-class imbalanced datasets using transfer learning techniques. The algorithm achieves improved accuracy, precision, and recall compared to traditional CNN models and outperforms other state-of-the-art algorithms, such as Random Forest and Support Vector Machines, in terms of classification accuracy and F1-score. Shin-Jye Lee [11] presented the usage of a trained deep convolutional neural network model to extract image features and the AdaBoost algorithm to assemble Softmax classifiers, resulting in improved accuracy and reduced retraining time cost. Ricardo Fuentes [20] proposed Adaptive Robust Transfer Learning (ART), a flexible pipeline for transfer learning with machine learning algorithms, providing theoretical guarantees for adaptive transfer and preventing negative transfer. The authors demonstrate the promising performance of ART through empirical studies on regression, classification, sparse learning, and a real-data analysis for a mortality study. Ke Zhao, Feng Jia and Haidong Shao [21] proposed a method called transfer adaptive boosting with squeeze-and-excitation attention convolutional neural network (SEACNN) to address the issue of unbalanced fault diagnosis in rolling bearings. The method combines an SEACNN model for feature extraction and identification, with an AdaBoost algorithm for handling unbalanced fault datasets. Transfer learning is also employed to transfer knowledge from one SEACNN estimator to the next, improving the identification performance. The proposed method is evaluated through extensive experiments, demonstrating its effectiveness in accurately classifying unbalanced datasets in fault diagnosis of rolling bearings. Yuki Kawana, Norimichi Ukita, Jia-Bin Huang, and Ming-Hsuan Yang[10] introduced an approach employing a Convolutional Neural Network (CNN) to capture intricate interdependencies. The network utilizes deep convolution and deconvolution layers to achieve comprehensive representations, resulting in resilient and precise pose estimation. The effectiveness of the proposed ensemble model is assessed on publicly available datasets, showcasing favorable performance in comparison to baseline models and state-of-the-art methods.

Chapter 3

Methodology

The proposed methodology encompasses a dual-phase approach. Initially, a foundational Convolutional Neural Network (CNN) undergoes training on a specified dataset. Subsequently, employing the fundamental tenets of Transfer Learning, the learned weights from the CNN are harnessed to train the ensemble model, thereby augmenting the overall accuracy of the CNN.

3.0.1 Training a CNN

A basic Convolutional Neural Network (CNN) undergoes training through a structured sequence of layers: convolutional layers, pooling layers, and finally fully connected layers[6]. Think of it like building blocks, where the lower layers discern simple things, and the higher layers understand more complex stuff. The initial convolutional layers focus on extracting local details from the input, generating distinct "feature maps" for different aspects. They use shared weights known as "kernel" to map the input to these feature maps. Then, a non-linear function like ReLU or sigmoid is used to improve the results.

After each convolutional layer, a max-pooling layer picks the most important information, reducing the data size and making things easier to handle. Following the convolutional layers, there are fully connected hidden layers that get the important features in a rearranged way. To make this work, the outputs from the convolutional layers are flattened into a single vector. These layers use non-linear functions to add complexity to the decision-making process.

At the top of this setup, there's a logistic regression model that uses the knowledge gathered from the previous layers. Its job is to create a final output, sorting the input into different categories. To do this, it uses the SoftMax function, which turns the output into a probability distribution, showing how likely each category is. This process helps make well-informed decisions.

3.0.2 Ensemble Configuration

After the basic CNN learns some things, an Ensemble of models is created, which is a collection of multiple Convolutional Neural Network (CNN) models. Each CNN model in the ensemble is considered a "weak learner" because it may not be individually highly accurate, but the ensemble aims to combine their strengths for better overall performance. Each CNN model in the ensemble is trained on the entire dataset (both features and labels). During training, the model learns to recognize patterns and relationships within the data that allow it to make predictions. After training each model, its performance is evaluated on the training set. The evaluation involves making predictions on the training set and comparing them to the actual labels.

The error is calculated by measuring the disagreement between the predicted labels and the actual labels. The error indicates how well or poorly the model is performing on the training set. The weights assigned to each model are calculated using the AdaBoost algorithm. AdaBoost assigns higher weights to models that perform well (have lower error) and lower weights to models that perform poorly. Once the models are trained and assigned weights, they are used to make predictions on a test set. For each model, predictions are made, and these predictions are weighted based on the previously assigned weights.

The final predictions for the ensemble are obtained by combining the weighted predictions of each individual model. The models with higher weights contribute more to the final prediction, while those with lower weights contribute less. The rationale behind using an ensemble is that even if individual models are not highly accurate, their diverse perspectives and strengths may complement each other. By combining the predictions of multiple weak learners with different focuses, the ensemble aims to achieve a more robust and accurate prediction on the test set. Flowchart for ensemble model can be seen in figure 1.

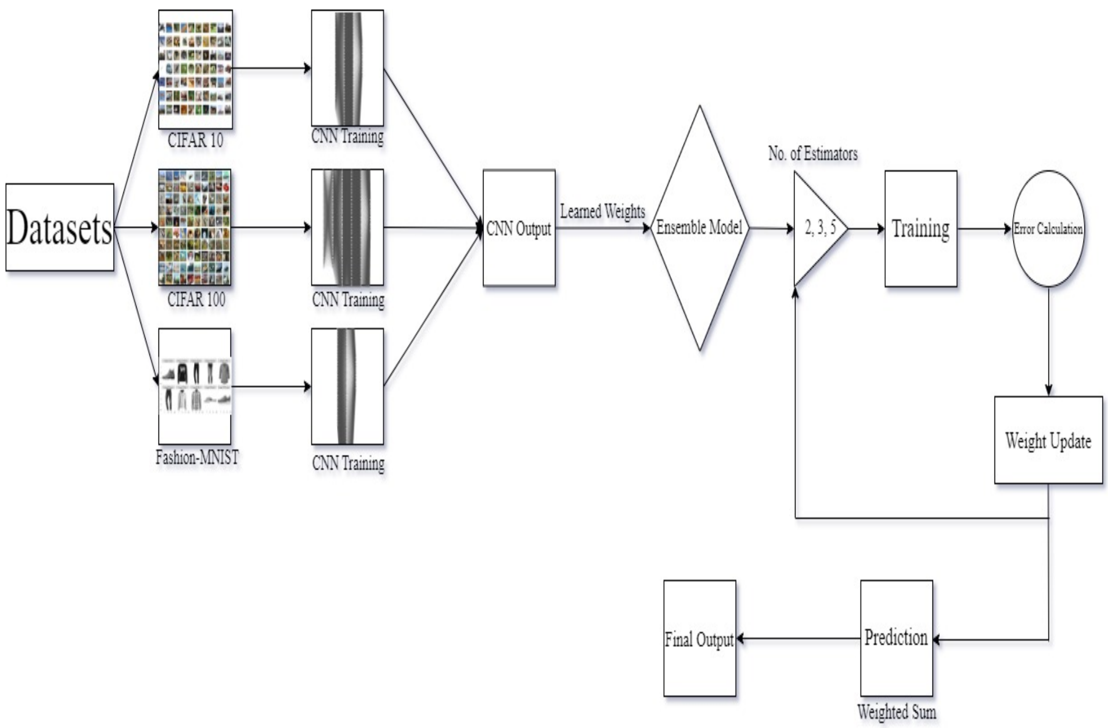


Figure 3.1: Framework of Ensemble Learning

So, If we divide ensemble configuration in simple steps it will be as follows:

1. **Ensemble of CNN models:** The approach begins by forming an ensemble, comprising several Convolutional Neural Network (CNN) models. Each CNN model within the ensemble is regarded as a 'weak learner' due to its potential lack of individual high accuracy. However, the ensemble is designed to amalgamate their respective strengths, ultimately enhancing overall performance

2. **Training Each Model:** Every CNN model in the ensemble undergoes training on the complete dataset, including both features and labels. This training process enables the model to acquire the ability to identify patterns and relationships within the data, facilitating accurate predictions.

3. **Assigning weights based on performance:** Following the training of each model, an assessment of its performance occurs on the training set. This assessment entails generating predictions on the training set and then comparing them with the corresponding actual labels.

Error Calculation: The error is determined by gauging the disparity between the predicted labels and the actual labels. The error serves as a metric indicating the effectiveness or inadequacy of the model on the training set.

4. **AdaBoost Weighting:** The determination of weights assigned to each model is carried out through the AdaBoost algorithm. AdaBoost allocates increased weights to models exhibiting better performance, characterized by lower error rates, and assigns lower weights to models with inferior performance. This strategy is implemented to amplify the impact of accurate models while diminishing the influence of less accurate ones.

5. **Making Predictions on the test set:** After the training and weight assignment for the models, they are employed to generate predictions on a test set. Each model produces predictions, and these predictions are weighted according to the assigned weights.

   Integration of Predictions: The ultimate predictions for the ensemble are derived by amalgamating the weighted predictions from each individual model. Models with higher weights exert a more substantial influence on the final prediction, while those with lower weights have a comparatively diminished impact.

6. **Leveraging Diversity of weak learners:** The justification for employing an ensemble lies in the potential synergy among individual models, even if their individual accuracies may not be exceptionally high. The ensemble leverages the diverse perspectives and strengths of these models to complement each other. Through the amalgamation of predictions from multiple weak learners with distinct focuses, the ensemble aspires to attain a more resilient and accurate prediction on the test set.

Chapter 4

Implementation and Results

In this section the dataset used are explained in detail. Results of ensemble model on Binary dataset and the other dataset are discussed in detail with comparison to other existing techniques.

4.1 Datasets

We have used three datasets for evaluation. CIFAR-10[4], CIFAR-100[4] and Fashion MNIST[9] dataset. An imbalance cats and dogs dataset is also used for comparing the results of AdaBoost on imbalance datasets.

4.1.1 CIFAR10 Dataset

The CIFAR-10 dataset is another popular benchmark dataset in the field of machine learning and computer vision. CIFAR-10 stands for the Canadian Institute for Advanced Research, which is the organization that created the dataset. The "10" in CIFAR-10 represents the number of different classes or categories present in the dataset. CIFAR-10 consists of color images, each of size 32x32 pixels. The dataset is divided into ten classes, and each class represents a different object or category. The classes are: airplane, automobile, bird, cat, deer, dog, frog, horse, ship, and truck. The dataset contains a total of 60,000 images. The images are split into 50,000 for training and 10,000 for testing, providing a standard split for evaluating model performance. Sample images from all classes are shown in figure 4.1.

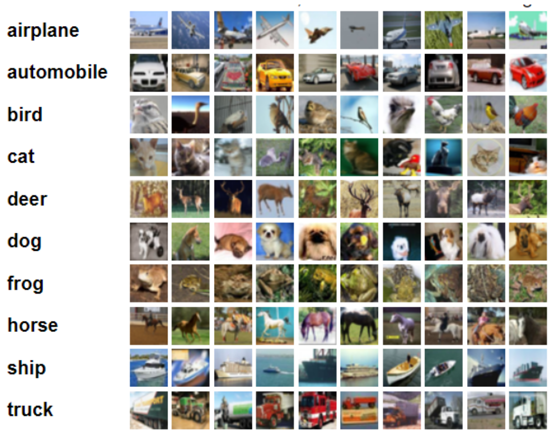


Figure 4.1: Sample Images from each class of CIFAR10

4.1.2 CIFAR100 Dataset

CIFAR-100, or the Canadian Institute for Advanced Research 100, is a widely used dataset in the field of machine learning and computer vision. It is an extension of the CIFAR-10 dataset and consists of 60,000 32x32 color images in 100 different classes, with each class containing 600 images. The dataset is divided into 50,000 training images and 10,000 testing images. Each image in CIFAR-100 belongs to one of the 100 classes, and these classes are further grouped into 20 super classes. The dataset is designed to be challenging, covering a diverse range of object categories. Some examples of classes in CIFAR-100 include "apple," "beaver," "clock," "forest," "man," and "woman."

Superclass	Classes
aquatic mammals	beaver, dolphin, otter, seal, whale
fish	aquarium fish, flatfish, ray, shark, trout
flowers	orchids, poppies, roses, sunflowers, tulips
food containers	bottles, bowls, cans, cups, plates
fruit and vegetables	apples, mushrooms, oranges, pears, sweet peppers
household electrical devices	clock, computer keyboard, lamp, telephone, television
household furniture	bed, chair, couch, table, wardrobe
insects	bee, beetle, butterfly, caterpillar, cockroach
large carnivores	bear, leopard, lion, tiger, wolf
large man-made outdoor things	bridge, castle, house, road, skyscraper
large natural outdoor scenes	cloud, forest, mountain, plain, sea
large omnivores and herbivores	camel, cattle, chimpanzee, elephant, kangaroo
medium-sized mammals	fox, porcupine, possum, raccoon, skunk
non-insect invertebrates	crab, lobster, snail, spider, worm
people	baby, boy, girl, man, woman
reptiles	crocodile, dinosaur, lizard, snake, turtle
small mammals	hamster, mouse, rabbit, shrew, squirrel
trees	maple, oak, palm, pine, willow
vehicles 1	bicycle, bus, motorcycle, pickup truck, train
vehicles 2	lawn-mower, rocket, streetcar, tank, tractor

Figure 4.2: Sample Images from each class of CIFAR100

4.1.3 Fashion MNIST Dataset

Fashion-MNIST is a dataset of Zalando's article images—consisting of a training set of 60,000 examples and a test set of 10,000 examples. Each example is a 28x28 grayscale image, associated with a label from 10 classes. Each training and test example is assigned to one of the following labels:

- T-shirt/top
- Trouser
- Pullover
- Dress
- Coat
- Sandal
- Shirt
- Sneaker
- Bag
- Ankle boot

4.2 Experimental Results

In this section, the experimental test results on the proposed model is explained. Performance of ensemble model is compared with Transfer Learning Algorithm and benchmark CNN using CIFAR-10, CIFAR100 and Fashion MNIST Dataset.

4.2.1 Experimental Results of AdaBoost with Decision Tree

Y Freundand RE Schapire[1] proposed AdaBoost demonstrating significant efficacy in tasks involving binary classification with decision tree as weak classifier, where the primary goal is to distinguish between two distinct classes. The algorithm's adaptability and its focus on misclassified instances during training make it particularly adept at addressing class imbalances. Its design is tailored to enhance the performance of weak classifiers, facilitating the amalgamation of their predictions to construct a robust classifier.

Moreover, AdaBoost showcases adaptability to the underlying data distribution by dynamically adjusting instance weights during the training phase. This flexibility proves beneficial, especially in scenarios where one class is underrepresented, contributing to the algorithm's success in handling imbalanced datasets.

On the contrary, training convolutional neural networks (CNNs) on imbalanced datasets poses challenges, potentially leading to suboptimal model performance. CNNs' inherent bias towards the majority class in the presence of imbalances may result in prioritizing the dominant class, potentially neglecting the minority class and yielding subpar generalization. It's crucial to note that accuracy can be a deceptive metric in imbalanced settings, as high accuracy may be achieved by predominantly predicting the majority class, even if performance on minority classes is inadequate.

While AdaBoost excels in binary and imbalanced data scenarios, utilizing a basic decision tree as a weak classifier may pose limitations in handling complex datasets or multiclass scenarios like CIFAR-10 and MNIST. Decision trees' simplicity may hinder their ability to capture intricate relationships within data, particularly in the presence of diverse classes.

To validate these points, we generated a binary class imbalanced dataset of cat and dog images. Initially balanced, the dataset comprised 279 training images for each class and 70 test images per class, totaling 558 training images and 140 test images. For experimental purposes, intentional efforts were made to create an imbalanced dataset. In this modified version, the training set retained 279 images for the dogs class while intentionally reducing the number of cat images to 71. This deliberate imbalance was introduced to explore and assess the performance of both ensemble models and CNNs under such conditions.

Results of AdaBoost on imbalance binary dataset and CIFAR-10 and MNIST[5] dataset are shown in the following table with varrying number of estimators.

Dataset	Number of Estimators	AdaBoost Accuracy	CNN Accuracy
Cats & Dogs	20	90.00 %	52.14 %
Cats & Dogs	40	96.54 %	52.14 %
MNIST	20	69.67 %	97.22 %
MNIST	40	73.51 %	97.22 %
CIFAR10	20	28.75 %	69.29 %
CIFAR10	20	30.37 %	69.29 %

Table 4.1: Results Using AdaBoost with Decision Tree as weak Classifier

4.2.2 Results for Using Past Knowledge for Better Accuracy:

The approach involves transferring what the first CNN has come to know to a second CNN that is frequently designed for a similar dataset. The aim is to leverage the useful pieces of information that the first CNN revealed concerning similar and unrelated problems. Since using old weights to improve the accuracy increases the speed with which later CNNs can learn, especially when the new job is somewhat similar to the first job. This helps the model to have prior knowledge and hence improve performance even when there is little labeled data. While some old weights can be of value, they might not always be beneficial, even more, so if the jobs are too dissimilar.
Architecture used for training CIFAR10, CIFAR100 and Fashion MNIST are given in the following subsequent tables.

Layers	Filters	Kernel Size	Activation
Conv2D	32	3*3	ReLU
Max Pooling	NA	2*2	NA
Conv2D	64	3*3	ReLU
Max Pooling	NA	2*2	NA
Dense	128	NA	ReLU
Dense	10	NA	Softmax

Table 4.2: CNN Architecture for CIFAR10 Dataset

Layers	Filters	Kernel Size	Activation
Conv2D	32	3*3	ReLU
Max Pooling	NA	2*2	NA
Dense	128	NA	ReLU
Dense	10	NA	Softmax

Table 4.3: CNN Architecture for Fashion MNIST Dataset

Rather than using a pre-trained model, we will first train a CNN and use its weights for the next CNN. We will freeze the convolution layers of the second model and change the fully connected layers. After the evaluation, there was almost 2.91% increase in the accuracy of the 2nd model as compared to the first model. An accuracy comparison for CIFAR10, CIFAR100 and Fashion MNIST is given in Table 4.5.

Layers	Filters	Kernel Size	Activation
Conv2D	32	3*3	ReLU
Conv2D	32	3*3	ReLU
Max Pooling	NA	2*2	NA
Conv2D	64	3*3	ReLU
Conv2D	64	3*3	ReLU
Max Pooling	NA	2*2	NA
Dropout	50%	NA	NA
Dense	512	NA	ReLU
Dropout	50%	NA	NA
Dense	10	NA	Softmax

Table 4.4: CNN Architecture for CIFAR100 Dataset

Dataset	CNN Accuracy	Accuracy Using Previous Weights
CIFAR10	70.13%	73.04%
CIFAR100	43.22%	46.73%
Fashion Mnist	91.61%	91.64%

Table 4.5: Results of Using Previous Weights

4.2.3 Result of Ensemble Model on CIFAR10

Experimental results for CIFAR10 using the ensemble model are discussed in this section. Ensemble model is tested for different number of estimators and all estimators are tested for different number of eppochs. For five estimators and each estimator tested for 15 training epochs ensemble model gave 76.27% accuracy respectively. CNN was trained for 15 epochs. Results for different number of estimators and epochs are given in the following table 4.6:

When number of epochs for CNN were changed to 20 from 15 it resulted in the change of accuracy of CNN. So, number of epochs for ensemble model were kept to 15 to check the effect of change in accuracy of ensemble model compariosn in CNN. Results for accuracy of CNN for 20 epochs and Ensemble model is in table 4.7:

Estimators	No. of Epochs	Ensemble Accuracy	CNN Accuracy
02	15	73.17 %	70.13 %
02	20	71.5 %	70.13 %
03	15	74.97 %	70.13 %
03	20	74.5%	70.13 %
05	15	76.25 %	70.13 %
05	20	76.03 %	70.13 %

Table 4.6: CIFAR10 Dataset Accuracy Using Encemble Model

Estimators	Ensemble Accuracy	Epochs for CNN	CNN Accuracy
02	72.89%	20	68.52%
03	75.46%	20	68.52%
05	76.47%	20	68.52%

Table 4.7: Results with increased CNN Epochs

4.2.4 Result of Ensemble Model on CIFAR100

Experimental results for CIFAR100 using the ensemble model are discussed in this section. Ensemble model is tested for different number of estimators and each estimator is tested for multiple number of epochs to inquire the results. CNN was trained for 15 epochs. Results for different number of estimators and epochs are given in the following table:

Estimators	No. of Epochs	Ensemble Accuracy	CNN Accuracy
02	15	47.33 %	43.22 %
02	20	48.80 %	43.22 %
03	15	47.16 %	43.22 %
03	20	49.23%	43.22 %
05	15	48.00 %	43.22 %
05	20	50.12 %	43.22 %

Table 4.8: CIFAR100 Dataset Accuracy Using Encemble Model

When number of epochs for CNN training were changed but number of epochs for ensemble model were kept to 15 accuracy of Ensemble model also changed with respect to CNN accuracy.

Estimators	Ensemble Accuracy	Epochs for CNN	CNN Accuracy
02	46.22%	20	45.62%
03	47.22%	20	45.62%
05	47.88%	20	45.62%

Table 4.9: Results with increased CNN Training Epochs

4.2.5 Result of Ensemble Model on Fashion MNIST

Experimental results for Fashion MNIST using the ensemble model are discussed in this section. Different numbers of CNN estimators in the AdaBoost and different numbers of learning epochs for each estimator are tested. CNN was trained for 15 epochs. Results for different number of estimators and epochs are given in the following table:

Estimators	Estimator Epochs	Ensemble Accuracy	CNN Accuracy
02	15	92.2 %	91.60 %
02	20	92.6 %	91.60 %
03	15	92.3 %	91.60 %
03	20	92.8%	91.60 %
05	15	92.7 %	91.60 %
05	20	92.5 %	91.60 %

Table 4.10: Fashion MNIST Accuracy Using Encemble Model

4.3 Comparison of Transfer Learning & Ensemble Model

Transfer learning seeks to improve the performance of target learners in specific domains by leveraging knowledge from different yet related source domains[16]. The goal is to enhance a learner in one domain by transferring valuable information from a related domain. In cases where obtaining training data is expensive or challenging, there is a need to develop high-performance learners trained with readily available data from diverse domains, commonly referred to as transfer learning[7].

In the realm of traditional machine learning, both training and testing data typically share the same input feature space and data distribution. Discrepancies in data distribution between the two sets can result in a degradation of the predictive learner's performance. The necessity for transfer learning arises when there is a limited supply of target training data, attributed to factors such as data rarity, high costs associated with data collection and labeling, or the inaccessibility of the data. In our experiments, we applied transfer learning to train models on the CIFAR-10 and CIFAR-100 dataset using ResNet50 and AlexNet[6] as pre-trained models[12]. The accuracy comparison of transfer learning using various models, along with an ensemble model and a simple CNN, is presented in the table below.

Dataset	AlexNet	ResNET50	Ensemble Accuracy
CIFAR10	36.12%	32.08%	76.47%
CIFAR100	44.10%	48.82%	50.12%

Table 4.11: Comparison of Transfer Learning & Ensemble Model

Utilizing transfer learning is a beneficial strategy to harness knowledge from related domains, yet it presents challenges related to adaptability and domain mismatch. The reliance on pre-trained models in transfer learning may hinder adaptability to the unique characteristics of the target dataset. The knowledge transferred from the source domain may not seamlessly align with the nuances of the target domain. This methodology assumes a shared set of features between the source and target domains. However, if there is substantial dissimilarity between the domains, the transferred knowledge may not effectively contribute to the target task. Neha Sharma,Vibhor Jain and Anju Mishra[12] concluded in their results that higher number of layers are required to get higher accuracy. The findings indicated that networks trained through transfer learning performed better than existing ones, demonstrating elevated accuracy rates. Specific objects such as "chair," "train," and "wardrobe" achieved flawless recognition with 147-layered networks, while objects like "cars" exhibited perfect recognition with 177-layered networks[12]. Additionally, the implementation of transfer learning often entails the use of pre-trained models, which can exhibit complex architectures.

Chapter 5

Conclusion and Future Work

5.1 Conclusion

This research focus on the use of an ensemble model, particularly incorporating AdaBoost, has emerged as an effective strategy for enhancing the accuracy of Convolutional Neural Networks (CNNs). The primary objective of this research was to boost the performance of CNNs by leveraging the strengths of diverse models through ensemble learning. The results obtained have showcased significant advancements compared to standalone CNNs.

A comparative analysis between ensemble models and alternative techniques, such as transfer learning, indicated that the ensemble approach not only surpassed in terms of accuracy but also demonstrated a reduction in the number of parameters. This reduction is particularly noteworthy as it directly translates into a decrease in computational costs, rendering the ensemble model more resource-efficient and practical for real-world applications.

A notable aspect of this study is the successful training of the AdaBoost on an imbalanced dataset. The AdaBoost approach exhibited superior results in addressing class imbalances compared to the standalone CNN. This implies that AdaBoost, as a component of the ensemble, contributes to the model's robustness in scenarios where class distribution is uneven.

5.2 Future Work

Moving forward, potential research directions in this domain could explore diverse avenues for further improvement. Firstly, investigating alternative ensemble techniques beyond AdaBoost, such as bagging or stacking, could yield additional insights into optimal model combinations for enhancing CNN performance. Additionally, exploring the impact of varying ensemble sizes and incorporating different base models within the ensemble may lead to the identification of more effective configurations.

Furthermore, addressing the interpretability of ensemble models remains a crucial aspect for broader adoption in real-world applications. Developing methodologies to interpret and explain the decisions made by the ensemble could enhance the model's trustworthiness and applicability in sensitive domains.

Finally, with the continuous evolution of technology, integrating ensemble models with emerging techniques like neural architecture search (NAS) or automated machine learning (AutoML) could pave the way for more efficient and adaptive models. These approaches have the potential to automate the process of selecting optimal architectures and hyperparameters, thereby reducing the burden on practitioners.

References

[1] Yoav Freund and Robert E Schapire. "A desicion-theoretic generalization of on-line learning and an application to boosting". In: *European conference on computational learning theory*. Springer. 1995, pp. 23–37.

[2] TG Dietterich et al. "Ensemble learning. The handbook of brain theory and neural networks". In: *Arbib MA* 2.1 (2002), pp. 110–125.

[3] Robi Polikar. "Ensemble based systems in decision making". In: *IEEE Circuits and systems magazine* 6.3 (2006), pp. 21–45.

[4] Alex Krizhevsky, Geoffrey Hinton, et al. "Learning multiple layers of features from tiny images". In: (2009).

[5] Li Deng. "The mnist database of handwritten digit images for machine learning research [best of the web]". In: *IEEE signal processing magazine* 29.6 (2012), pp. 141–142.

[6] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. "Imagenet classification with deep convolutional neural networks". In: *Advances in neural information processing systems* 25 (2012).

[7] Karl Weiss, Taghi M Khoshgoftaar, and DingDing Wang. "A survey of transfer learning". In: *Journal of Big data* 3.1 (2016), pp. 1–40.

[8] Wonji Lee, Chi-Hyuck Jun, and Jong-Seok Lee. "Instance categorization by support vector machines to adjust weights in AdaBoost for imbalanced data classification". In: *Information Sciences* 381 (2017), pp. 92–103.

[9] Han Xiao, Kashif Rasul, and Roland Vollgraf. "Fashion-mnist: a novel image dataset for benchmarking machine learning algorithms". In: *arXiv preprint arXiv:1708.07747* (2017).

[10] Yuki Kawana et al. “Ensemble convolutional neural networks for pose estimation”. In: *Computer Vision and Image Understanding* 169 (2018), pp. 62–74.

[11] Shin-Jye Lee et al. “Image classification based on the boost convolutional neural network”. In: *IEEE Access* 6 (2018), pp. 12755–12768.

[12] Neha Sharma, Vibhor Jain, and Anju Mishra. “An analysis of convolutional neural networks for image classification”. In: *Procedia computer science* 132 (2018), pp. 377–384.

[13] Aboozar Taherkhani, Georgina Cosma, and T Martin McGinnity. “AdaBoost-CNN: An adaptive boosting algorithm for convolutional neural networks to classify multi-class imbalanced datasets using transfer learning”. In: *Neurocomputing* 404 (2020), pp. 351–366.

[14] Tsehay Admassu Assegie, R Lakshmi Tulasi, and N Komal Kumar. “Breast cancer prediction model with decision tree and adaptive boosting”. In: *IAES International Journal of Artificial Intelligence* 10.1 (2021), p. 184.

[15] Haoyu Zhang et al. “Boosting CNN for Hyperspectral Image Classification”. In: *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*. IEEE. 2021, pp. 3673–3676.

[16] Fuzhen Zhuang et al. “A Comprehensive Survey on Transfer Learning”. In: *Proceedings of the IEEE* 109.1 (2021), pp. 43–76. DOI: 10.1109/JPROC.2020.3004555.

[17] S Gowthami and R Harikumar. “Performance analysis of Incremental boosting based Transfer Learning in Deep CNN”. In: *2022 3rd International Conference on Communication, Computing and Industry 4.0 (C2I4)*. IEEE. 2022, pp. 1–6.

[18] Suyash Kumar, Prabhjot Kaur, and Anjana Gosain. “A comprehensive survey on ensemble methods”. In: *2022 IEEE 7th International conference for Convergence in Technology (I2CT)*. IEEE. 2022, pp. 1–7.

[19] Neelesh Mungoli. “Adaptive Ensemble Learning: Boosting Model Performance through Intelligent Feature Fusion in Deep Neural Networks”. In: *arXiv preprint arXiv:2304.02653* (2023).

[20] Boxiang Wang, Yunan Wu, and Chenglong Ye. “The ART of Transfer Learning: An Adaptive and Robust Pipeline”. In: *Stat* (2023), e582.

[21] Ke Zhao, Feng Jia, and Haidong Shao. "Unbalanced fault diagnosis of rolling bearings using transfer adaptive boosting with squeeze-and-excitation attention convolutional neural network". In: *Measurement Science and Technology* 34.4 (2023), p. 044006.