Transfer-Ensemble based Deep learning approach for improved Weather Image Classification Task

Thesis submitted in partial fulfillment of the requirements for the degree of

Master of Engineering
in
Computer Science and Engineering

Submitted by

Dibyendu Chattoraj
Regn. No. - 163564 of 2022 - 2023
Exam Roll No. - M4CSE24010
Class Roll No. - 002210502013
Session - 2022 - 2024

Under the supervision of

Prof. Susmita Ghosh

Department of Computer Science and Engineering
Jadavpur University
188, Raja S.C. Mallick Rd,
Kolkata - 700032, West Bengal, India

Department of Computer Science and Engineering
Faculty of Engineering and Technology
Jadavpur University, Kolkata - 700032

CERTIFICATE OF RECOMMENDATION

This is to certify that the work embodied in this thesis entitled **"Transfer-Ensemble based Deep learning approach for improved Weather Image Classification Task"** has been satisfactorily completed by **Dibyendu Chattoraj** (Registration Number 163564 of 2022 - 2023; Class Roll No. 002210502013; Examination Roll No. **M4CSE24010**). It is a bonafide piece of work carried out under my supervision and guidance at Jadavpur University, Kolkata, for partial fulfilment of the requirements for the awarding of the **Master of Engineering** in **Computer Science and Engineering** degree of the Department of Computer Science and Engineering, Faculty of Engineering and Technology, Jadavpur University, during the academic year 2022 - 2024.

_____

Prof. Susmita Ghosh,
Department of Computer Science and Engineering,
Jadavpur University.
(Supervisor)

Forwarded By:

_____

Prof. Nandini Mukherjee (Mukhopadhyay),
HOD, Department of Computer Science and Engineering,
Jadavpur University Kolkata 700032.

_____

Prof. Dipak Laha,
DEAN, Faculty of Engineering and Technology,
Jadavpur University Kolkata 700032.

Department of Computer Science and Engineering
Faculty of Engineering and Technology
Jadavpur University, Kolkata - 700032

CERTIFICATE OF APPROVAL

This is to certify that the thesis entitled **"Transfer-Ensemble based Deep learning approach for improved Weather Image Classification Task"** is a bonafide record of work carried out by **Dibyendu Chattoraj** (Registration Number 163564 of 2022- 2023; Class Roll No. 002010502013; Examination Roll No. **M4CSE24010**) in partial fulfilment of the requirements for the award of the degree of **Master of Engineering** in **Computer Science and Engineering** in the Department of Computer Science and Engineering, during the period of September 2022 to May 2024. It is understood that by this approval, the undersigned do not necessarily endorse or approve any statement made, opinion expressed or conclusion drawn therein, but approve the thesis only for the purpose of which it has been submitted.

_____

Signature of Examiner

Date:

_____

Signature of Supervisor

Date:

Department of Computer Science and Engineering
Faculty of Engineering and Technology
Jadavpur University, Kolkata - 700032

DECLARATION OF ORIGINALITY AND COMPLIANCE OF ACADEMIC ETHICS

I hereby declare that the thesis entitled **"Transfer-Ensemble based Deep learning approach for improved Weather Image Classification Task"** contains literature survey and original research work by the undersigned candidate, as a part of his degree of **Master of Engineering** in **Computer Science and Engineering, JadavpurUniversity**. All the information has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all materials and results that are not original to this work.

Name: Dibyendu Chattoraj

Examination Roll No.: M4CSE24010

Registration No.: 163564 of 2022 - 2023

Thesis Title: Transfer-Ensemble based Deep learning approach for improved Weather Image Classification Task

_____

Signature of the Candidate:

Acknowledgements

I would like to start by thanking my parents, my brother, and my teachers for helping me deploy all the right resources and for shaping me into a better human being. I would like to express my deepest gratitude to my advisor, **Prof. Susmita Ghosh**, Department of Computer Science and Engineering, Jadavpur University for her admirable guidance, care, patience, mental support and for providing me with an excellent atmosphere for doing research. My thanks to Mr. **Abhiroop Chatterjee**, a devoted tutee of **Prof Susmita Ghosh**. Iam deeply grateful to them for the discussions that helped to enrich the content of this thesis. I am also thankful to Laboratory for giving me the proper laboratory facilities for carrying out my work. This thesis would not have been completed without the inspiration and support of a number of wonderful – my thanks and appreciation to all of them for being part of this journey and making this thesis possible.

Dibyendu Chattoraj

Examination Roll No.: **M4CSE24010**

Registration No.: 163564 of 2022 - 2023

Department of Computer Science and Engineering,

Jadavpur University.

Abstract

Weather prediction is crucial for various domains, including transportation systems, particularly for autonomous vehicles. Accurate forecasts aid in anticipating traffic flow, ensuring safety, and mitigating the impacts of severe weather on vehicle operations. Early warning systems rely on precise weather forecasts to minimize accidents and congestion, highlighting the importance of predictive technologies in enhancing transportation resilience.

In this research work, the goal is to propose a weather phenomenon classification approach using the ensemble techniques and state of the art deep neural networks namely Inception v3, RestNet 50 and VGG16 deep neural network models. The goal is to categorize weather events into 11 distinct classes, leveraging deep learning for improved accuracy. Focusing on outperforming existing models on the Weather Phenomena Dataset (WEAPD), the study employs data augmentation ensemble techniques and meticulous hyper- parameter tuning. The optimized VGG16 model with data augmentation individually achieves an impressive overall accuracy of 92.10%, surpassing other established models. Moreover, a novel ensemble technique of geometric harmonic mean (GHE) of the three deep neural networks (VGG16, RestNet50 and Inception V3) has achieved an overall best accuracy of 92.82% Another approach for making better decision in ensembled models is by using cubic custom mean which gives an average classification of 92.86%. This research contributes to advancing weather event classification accuracy, with potential implications for meteorology and related applications. We have also validated and refined the model through iterative experimentation, ensuring an effective balance between model complexity and efficient training.

Contents

Abstract i

Table of Contents ii

List of Figures iv

List of Tables vi

List of Abbreviations vii

1 Chapter 1: Introduction and Background ……1
- 1.1 Weather Phenomena's.....1
- 1.2 Importance of Weather Classification 2
- 1.3 Deep Learning 3
- 1.4 Deep Learning Networks 4
  - 1.4.1 Convolution Neural Networks ……..…5
- 1.5 Transfer Learning Techniques 7
- 1.6 Overfitting 8
- 1.7 Image Augmentation 9
- 1.8 Ensemble Techniques ……...…….....10
- 1.9 Novel Approach ……...……11
- 1.10 Contribution ……...……...12
- 1.11 Objectives And aims ……12

2 Chapter 2: Literature Survey.…….……...14
- 2.1 Related Research and Previous Works 14
- 2.2 Comparative Analysis of Previous Research 21

3 Chapter 3 : Proposed Methodology.…….…….22
- 3.1 VGG-16 Model ……22
- 3.2 VGG-16 With Data Augmentation ……...…….....23
- 3.3 Ensemble Techniques …….…….....27

4 Chapter 4: Dataset and Experimental Setup.……...….29
- 4.1 Dataset Used …….……29
- 4.2 Dataset Preprocessing ……30
- 4.3 Experimental Setup ……30
- 4.4 Fine-tuned parameters ……....34

5 Chapter 5: Results and Analysis.…….…….....…….35
- 5.1 VGG-16 Results ……...35
- 5.2 VGG-16 With Data augmentation Results……36
- 5.3 Ensemble Results. …….…….....…...37
- 5.4 Comparison Analysis …….…….…..……40
- 5.5 Comparison With Other State of The Arts Models ……41
- 5.6 Discussion ……...……...41

6 Chapter 6: Conclusion and Future Scope.……43
6.1 Conclusions Of the Present Work 43
6.2 Scope Of Future Works 43

Bibliography

List of Figures

1 Weather Phenomena's images 2

2 Schematic diagram of a basic convolutional neural network 6

3 Transfer learning processnos..... .7

4 Overfitting Curve ……..…...….9

5 VGG -16 Architecture..... 23

6 (a) Original Image of dew Class ……..……26

6 (b) Augmented images of dew class 26

7 VGG16 with data augmentation flow chart..... 27

8 Ensemble Learning flow chart..... 28

9 (a) Feature Map of 1st Convolutional layer of VGG16 trained model….…….31

9 (b) Feature Map of last three Convolutional layers of VGG16 trained model ……..……32

10 (a) Feature Map of 1st Convolutional layer of Inception V3 trained model….32

10 (b) Feature Map of last three Convolutional layers of Inception V3 trained model……....…….33

11 (a) Feature Map of 1st Convolutional layer of Rest Net 50 trained model. 33

11 (b) Feature Map of last three Convolutional layers of Rest Net 50 trained model……...34

12 (a) Accuracy and Loss Curves for a simulation of VGG16 35

12 (b) ROC curve and Confusion matric for VGG-16 Model without augmentation…….36

13 (a) Accuracy and Loss Curves for a simulation of VGG16 37

13 (b) ROC curve and Confusion matric for VGG-16 Model with augmented weather data……..…37

14 Accuracy and Loss Curves for a simulation of (a) VGG16 38

14 Accuracy and Loss Curves for a simulation of (b) RestNet ……39

14 Accuracy and Loss Curves for a simulation of (b) Inception v3 …….39

15 The ROC curve and confusion matrix of the proposed Geometric – Harmonic Ensemble Model on the testing set. 40

List of Tables

1 Distribution of images among classes in the WEAPD Dataset……....29

2 Experimental setup and Hyperparameters Used. ……..….30

3 Performance scores without augmentation…….35

4 Performance scores with augmentation……...…...36

5 Performance scores with Maximum Ensemble having 92.02%.....37

6 Performance scores with GHE ensemble having average 92.75% accuracy ……38

7 Performance scores with Average Ensemble having 92.52%.....38

8 Comparison with state-of-the-art methods 41

List of Abbreviations

WEAPD Weather Phenomena Dataset (WEAPD)

GHE – Geometric Harmonic Mean Ensemble

VGG - Visual Geometry Group

CNN - Convolutional Neural Networks

CHAPTER 1

Introduction and Background

1.1 Weather Phenomena's

Throughout human history, the pursuit of understanding and predicting the world has been an enduring endeavor, offering valuable competitive advantages in various domains, particularly in weather forecasting. Weather conditions, encompassing elements like temperature, humidity, and wind, profoundly influence many aspects of human livelihood.

The analytical support provided by weather forecasting has far-reaching implications for intelligent transportation systems. By predicting traffic flow, authorities can optimize routes, reduce congestion, and enhance overall transportation efficiency. Air visibility analysis helps mitigate risks associated with reduced visibility in adverse weather conditions, reducing the likelihood of accidents and improving travel safety. Furthermore, as autonomous vehicles become increasingly prevalent, their reliance on accurate weather predictions intensifies. These vehicles must navigate through various weather scenarios, making real-time and precise forecasting paramount for ensuring their safe and efficient operation. By anticipating weather conditions, autonomous vehicles can adjust their driving strategies and respond to hazards appropriately, thus reducing the chances of accidents and enhancing passenger safety. Traditional methods of weather phenomena classification heavily rely on human observation, which is time-consuming and prone to errors. As a result, there is a pressing demand for the development of automated technologies that are highly precise, efficient, and capable of accurately classifying weather phenomena. However, ordinary machine learning approaches have limitations in learning the complex characteristics of weather phenomena. Weather phenomena exhibit intricate patterns and dynamics that are difficult to capture using conventional machine learning methods. The conventional machine learning algorithms often struggle to discern the intricate relationships between various weather variables, leading to suboptimal classification accuracy.

Moreover, weather patterns can be highly nonlinear and dynamic, making it challenging for traditional machine learning models to capture the full extent of their complexity.



Fig.1. Weather Phenomena's images

Additionally, weather data is inherently high-dimensional and spatiotemporal, requiring advanced techniques to effectively handle such large and complex datasets. Ordinary machine learning models may lack the capacity to process and extract relevant features from such data comprehensively. As a result, their ability to accurately classify weather phenomena is limited. To overcome these challenges, researchers and scientists are exploring advanced artificial intelligence techniques, such as deep learning and neural networks. These methods have shown promising results in various fields, including image recognition and natural language processing. By leveraging the power of deep learning, it becomes possible to extract intricate patterns and relationships from weather data, enabling more accurate classification of weather phenomena.

1.2 Importance Of Weather Classification

The significance of weather prediction extends beyond our day-to-day plans. It plays a crucial role in supporting intelligent transportation systems, aiding in traffic flow prediction and air visibility analysis. Autonomous vehicles heavily rely on accurate sensing and prediction of external environmental factors. Severe weather conditions, such as heavy rain or dense fog, increase the likelihood of traffic accidents and congestion. Therefore, precise and timely weather forecasts are fundamental for early warning systems that can mitigate the impacts of weather-

related disasters on autonomous vehicles, like urban flooding or sewage overflow.

The analysis and classification of weather phenomena has a significant impact on wide range of applications, spanning from environmental monitoring to weather forecasting and assessing environmental quality. Each type of weather phenomenon has distinct effects on various sectors, making accurate distinction essential for optimizing agricultural planning and improving numerous aspects of daily life.

Agriculture is significantly impacted by different weather phenomena, making precise classification crucial for farmers' decision-making processes. Understanding the weather patterns allows for better crop management, irrigation planning, and pest control strategies, ultimately leading to improved agricultural productivity and sustainability. Weather phenomena also hold substantial influence over vehicle assistant driving systems and daily activities. For instance, knowing the weather conditions helps drivers anticipate potential hazards, such as slippery roads during rainfall or reduced visibility in dense fog. Likewise, individuals rely on weather forecasts to make informed choices about clothing, travel plans, and outdoor activities. Solar technologies are heavily affected by weather conditions, particularly solar panels' efficiency and energy production. Accurate weather classification aids in predicting cloud cover, solar radiation, and temperature changes, thus optimizing the use of solar energy and contributing to a more sustainable energy future.

Additionally, weather conditions significantly impact the functionality of visual systems, including outdoor video surveillance. Understanding weather phenomena allows for adjustments to surveillance systems, compensating for factors like rain or fog that might hinder visibility. Furthermore, the interconnectivity of weather conditions cannot be overlooked. Weather from the previous day can have a cascading effect on upcoming days, influencing atmospheric conditions and potentially leading to prolonged weather patterns. Understanding these interconnected phenomena is crucial for meteorologists to develop more accurate and reliable weather forecasts. Moreover, certain local or regional weather conditions pose risks to traffic safety, contributing to accidents on expressways. By effectively classifying weather phenomena, transportation authorities can implement appropriate measures to enhance traffic safety and reduce the likelihood of weather-related accidents.

1.3 Deep Learning

The limitations of traditional machine learning methods in accurately learning the characteristics of weather phenomena have highlighted the need for advanced

technologies. Ordinary machine learning approaches struggle to capture the intricate patterns and dynamics inherent in weather data. As a result, the research focus has shifted towards leveraging deep learning and neural networks to develop high-precision, efficient, and automated technologies for weather phenomena classification. By harnessing the power of deep learning algorithms, scientists and meteorologists aim to enhance weather forecasting and improve our understanding of weather patterns and their impact on various aspects of human life.

Deep learning algorithms, particularly convolutional neural networks (CNNs), are well-suited for handling high-dimensional data and can automatically learn relevant features directly from raw weather data. This capacity for automated feature extraction enables deep learning models to capture the complex spatiotemporal patterns present in weather data, leading to improved classification performance.

Deep neural networks were employed, which build upon the aspects of the shallow neural networks used previously. The layers in the deep networks were fully connected, ensuring comprehensive exploration of feature combinations during the learning process. Similar to the shallow networks, the deep networks utilized on-line learning, updating weights after processing each individual data point, and used the tanh activation function for the layers. The scaled conjugate gradient descent (SCG) backpropagation algorithm, known for its quicker convergence and computational efficiency, was also employed by the deep networks.

1.4 Deep Learning Networks

Deep learning networks, a subset of machine learning, have revolutionized the field by enabling computers to learn and make decisions from data in a manner that mimics the human brain. These networks are characterized by their depth, as they consist of multiple layers of interconnected nodes, or artificial neurons, that process and transform input data to produce desired output.

Key components of deep learning networks include:

Neural Networks: At the core of deep learning are artificial neural networks, inspired by the structure and function of the human brain. These networks consist of layers of interconnected nodes, with each connection having a weight that is adjusted during the training process.

Deep Architecture: Unlike traditional machine learning models, deep learning networks have multiple hidden layers, allowing them to learn complex hierarchical representations of data. The depth of these networks enables them to automatically

extract relevant features from raw input, making them highly effective for tasks such as image and computer vision tasks.

Training with Backpropagation: Deep learning networks are trained using a process called backpropagation. During this training, the network is exposed to labeled data, and the difference between the predicted output and the true output (known as the loss) is calculated. The network then adjusts its weights using optimization algorithms to minimize this loss, iteratively improving its ability to make accurate predictions.

Activation Functions: Nodes within a deep learning network typically apply activation functions to their input. These functions introduce non-linearities into the model, allowing it to learn complex relationships and patterns in the data.

Convolutional Neural Networks (CNNs): Commonly used for image and video processing, CNNs are a type of deep learning network that leverages convolutional layers to automatically learn spatial hierarchies of features.

Transfer Learning: Deep learning networks can leverage pre-trained models on large datasets, allowing them to benefit from knowledge gained in one domain to improve performance on a different but related task.

Deep learning networks have achieved remarkable success across various domains, including computer vision, natural language processing, speech recognition, and game playing. Their ability to automatically learn intricate representations from data has contributed to breakthroughs in technology and has led to significant advancements in fields such as healthcare, finance, and autonomous systems. Despite their success, challenges remain, including the need for large labeled datasets, computational resources, and interpretability of complex models. Continued research and development in the field of deep learning are ongoing, with the potential for further innovations and applications in the future.

1.4.1 Convolution Neural Networks

Convolutional Neural Network (CNN) is a deep learning algorithm that has proven highly effective in providing powerful feature representations for images. Its success can be attributed to several key components, including its deep structure, local receptive fields, spatial subsampling, and shared weights. The deep structure of CNN allows it to learn hierarchical representations of images, starting with simple features like edges and gradually building up to more complex and abstract features. This enables CNN to capture intricate patterns and relationships within the

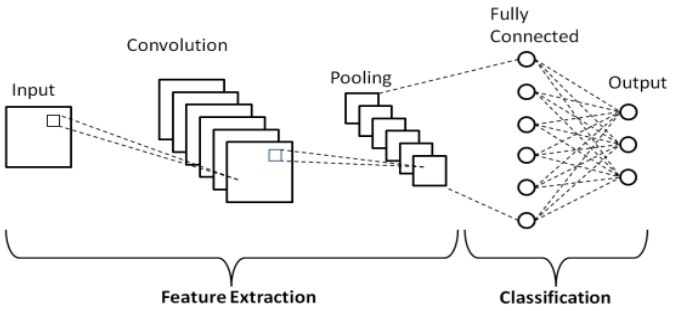


Fig.2. Schematic diagram of a basic convolutional neural network (source: www.researchgate.net)

data, leading to superior image classification performance. Local receptive fields in CNN allow the network to focus on small, localized regions of the image at a time. By convolving small filters over the input, the network can efficiently learn local features, enhancing its ability to detect specific patterns in different areas of an image. Spatial sub-sampling, achieved through techniques like max-pooling, reduces the spatial dimensions of the feature maps while retaining the most relevant information as show in Fig 2. This process helps in making the network more robust to spatial translations and increases its resistance to variations in object positions within an image. The use of shared weights in CNN significantly reduces the number of parameters, making it computationally efficient and scalable. Shared weights allow the network to learn and recognize similar patterns across different parts of the image, contributing to its ability to generalize well to unseen data.

Since the groundbreaking success of the Alex Net model in the ImageNet Large-Scale Visual Recognition Challenge (ILSVRC), CNN has become the go-to architecture for various computer vision tasks. It has found applications in fields like face recognition, where it can learn discriminative features for accurate identification. CNN is also widely used for regression prediction, allowing it to predict continuous values, such as predicting the age or position of objects in an image. Additionally, CNN's ability to detect and localize objects within images makes it an indispensable tool for object detection tasks. CNN's deep structure, local receptive fields, spatial sub-sampling, and shared weights contribute to its exceptional performance in image-related tasks. Its success in the ILSVRC competition opened doors for its application in various domains, including face recognition, regression prediction, and object detection, revolutionizing the field of computer vision and impacting numerous real-world applications.

1.5 Transfer Learning Techniques

Transfer learning involves the reuse of a pre-trained model on a source task to aid learning in a target task. This approach is particularly beneficial when the source and target tasks share some underlying patterns or features. There are several transfer learning techniques, and they can be broadly categorized into the following:

Feature Extraction: In feature extraction, a pre-trained model, typically trained on a large dataset for a specific task, is used as a fixed feature extractor. In Fig 3 The pretrained model on ImageNet dataset is used as a feature extractor for military images. The earlier layers of the model, which capture generic and low-level features, are retained, while the final layers are replaced or retrained for the target task.

Fine-tuning: Fine-tuning involves adjusting the weights of the pre-trained model on the source task to better suit the target task. It allows the model to adapt its learned representations to the nuances of the new task by updating the weights through additional training on the target dataset. Fine-tuning is particularly useful when the target task has some unique characteristics not well captured by the source task.

Domain Adaptation: Domain adaptation aims to transfer knowledge from a source domain to a target domain, even when the distribution of data between the two domains differs. Techniques in domain adaptation attempt to align the feature distributions between the source and target domains, mitigating the impact of domain shift on model performance.

Multi-Task Learning: In multi-task learning, a model is trained on multiple tasks simultaneously, with the idea that knowledge gained from one task can improve the performance on another. This approach is effective when there is a shared underlying representation across tasks, promoting the model's ability to generalize to new tasks.

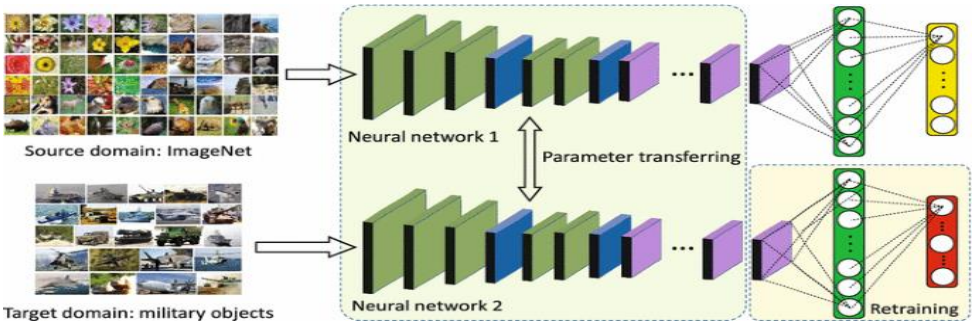


Fig.3. Transfer learning process (source: link.springer.com)

Transfer learning techniques have proven successful in various applications, including computer vision, natural language processing, and speech recognition. They have been instrumental in achieving state-of-the-art performance on tasks with limited labeled data, reducing the need for massive datasets for every specific application. As the field continues to advance, transfer learning methods are expected to play a crucial role in addressing challenges and accelerating progress in machine learning applications across diverse domains.

1.6 Overfitting

Overfitting is a common challenge in machine learning, occurring when a model learns the training data too closely, capturing noise and specific patterns that may not generalize well to new, unseen data. Several methods are employed to mitigate overfitting and enhance a model's ability to generalize:

Regularization Techniques: Regularization methods add penalties to the model's learning process, discouraging it from fitting the training data too closely. L1 and L2 regularization introduce penalties based on the absolute and squared values of model weights, respectively, while elastic net combines both.

Dropout: Commonly used in neural networks, dropout randomly deactivates neurons during training, preventing the model from relying too heavily on specific features. This regularization technique encourages the network to learn more robust representations.

Data Augmentation: Data augmentation artificially expands the training dataset by applying transformations like rotation, scaling, and flipping to existing data. This exposes the model to a broader range of variations, promoting generalization.

Cross-Validation: Cross-validation assesses a model's performance on multiple subsets of the training data. Techniques like k-fold cross-validation split the data into folds, training the model on different subsets and evaluating its performance on others to ensure consistency.

Early Stopping: Early stopping monitors the model's performance on a validation set during training and halts the process when performance starts to degrade. This prevents the model from learning noise in the training data and helps find an optimal point before overfitting occurs.

Ensemble Learning: Ensemble methods combine predictions from multiple models to improve overall performance. Techniques like bagging and boosting create diverse models, reducing overfitting by combining their strengths.

Simplifying Model Architecture: Complex models with excessive parameters are more prone to overfitting. Simplifying the architecture by reducing layers, nodes, or features helps prevent overfitting, especially with limited data.

Implementing these methods, either individually or in combination, assists in finding a balance between fitting the training data well and generalizing effectively to new data. By addressing overfitting, these techniques contribute to building more robust and reliable machine learning models.

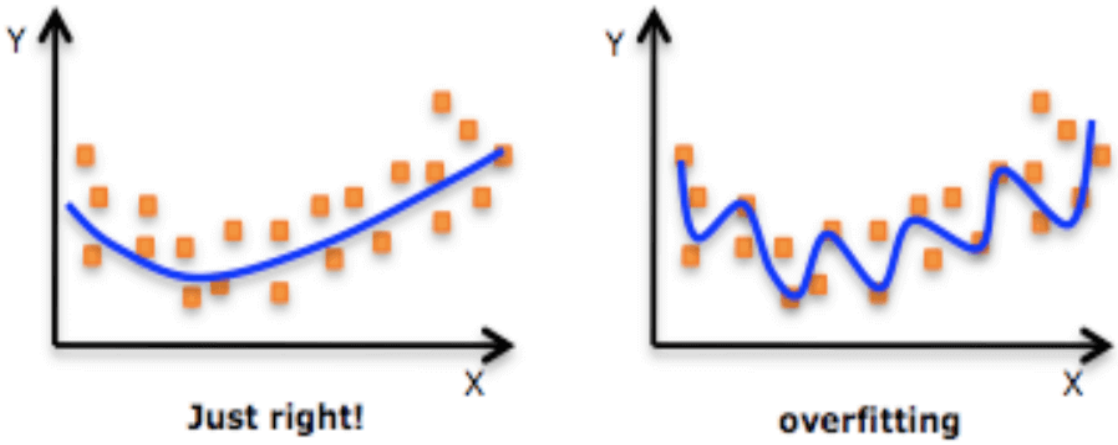


Fig.4. Overfitting Curve (source: www.datarobot.com)

1.7 Image Augmentation

Image augmentation is a technique widely used in computer vision and deep learning to artificially increase the diversity of a training dataset by applying various transformations to the original images. The goal is to expose the model to a broader range of variations, helping to improve its generalization and robustness to different conditions.

The primary goal of data augmentation is to improve the generalization ability of machine learning models by exposing them to a wider range of variations and scenarios that they might encounter during prediction. By increasing the diversity of training examples, data augmentation helps prevent overfitting and enhances the robustness of models to variations in input data.

In image classification tasks, for example, data augmentation techniques such as rotation and flipping can create new training samples that capture variations in object orientation and viewpoint. Similarly, in natural language processing tasks, techniques like adding synonyms or swapping words can generate new textual samples with slightly different semantics.

Data augmentation is particularly useful in scenarios where collecting large amounts of labeled data is expensive or impractical. By leveraging existing data more effectively, data augmentation allows machine learning models to achieve better performance with less labeled data, ultimately leading to more reliable and accurate predictions.

Furthermore, data augmentation can also be tailored to specific domains or applications. For example, in medical imaging, techniques such as elastic deformation or intensity variation can simulate realistic anatomical variations or imaging artifacts, allowing models to learn to be more robust to such factors. In autonomous driving, data augmentation can simulate different weather conditions, lighting conditions, or road surfaces, enabling models to generalize better to diverse real-world scenarios.

1.8 Ensemble Techniques

Ensembling techniques in deep learning are a powerful strategy employed to harness the collective intelligence of multiple individual models. By amalgamating the predictions of diverse models, ensembling aims to overcome the limitations of any single model and attain a more accurate and robust final prediction.

The essence of ensembling lies in the acknowledgment that different models may excel at capturing distinct aspects of the underlying data distribution or learning different representations of the input space. Leveraging this diversity, ensembling techniques seek to aggregate the varied insights and perspectives offered by these models to arrive at a more comprehensive and reliable prediction.

Moreover, ensembling acts as a form of wisdom of the crowd, where the collective judgment of multiple models can often outperform that of any individual model. This is particularly beneficial in scenarios where the data may be noisy, complex, or subject to inherent variability, as ensembling can help smooth out discrepancies and uncertainties present in individual predictions. Ensembling not only enhances the accuracy of predictions but also bolsters the robustness and generalization ability of the model. By combining predictions from multiple models trained on different subsets of the data or employing diverse architectures and hyperparameters, ensembling mitigates the risk of overfitting to specific anomalies of the training data and promotes a more robust understanding of the underlying patterns in the data.

Overall, ensembling techniques in deep learning represent a sophisticated approach to model aggregation, harnessing the strengths of diverse models to achieve superior performance and reliability. Whether through bagging, random forests, or other ensemble methods, the principle remains the same. y embracing diversity and collaboration, ensembling empowers deep learning models to unlock their full potential and excel in a wide range of real-world applications.

1.9 Novel Approach

In this research we proposed a decision-making criterion for ensembling of deep learning models for the prediction of weather events. We utilized the Weather Phenomena Dataset (WEAPD [14]) comprising 11 distinct weather classes, and adopted an ensemble approach incorporating VGG16, Inception V3, and ResNet50 pre-trained models for weather image classification. The ensemble model construction involved removing the top classification layers from the individual models, creating an ensemble model with shared input layers. Dynamic learning rate adaptation, contingent on validation loss, was employed during training to optimize convergence. Additionally, dropout regularization was integrated into the ensemble model to mitigate overfitting. The decision-making process was diversified by employing three strategies—average, maximum, and Geometric-Harmonic Ensemble (GHE) —to aggregate predictions from the individual models, providing a comprehensive evaluation of their collective performance.

The geometic harmonic mean, a fundamental mathematical concept, possesses several desirable properties that make it suitable choice for combining predictions in machine learning tasks. Firstly, it guarantees non-negativity, ensuring that the result of the ensemble remains interpretable and valid, especially when dealing with predictions or scores that must be positive values. Secondly, its scale invariance property is invaluable in situations where input values come from different scales across models. By treating each model's predictions equally regardless of their magnitude, the ensemble maintains fairness and consistency. Moreover, the geometric harmonic mean exhibits robustness to outliers, although not strictly a mathematical property, its tendency to be less affected by extreme values safeguards against the skewing of results, enhancing the reliability of the ensemble. Furthermore, its multiplicative nature preserves the multiplicative relationship between probabilities or likelihoods, making it particularly suitable for combining such values. By leveraging these properties, the geometric harmonic mean provides a robust, balanced and mathematically sound approach to combining predictions from diverse models like VGG16, RestNet50, and Inception V3, ensuring the integrity and effectiveness of the ensemble in machine learning applications.

The present work utilized the pre-trained deep neural networks with dynamic learning rate for weather image classification using WEAPD and achieved an impressive accuracy of 92.82% on holdout test images. The application of transfer learning helps facilitate the development of a resilient and potent classifier. Incorporating the Geometric-Harmonic Ensemble (GHE) technique into our methodology further enhances the accuracy and robustness of weather image classification. GHE combines the strengths of geometric mean (GM) and harmonic mean (HM) by averaging their results. The GM emphasizes consensus among models, while the HM gives more weight to smaller values, ensuring that no single model dominates the ensemble. By striking a balance between consensus and diversity, GHE effectively improves classification performance.

1.10 Contribution

In this research, a novel approach for enhancing image classification of weather images is proposed. Building upon existing ensemble decision-making methods, a cubic ensemble technique is introduced for creating a more robust classifier. Unlike traditional ensembles that combine predictions linearly, cubic ensemble involves cubing the predictions generated by state-of-the-art models and calculating the mean of these cubed predictions. This innovative approach aims to improve the accuracy and reliability of weather image classification by leveraging the collective intelligence of multiple models in a novel way. Through experimentation and evaluation, the efficacy of the cubic ensemble method is demonstrated, offering promising results for advancing the field of weather image classification.

1.11 Objectives And Aims

In this research, our primary goal is to propose a novel method for classifying various weather phenomena. To achieve this, we leverage the power of the VGG16 deep neural network architecture, a well-established model in the field of deep learning. Our focus is on categorizing data into 11 distinct weather classes, and we aim to enhance classification accuracy by exploiting the capabilities of deep learning. To optimize the performance of our model, we engage in meticulous hyperparameter tuning, adjusting the configuration settings to find the most effective combination. Additionally, we employ data augmentation techniques to artificially diversify our training dataset, enhancing the model's ability to recognize different patterns and scenarios.

In the pursuit of improved performance, we incorporate dynamic learning rate strategies. This involves adjusting the learning rate during the training process, allowing the model to adapt more effectively to the nuances of the data. Furthermore, we explore ensemble techniques, which involve combining the predictions of multiple models to create a more robust and accurate final prediction. This approach aims to surpass the performance of existing models that are commonly used in weather phenomena classification.

Our research methodology reflects a comprehensive effort to push the boundaries of weather classification accuracy, utilizing a sophisticated deep learning architecture, optimizing hyperparameters, employing data augmentation, incorporating dynamic learning rates, and exploring ensemble techniques. Through these strategies, we aspire to achieve superior results compared to other established models in the domain, contributing to advancements in the field of weather phenomena classification.

CHAPTER 2

Literature Survey

Weather image classification plays an important role in enhancing the understanding of meteorological conditions, contributing to numerous applications that impact various sectors. By using advanced machine learning techniques, such classification involves the categorization of diverse weather phenomena, such as rain, fog, lightning, and snow, based on visual data. Beyond meteorology, the application extends to precision agriculture, transportation management, disaster preparedness, and environmental monitoring, where real-time recognition of weather patterns facilitates informed decision-making and mitigates the impact of unpredictable weather conditions.

Many Researches have been conducted in past for classification of weather events.

2.1 Related Research and Previous Works

In 2017, Sebastian Stabinger and Antonio Rodríguez-Sánchez [1] explored the burgeoning realm of image classification methodologies. They noted the ascendancy of Convolutional Neural Networks (CNNs) as premier tools in this domain, often surpassing human performance on various datasets. While most datasets at the time were structured around concrete object classes, the researchers proposed a groundbreaking alternative: an image classification dataset featuring abstract classes. Despite these classes being seemingly straightforward for human classification, Stabinger and Rodríguez-Sánchez discovered that CNNs faced notable challenges when confronted with variations within the dataset. Through rigorous analysis, they evaluated the classification capabilities of prevalent CNN architectures on this innovative dataset, pinpointing avenues for further research and exploration.

In 2020, Saiharsha B., Abel Lesle A., B. Diwakar, Karthika R., and Ganesan M. [2] conducted a study comparing the performance of VGG-Net and conventional CNN architectures using the Fashion MNIST dataset. VGG-Net, a relatively recent and notably accurate CNN architecture, sparked interest in understanding its effectiveness across different scenarios. The researchers employed a 3-layer CNN architecture for their investigation, achieving impressive training accuracy of 98.92% with minimal training loss of 0.02. Through their experiments, they attained a maximum test accuracy of 90.77% in classifying 10,000 images

representing 10 distinct types within the Fashion MNIST dataset. Their findings shed light on the comparative performance of VGG-Net and traditional CNN architectures, providing valuable insights into their efficacy for image classification tasks.

In 2022, Abebaw Alem [3] from the IT Department at Debre Tabor University in Ethiopia and Shailender Kumar. User from the Department of Computer Science and Engineering at Delhi Technological University in India delved into the realm of deep learning-based land cover and land use (LCLU) classification systems. Recognizing the pivotal role of such systems in remote sensing communities, the researchers acknowledged the multifaceted nature of remote sensing images, which pose challenges in analysis and interpretation due to varying factors such as sensor technology, environmental conditions, and seasonal changes. The significance of effectively analyzing these images lies in their application in environmental monitoring, agricultural decision-making, and urban planning, all of which could greatly benefit from the support of deep learning systems.

In 2021, W A Ezat, M M Dessouky, and N A Ismail [4] explored the intricacies of image classification within the domain of computer vision, recognizing its complexity and the resource-intensive nature of training deep learning models. Leveraging the convolutional neural network (CNN), a widely employed architecture in deep learning, the researchers investigated the performance of a CNN model pre-trained on ImageNet for classifying images from the PASCAL VOC 2007 dataset. Employing a transfer learning approach, they aimed to enhance the CNN model's performance while minimizing computational time and resource utilization. Through their study, they scrutinized the behavior and performance of the deep learning CNN model, comparing its results with those obtained from alternative methods such as the Super-vector coding of local image descriptors method, SVM method, and Region Ranking SVM method, all tested on the PASCAL VOC 2007 dataset.

In 2017, Luis Perez and Jason Wang [5] delved into the realm of data augmentation in image classification, aiming to explore and compare various solutions to this pervasive problem. Building upon previous research that showcased the efficacy of simple techniques like cropping, rotating, and flipping input images, the study artificially limited access to data to a small subset of the ImageNet dataset.Each data augmentation technique was meticulously evaluated and compared. Notably, traditional transformations emerged as one of the more successful strategies. However, the researchers also ventured into experimenting with Generative Adversarial Networks (GANs) to generate images of different styles, expanding the scope of augmentation possibilities.

In 2015, Asma ElAdel, Ridha Ejbali, Mourad Zaied, and Chokri Ben Amar [6] introduced a novel scheme for image classification, revolutionizing the landscape of Convolutional Neural Networks (CNNs). The proposed scheme presents a unique shallow architecture of CNN that achieves deep learning capabilities. In this scheme, connection weights between the input layer and the hidden layer are computed for each image based on Multiresolution Analysis (MRA) at varying levels of abstraction. Subsequently, employing the Adaboost algorithm, the best features representing each image class are selected along with their corresponding weights. These weights are then utilized as the connection weights between the hidden layer and the output layer.

In 2015 Zheng Zhang and Huadong Ma [7] address the challenging task of multi-class weather classification from single images, crucial in numerous outdoor computer vision applications. They note the scarcity of research tackling this difficulty and the limitations of existing methods, often confined to fixed scenes. To overcome these hurdles, they propose a novel approach based on multiple weather features and multiple kernel learning, adaptable to any scenario. Their method involves extracting and properly processing diverse weather features, which are then combined into high-dimensional vectors. Through multiple kernel learning, an adaptive classifier is trained to categorize weather conditions effectively achieving an accuracy of 59%. The researchers compile the MWI (Multiclass Weather Image) dataset comprising 20,000 outdoor images for evaluation. Experimental results demonstrate the efficiency of their approach in recognizing weather patterns across the MWI dataset, marking a significant advancement in outdoor weather classification within computer vision.

In their 2016 paper, Ziqi Zhu, Li Zhuo, Panling Qu, Kailong Zhou, and Jing Zhang [8] address the critical issue of automatic recognition of extreme weather, a task vital for ensuring driving safety and protecting lives and property. Traditional methods often rely on multiple sensors with limited accuracy and artificial participation. To overcome these challenges, the researchers propose a novel approach based on computer vision techniques utilizing convolutional neural networks (CNNs). By leveraging big data, they generate features and recognition models for extreme weather. Additionally, they introduce the "Weather Dataset," a large-scale dataset containing 16,635 extreme weather images divided into four classes: sunny, rainstorm, blizzard, and fog. The recognition model is developed through pre-training on the ILSVRC-2012 Dataset using Google Net, followed by fine-tuning on Weather Dataset. Experimental results demonstrate the method's high performance, achieving a recognition accuracy rate of 94.5% and meeting the requirements of real-world applications. This research marks a significant advancement in automatic extreme weather recognition, offering promising implications for highway traffic condition warnings, automobile assistance systems, and climate analysis.

In their 2018 paper, Zhe Li, Yi Jin, Yidong Li, Zhiping Lin, and Shan Wang [9] tackle the challenge of imbalanced label distributions in weather classification, a fundamental task with wide-ranging computer vision applications. Recognizing the rarity of certain weather types, such as rain and snow, compared to more common categories like sunny and hazy days, the researchers propose a data augmentation approach using generative adversarial networks (GANs). Their method aims to enhance the diversity of image data, thereby improving classification accuracy. Specifically, the researchers design a framework utilizing a deep convolutional generative adversarial network (DCGAN) as a generator to produce images and balance the imbalanced dataset. They employ a Convolutional Neural Network (CNN) model as a classifier to evaluate the classification results. To assess the performance of DCGAN, the researchers introduce an evaluation method using three benchmark datasets for comparative experiments. The empirical findings of their study demonstrate the effectiveness of DCGAN in generating high-quality weather images and improving classification accuracy through data augmentation. This research contributes to advancing weather classification methodologies, offering promising implications for various computer vision applications.

In their 2021 study, Mesut Toğaçar, Burhan Ergen, and Zafer Cömert [10] explore the significance of weather information transmission on people's daily lives and agricultural activities. Recognizing the potential of weather images for making predictions with computer systems, the researchers highlight the growing importance of integrating artificial intelligence (AI) technologies into such systems. Their dataset comprises four weather classes: cloudy, rain, shine, and sunrise. To enhance classification accuracy, the researchers employ a combination of the Google Net and VGG-16 models alongside a spiking neural network (SNN). They extract features from Google Net and VGG-16 models and combine them as input for the SNN. Through this collaborative approach, the SNN contributes significantly to the success of weather classification. The experimental results demonstrate impressive classification accuracy rates for the cloudy, rain, shine, and sunrise classes, reaching 98.48%, 97.58%, 97%, and 98.48%, respectively, when used in conjunction with the SNN. This study underscores the effectiveness of integrating SNNs with deep learning models for achieving successful weather classification outcomes, showcasing the potential of AI technologies in enhancing weather prediction systems.

In contrast, a more recent study in 2022 [11] shifted the focus to the agricultural sector and addressed the challenges posed by diverse weather conditions. The authors utilized the same 11-class WEAPD dataset employed in the current article. They explored various pre-trained models, including Inception, ResNetV2, Xception, MobileNet, and DenseNet201, achieving accuracies of 73%, 76%, 81%, and 83%, respectively. This research highlights the importance of considering a broader range of weather classes, reflecting the nuanced conditions experienced in real-world agricultural settings.

Additionally, another significant contribution [12] emphasized the critical implications of weather phenomenon recognition. The proposed solution, MeteCNN, introduced a deep CNN architecture with a squeeze-and-excitation module designed to enhance feature importance by adaptively recalibrating the network's responses. This approach showcases the evolving sophistication in the field, incorporating advanced architectures and techniques to address the intricacies of weather recognition and classification.

They also introduced the said WEAPD database comprising 6,877 images representing 11 diverse weather phenomena. MeteCNN exhibited a classification accuracy of approximately 92.6%. In 2023, the weather-image classification model is contextualized by the authors through the integration of a Vision Transformer (VIT) [13]. This experiment was carried out on the WEAPD and achieved an accuracy of 87.69%.

In their 2020 study, Zabit Hameed et al. [14] address the critical issue of breast cancer diagnosis, recognizing its significance as a major public health concern worldwide. With early detection being crucial for increasing survival rates, the authors emphasize the importance of accurate histopathological analysis. However, traditional methods are labor-intensive and prone to disagreement among pathologists. To overcome these challenges, the researchers propose an ensemble deep learning approach for the definitive classification of non-carcinoma and carcinoma breast cancer histopathology images using their collected dataset. The study involves training four different models based on pre-trained VGG16 and VGG19 architectures. Through 5-fold cross-validation operations on each model, including fully-trained and fine-tuned variants of VGG16 and VGG19, the authors evaluate their performance. Subsequently, an ensemble strategy is employed, averaging the predicted probabilities from the individual models. Notably, the ensemble of fine-tuned VGG16 and VGG19 models demonstrates competitive classification performance, particularly for carcinoma images. The results indicate that the ensemble approach achieves a sensitivity of 97.73% for carcinoma class and an overall accuracy of 95.29%. Additionally, it yields an F1 score of 95.29%.

These findings highlight the effectiveness of the proposed deep learning approach for automatic classification of complex histopathology images of breast cancer, particularly for carcinoma images. This research contributes to advancing the field of breast cancer diagnosis, offering promising implications for improving diagnostic processes and patient outcomes.

In their 2022 study, Qian Pan et al. [15] address the significant impact of rust disease on wheat growth and yield, particularly the challenges associated with distinguishing between stem rust and leaf rust, which can be time-consuming through manual detection methods. To address this issue, the authors propose a novel method for identifying wheat rust based on ensemble learning (WR-EL). The WR-EL method integrates multiple convolutional neural network (CNN) models, including VGG, ResNet 101, ResNet 152, DenseNet 169, and DenseNet 201, utilizing bagging, snapshot ensembling, and the stochastic gradient descent with

warm restarts (SGDR) algorithm. By combining these models, the WR-EL method aims to improve identification accuracy. Comparative analysis with five individual CNN models demonstrates that the WR-EL method yields significant improvements in identification accuracy, with increases of 32%, 19%, 15%, 11%, and 8% across various metrics. Additionally, the authors introduce the SGDR-S algorithm, which further enhances the f1 scores for healthy wheat, stem rust wheat, and leaf rust wheat compared to the SGDR algorithm. Overall, the proposed method offers a more accurate means of identifying wheat rust disease, potentially serving as a timely prevention and control measure. By mitigating economic losses and improving wheat yield and quality, this approach has promising implications for agricultural practices and food security.

In recent years, Dan Xue; Xiaomin Zhou; Chen Li; Yudong Yao; Md Mamunur Rahaman; Jinghua Zhang [16] in 2020 there has been growing interest in leveraging Transfer Learning (TL) and Ensemble Learning (EL) techniques for cervical histopathology image analysis. However, there remains a lack of research focusing on differentiating the stages of cervical histopathological images. To address this gap, this article presents an innovative Ensembled Transfer Learning (ETL) framework designed to classify well, moderately, and poorly differentiated cervical histopathological images. The framework begins with the development of TL structures based on Inception-V3, Xception, VGG-16, and ResNet-50 models. Subsequently, a weighted voting-based EL strategy is introduced to enhance classification performance. The proposed algorithm is evaluated using a dataset comprising 307 images stained with three immunohistochemistry methods (AQP, HIF, and VEGF). Experimental results demonstrate the effectiveness of the ETL framework, achieving the highest overall accuracy rates of 97.03% and 98.61% on AQP staining images and poorly differentiated VEGF staining images, respectively. Additionally, an experiment on the Herlev dataset for classifying benign cells from malignant ones yields an impressive overall accuracy of 98.37%. This research significantly contributes to advancing cervical histopathology image analysis by providing a robust framework capable of accurately classifying images based on differentiation stages. The utilization of TL and EL techniques underscores the potential of the proposed ETL framework in improving diagnostic accuracy for cervical cancer detection.

Remote sensing technology has been instrumental in providing extensive agricultural data, particularly in monitoring paddy growing ecosystems. However, challenges persist in data fusion techniques, impacting the accuracy of image classification in paddy rice fields. In this study, a novel approach integrating various learning concepts, bolstered by a probability progress Fuzzy Dempster-Shafer (FDS) analysis, is proposed to enhance models and analyze different types of image data. The study [17] Led by Tsu Chiang Lei, Shiuan Wan, Shih-Chieh Wu, and Hsin-Ping Wang, this research underscores the efficacy of the proposed approach in improving image classification accuracy in paddy rice fields, offering promising implications for agricultural monitoring and management leverages FDS to generate probability models for classification systems, integrating Logistic Regression (LR), Support Vector Machine (SVM), and Neural Network (NN) approaches within the developed FDS system. Satellite imagery and aerial photos serve as the primary analysis materials, with the classification accuracy significantly improved to 97.27% and a kappa value of 0.93 achieved. Moreover, the study reports overall classification accuracies ranging between 85% and 90%

for multi-period mid-scale satellite images of paddy fields, and between 91% and 95% for multi-spectral numerical aerial photos. Notably, the FDS implementation further enhances the accuracy of these classification results.

In the study [18] , the authors address the computational complexity issues of existing graph-based, semi-supervised hyperspectral image (HSI) classification models by proposing two novel approaches: fast anchor graph regularization (FAGR) and semi-supervised anchor graph ensemble (SAGE). Firstly, they introduce the FAGR model, which utilizes a simplified anchor-based graph construction method and a new adjacency matrix among anchors to significantly reduce computational complexity while maintaining high classification performance. Furthermore, we propose the SAGE model, which is an ensemble realization of multiple FAGR classifiers. Each component of SAGE operates on a randomly selected subset of features. A meta-classifier aggregates the outputs of these component classifiers to produce an ensemble classification result.

Extensive experimentation using three real-world HSI datasets demonstrates the effectiveness of FAGR and SAGE. Particularly, SAGE achieves a remarkable 95.78% classification accuracy on the Indian Pines dataset, even with limited labeled samples. Compared to existing models, SAGE offers shorter execution times and superior classification accuracy. Led by Ziping He and colleagues from the School of Electronic and Information Engineering at Hebei University of Technology, this research contributes valuable insights into addressing computational complexity challenges in HSI classification, offering practical solutions for improved classification accuracy and efficiency.

Recently in 2024 [19], Oral squamous cell carcinoma (OSCC) poses a significant health concern globally, with its high mortality rate largely attributed to poor prognosis. Traditional diagnosis methods relying on pathologists' visual inspection of tissue biopsy sections are prone to human errors and subjectivity. Leveraging advancements in artificial intelligence, deep learning models have emerged as valuable tools in medical diagnosis, offering enhanced accuracy and efficiency in image classification tasks. In this study, we employ a customized deep learning model to aid pathologists in OSCC detection from histopathological images. Our approach involves analyzing a comprehensive dataset comprising 696 histopathological oral images, with 80% allocated to the training set, 10% for validation, and the remaining 10% for testing purposes. We utilize 2D empirical wavelet transform to extract features from the images, followed by an ensemble of two pretrained models, ResNet50 and DenseNet201, for classification into normal and OSCC classes. The efficacy of our model is evaluated based on accuracy, sensitivity, and specificity metrics. Our simulation results demonstrate that the proposed model achieves a classification accuracy of 92%, indicating its effectiveness in binary classification of oral histopathological images. This method holds promise for assisting pathologists in OSCC diagnosis, potentially reducing diagnostic errors and workload. Led by Bhaswati Singha Deo, Mayukha Pal, Prasanta K. Panigrahi, and Asima Pradhan, this research contributes to advancing OSCC detection methodologies, offering a reliable and efficient solution for improved patient outcomes.

2.2 Comparative Analysis of Previous Research

Various methods for weather image classification are compared based on their performance metrics using the WEAPD [20] dataset. The dataset consists of weather images, and the objective is to accurately classify them into different weather classes. Several state-of-the-art methods have been employed, and their performance is evaluated in terms of accuracy, recall, precision, and F1 score.

These results provide a comprehensive overview of the performance of different methods in weather image classification, highlighting the strengths and effectiveness of each approach. The metrics indicate the models' abilities to correctly classify images, with VIT-DA emerging as the top-performing method in terms of accuracy on the WEAPD dataset.

CHAPTER 3

Proposed Methodology

In this research, the main objective is to present a technique for weather phenomena classification, utilizing the pre-trained VGG16 [21] deep neural network, data augmentation and ensemble techniques. Targeting 11 distinct weather classes, we harness the capabilities of deep learning to improve classification accuracy for the Weather Phenomena Dataset (WEAPD) [20]. Through thorough hyperparameter tuning along with data augmentation and dynamic learning rate, the optimized ensemble model is aimed to generalize better and achieve higher accuracy results.

3.1 VGG-16 Model

The VGG16 model [11], designed by the Visual Geometry Group, features 16 layers, including 13 convolutional and 3 fully connected layers. Utilizing 3x3 convolutional filters with ReLU activations, it excels in capturing hierarchical features. Max-Pooling layers strategically reduce spatial dimensions, preserving essential features. Through transfer learning on datasets like ImageNet, VGG16 is adept at image classification tasks. Its hierarchical feature extraction, systematic pooling, and parameter efficiency make it a robust choice for various computer vision tasks.

The existing VGG-16 model when trained from scratch on the ImageNet dataset could exhibit a tendency to overfit on the weather image dataset, potentially leading to poor performance on unseen test data. To mitigate this and also to reduce computation burden for training, the VGG16 model [11] pre trained on ImageNet dataset is used as a backbone architecture for our task. Initially, the pre-trained model is fine-tuned with a lower learning rate thereby preserving valuable features obtained from ImageNet dataset while making nuanced adjustments. This cautious approach prevents drastic modifications in the architecture which could otherwise compromise pre-existing knowledge, enabling profound fine-tuning to learn intricate weather specific patterns. The balance between conservative learning and layer unfreezing ensures optimal refinement of parameters, allowing the model to generalize better and excel in classifying complex weather patterns, enhancing efficacy in weather image classification tasks.

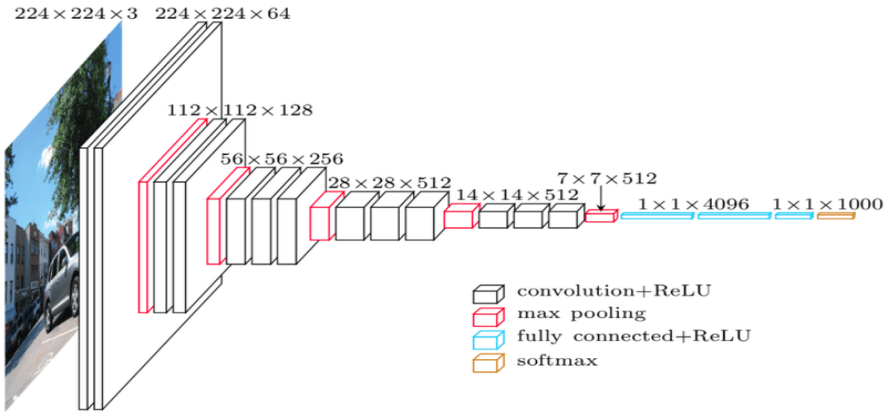


Fig.5. VGG -16 Architecture (source: datagen.tech)

To address the challenge of overfitting and enhancing generalization ability, dropout and regularization techniques can been used. Introduction of dropout, in turn, increases the bias and will also eliminate the redundant connections resulting into reduced network complexity. Regularization will facilitate better convergence. Furthermore, the model was trained with a low learning rate, dynamically adjusted based on the model's performance on validation data. On top of it, an optimal batch size was employed to improve efficacy in handling unseen weather data. Two fully connected dense layers (512 and 256 neurons) and L2 regularization were introduced in the model. A fixed learning rate may not be optimal for the entire training duration, as it might lead to convergence issues or slow down the learning process. Dynamic learning rate strategies, such as learning rate schedules, adaptive methods like Adam can dynamically modify the learning rate based on certain conditions or performance metrics. This allows the model to navigate the optimization landscape more effectively, speeding up convergence during the initial stages and fine-tuning as the optimization progresses.

3.2 VGG-16 With Data Augmentation

If the dataset used is imbalanced, with a disproportionately low number of instances representing specific weather conditions or scenarios (e.g., Rainbow, Lightning) among 11 different weather phenomena as shown in Table 1. A model trained on a limited set of instances for a specific class as compared to other classes may lead to an inferior training thereby struggling to generalize to unseen weather conditions. To address the issue of data imbalance and the limited number of weather instances, expanding training set through augmentation might be beneficial. Several augmentation techniques exist in literature. By applying several transformations like rotations, flips, zooms, and adjustments in brightness or contrast, augmentation method artificially expands the dataset, reducing overfitting and improving generalization.

Image augmentation involves applying a set of transformations to the existing images in a dataset, creating new variations while preserving the essential characteristics of the original data. Common transformations include:

Rotation: Rotating images by a certain angle to simulate different perspectives and orientations.

The rotation operation can be expressed as follows:

$$(x', y') = (cos(\theta)sin(\theta) - sin(\theta)cos(\theta))(x, y) \quad \text{..... (i)}$$

Where:
cos(θ) represents the cosine of the rotation angle θ.
sin(θ) represents the sine of the rotation angle θ.

This formula computes the new coordinates.
(x', y') of a pixel after applying a rotation of angle.
θ to the original coordinates (x, y).

Flipping: Horizontally or vertically flipping images to introduce mirror effects and diversify spatial orientations. for flipping the image both horizontally and vertically, we can combine the below formulas for horizontal and vertical flipping:

$$x' = W - x - 1 \quad \text{.....(ii)}$$
$$y' = H - y - 1 \quad \text{.....(iii)}$$

(x, y) as the coordinates of a pixel in the original image.
(x', y') as the coordinates of the pixel after flipping.
W as the width of the image.
H as the height of the image.

Zooming: Scaling images in or out to mimic variations in the object's size or distance from the camera.

We can use the following formula for performing zooming operation in an image.

$$x' = Sx \times x \quad \text{.....(iv)}$$
$$y' = Sy \times y \quad \text{.....(v)}$$

Where:

x, y) as the coordinates of a pixel in the original image.
(x', y') as the coordinates of the pixel after zooming.
Sx as the scaling factor along the x-axis.
Sy as the scaling factor along the y-axis.

Translation: Shifting images horizontally or vertically to simulate changes in the object's position within the frame.

The translation operation can be represented as follows:

$$x' = x + Tx \quad \text{..... (vi)}$$
$$y' = y + Ty \quad \text{..... (vii)}$$

Where:

(x, y) as the coordinates of a pixel in the original image.
(x', y') as the coordinates of the pixel after translation.
Tx as the translation amount along the x-axis.
Ty as the translation amount along the y-axis.

These formulas compute the new coordinates (x', y') of a pixel after applying translation with amounts Tx along the x-axis and Ty along the y-axis to the original coordinates (x, y).

Brightness and Contrast Adjustment: Altering the brightness and contrast levels of images to account for varying lighting conditions.

Brightness and contrast adjustment are common techniques in image augmentation used to enhance or modify the visual appearance of images. Let's denote:

I (x, y) as the intensity of the pixel at coordinates in the original image.
I' (x, y) as the intensity of the pixel at coordinates in the adjusted image.
α as the brightness factor.
β as the contrast factor.
The brightness and contrast adjustment operations can be represented as follows:

(a) Brightness Adjustment:

$$I'(x, y) = \alpha \times I(x, y) \quad \text{..... (viii)}$$

In this operation, each pixel intensity is multiplied by the brightness factor $\alpha\alpha$. Increasing $\alpha\alpha$ makes the image brighter, while decreasing it makes the image darker.

(b) Contrast Adjustment:

$$I'(x, y) = \beta \times (I(x, y) - \mu) + \mu \qquad \ldots\ldots. (ix)$$

Where μ is the mean intensity of the image. In this operation, the mean intensity of the image is subtracted from each pixel intensity, then the result is multiplied by the contrast factor β and added back to the mean intensity. Increasing β increases the contrast of the image, while decreasing it decreases the contrast.

These adjustments allow for fine-tuning the appearance of images, making them brighter, darker, or adjusting the range of intensities to emphasize or de-emphasize details in the image.

Noise Injection: Adding random noise to images to simulate the presence of noise in real-world scenarios. For example:
Gaussian Noise:

$$I'(x, y) = I(x, y) + noise(0, \sigma^2) \qquad \ldots\ldots\ldots (x)$$

Where:
$I'(x, y)$ is the pixel intensity in the noisy image, I (x, y) is the original pixel intensity, and noise $(0, \sigma^2)$ represents random noise drawn from a Gaussian distribution with mean 0 and standard deviation σ.

Color Jittering: Introducing small random changes to the color channels of images to account for variations in lighting and color conditions.

By applying these transformations, image augmentation aids in creating a more diverse training dataset. This is particularly beneficial when dealing with limited labeled data, as it helps prevent overfitting by providing the model with a richer set of examples to learn from as shown in Fig 5.



(a) (b)

Fig.6. (a) Original Image of dew Class (b) Augmented images of dew class

By applying several transformations like rotations, flips, zooms, and adjustments in brightness or contrast, augmentation method artificially expands the dataset, reducing overfitting and improving generalization.



Fig. 7. VGG16 with data augmentation flow chart.

These augmented images are used to fit into VGG-16 model for better prediction of weather phenomenon classes.

3.3 Ensemble Techniques

Ensemble learning, a powerful technique in machine learning, involves combining the predictions of multiple individual models to produce a final prediction that often outperforms any single model alone. Just as diverse perspectives can lead to better decision-making in human groups, ensemble methods leverage the collective intelligence of various models to enhance predictive accuracy, robustness, and generalization.

The effectiveness of VGG16, Inception V3, and ResNet50 in image classification stems from their unique architectural features. VGG16 stands out for its simple yet deep structure, which aids in capturing intricate visual details through multiple convolutional layers. Inception V3 introduces innovation through its inception modules, employing filters of different sizes to efficiently capture features across various scales, thereby handling a wide range of visual complexities while keeping computational costs in check. ResNet50 takes a step further with its use of residual connections, which mitigate the vanishing gradient problem and enable the training of exceptionally deep networks, allowing for the exploration of complex feature hierarchies.

This study introduces a decision-making framework that combines predictions from different models for weather image classification. To conduct the research, we utilized the Weather Phenomena Dataset (WEAPD), which consists of 11 distinct weather classes. We incorporated three pre-trained models—VGG16, Inception V3, and ResNet50—and employed an ensemble approach to create a unified model. This involved removing the top classification layers from each model and sharing input layers to construct the ensemble model. During training, we used dynamic learning rate adjustment based on validation loss to enhance convergence. We also implemented dropout regularization within the ensemble model to prevent overfitting. To diversify the decision-making process, we employed three aggregation strategies—average, maximum, and Geometric-Harmonic Ensemble (GHE)—to combine predictions from the individual mode ls. These strategies provide a comprehensive evaluation of the collective performance of the models.

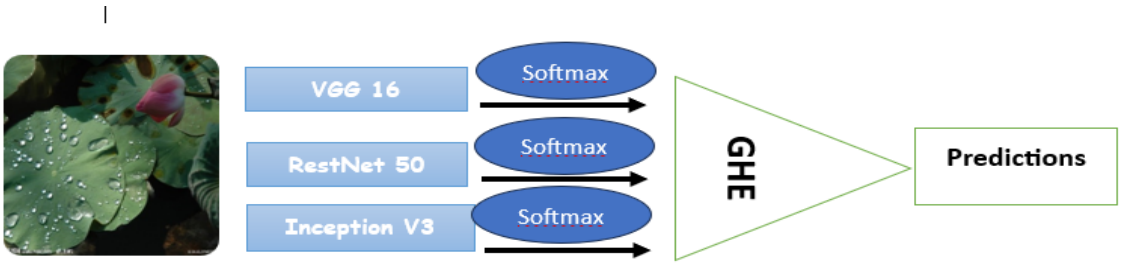


Fig. 8. Ensemble Learning flow chart.

Formula for calculating Geometric mean of prediction is:

$$GM = \sqrt[3]{p_{vgg} \times p_{res} \times p_{inc}}$$

Formula for calculating Harmonic mean of prediction is:

$$HM = \frac{3}{\frac{1}{p_{vgg}} + \frac{1}{p_{res}} + \frac{1}{p_{inc}}}$$

The Geometric-Harmonic Ensemble is finally calculated as:

$$GHE = \frac{GM + HM}{2}$$

Where:
p_{vgg} is the prediction of vgg16 trained model.
p_{res} is the prediction of restnet 50 trained model.
p_{inc} is the prediction of inception V3 trained model.

Combining these distinct architectures allows us to leverage their complementary strengths in creating a more robust image classification system. VGG16's depth aids in detailed feature extraction, while Inception V3's diverse filter sizes enhance the representation of features across different scales. Additionally, ResNet50's residual connections ensure stable training of deep networks, facilitating the extraction of intricate features crucial for discerning subtle patterns in weather images. This fusion of architectures enhances our capability to detect subtle variations and patterns in weather phenomena, leading to a more reliable and accurate classification system for meteorological applications.

CHAPTER 4

Dataset and Experimental Setup

4.1 Data Set Used

In this study, we utilized the WEAPD dataset, as referenced by [10], which consists of a total of 6,877 images distributed in to 11 weather classes as shown in Table 2. To ensure a robust evaluation of our proposed methods, we carefully partitioned the dataset into distinct training, validation, and testing sets, following an 8:1:1 ratio. This division ensures that each image is exclusively assigned to one of the three subsets, avoiding any overlap between them. The training set is used to teach the model, the validation set aids in fine-tuning and optimizing hyperparameters, and the testing set serves as an independent dataset for evaluating the final performance of our model. This meticulous dataset organization is crucial for maintaining the integrity of the evaluation process and ensuring the reliability of our findings.

Table.1. Distribution of images among classes in the WEAPD Dataset.

Types of Classes	Number of Images
Hail	592
Rainbow	238
Snow	621
Rain	527
Lightning	378
Dew	700
Sandstorm	692
Frost	475
Fog/smog	855
Rime	1160
Glaze	639
Total Images	6877

The WEAPD [10] dataset is imbalanced, with a disproportionately low number of instances representing specific weather conditions or scenarios (e.g., Rainbow, Lightning) among 11 different weather phenomena (Table 1). A model trained on a limited set of instances for a specific class as compared to other classes may lead to an inferior training thereby struggling to generalize to unseen weather conditions. To address the issue of data imbalance and the limited number of weather instances, expanding training set through augmentation might be beneficial. Several

augmentation techniques exist in literature.

4.2 Data Set Preprocessing

In this WEAPD Dataset, every input within our dataset undergoes a resizing process to achieve a uniform dimension of 224x224 pixels. This standardization is essential for neural networks as it ensures that each input image shares the same size, facilitating consistency in the model's interpretation of visual features. Furthermore, the pixel values of these resized images are normalized to fall within the [0, 1] range. Normalization is a common practice in preprocessing for machine learning models, enhancing convergence during training. By resizing and normalizing the images, we establish a standardized and manageable input format for our neural network, enabling it to effectively learn and extract meaningful patterns from the data.

4.3 Experimental Setup

Dataset splitting is vital in machine learning, with 80% for training ,10% for validation and 10 % for testing, maintaining class balance through randoms sampling. Training involves backpropagation and gradient descent, updating weights to minimize the categorical cross-entropy loss function. This iterative

process enables the model to learn patterns and improve predictive capabilities. The carefully curated split and optimization techniques ensure the model is evaluated realistically, gauging its ability to generalize to new, unseen data points.

Table.2. Experimental setup and Hyperparameters Used.

Parameters	Values
Learning Rate	0.00001
Batch Size	32
Max Epochs	100
Optimizer	Adam
Loss Function	Categorical Cross-entropy
Kernel_regularizer	0.001

Dropout, regularization, and optimal batch sizes collectively combat overfitting in machine learning models as shown in Table 2. Dropout introduces randomness by deactivating neurons, preventing reliance on specific features. Regularization penalizes excessive weight magnitudes, encouraging simplicity. Adjusting batch size introduces noise in weight updates, promoting generalization. These strategies synergize to enhance a model's adaptability to diverse data and improve its performance on unseen datasets.

Feature maps, especially those extracted from deep convolutional neural networks (CNNs) like VGG16, serve as powerful representations of hierarchical visual features learned by the model during training. These maps encapsulate information about various patterns, textures, and structures present in the input images. They play a crucial role in tasks such as object detection, image segmentation, and feature visualization. By examining feature maps at different layers, many usefull insights can be drawn about how the network processes and interprets visual information, aiding in model interpretation and debugging. Additionally, feature maps can be utilized for transfer learning, where the representations learned by pre-trained models can be leveraged for tasks with limited labeled data, thus enhancing generalization and efficiency in model training. Overall, feature maps serve as interpretable and transferable representations of visual information, facilitating a wide range of computer vision tasks.

Detailed analysis of the feature maps reveals that initial layers in the convolutional neural network (CNN) adeptly capture texture and detailed weather profiles, while deeper layers prioritize complex, high-level features at the expense of some profile details. This parallels prior research, indicating a progressive learning process across CNN layers, where shallow layers emphasize specifics and deeper layers discern intricate semantics in weather images. The study shows the CNN's ability to extract image information at varying hierarchical depths during the training process.

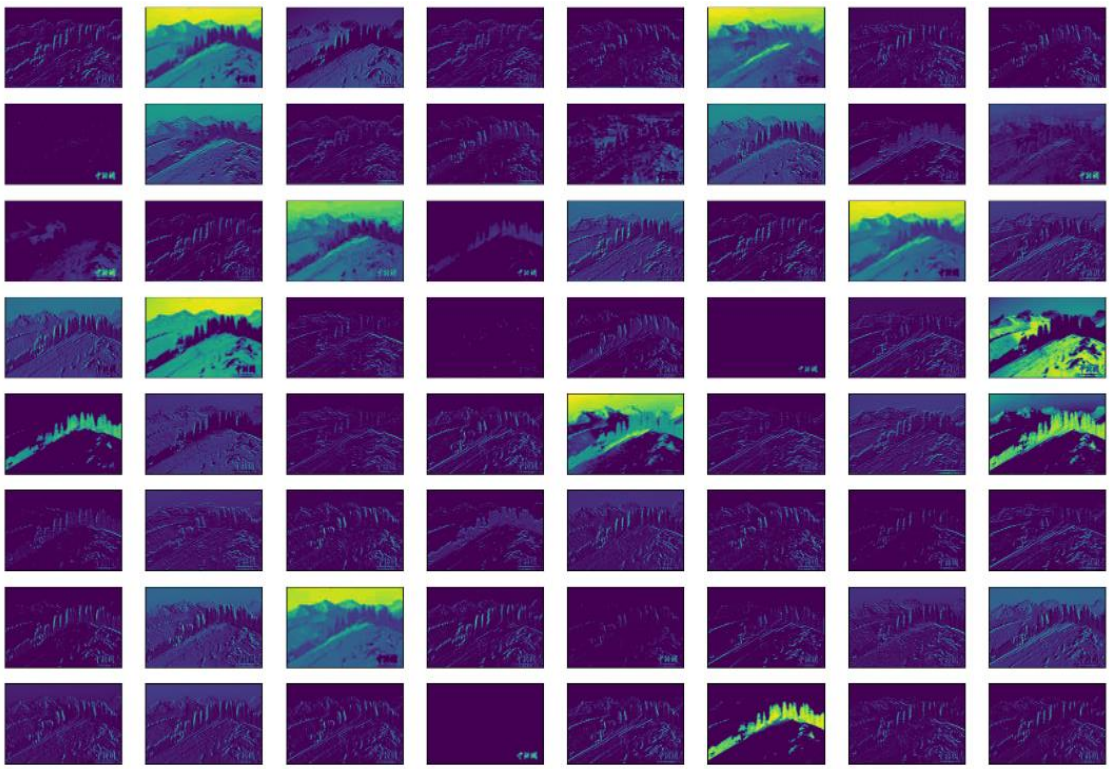


Fig. 9. (a) Feature Map of 1st Convolutional layer of VGG16 trained model.

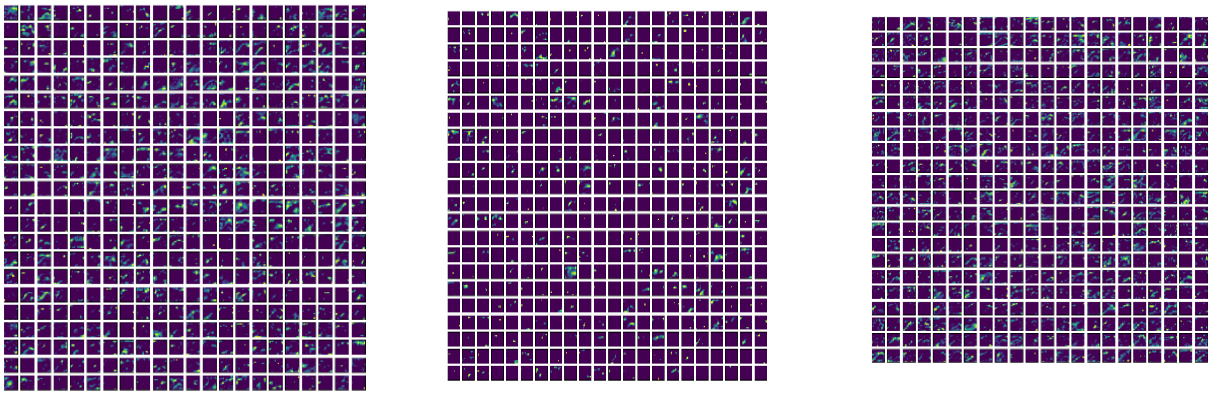


Fig. 9. (b) Feature Map of last three Convolutional layers of VGG16 trained model

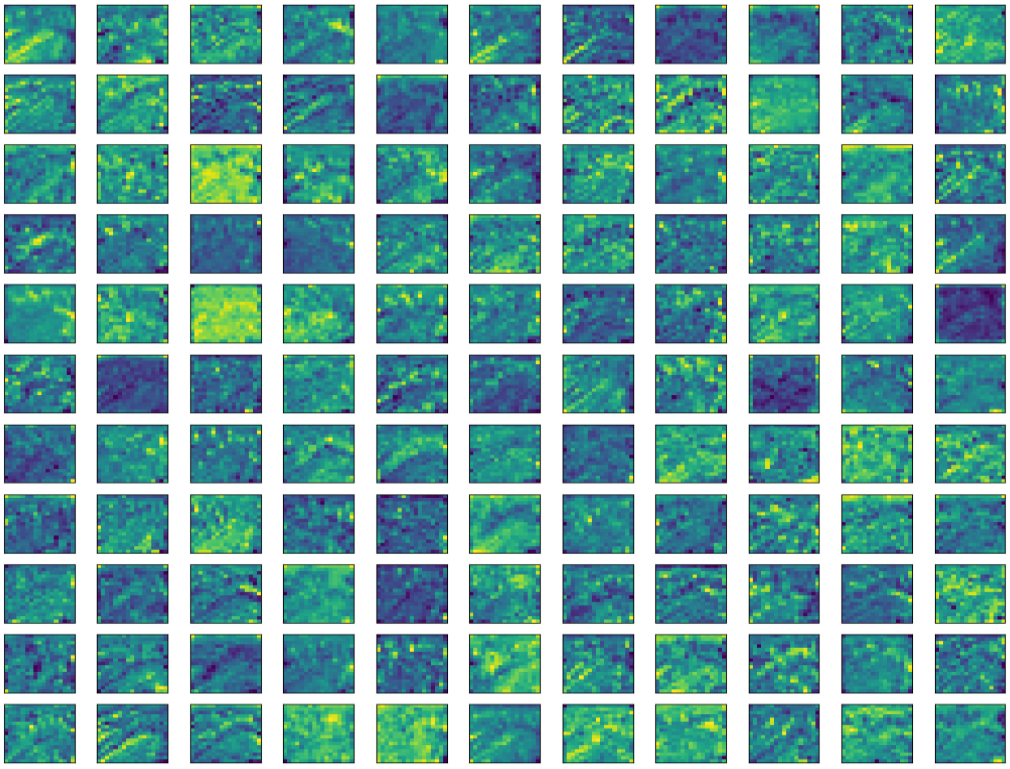


Fig. 10 (a) Feature Map of 1st Convolutional layer of Inception V3 trained model

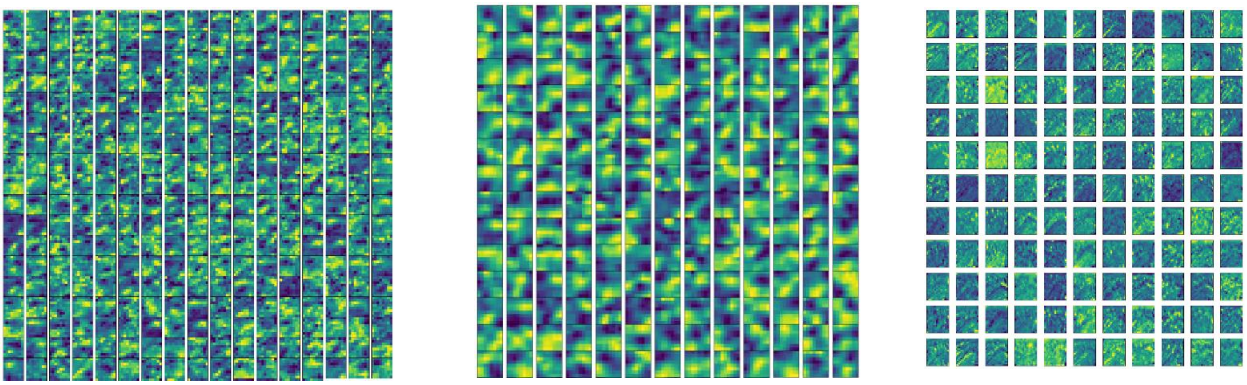


Fig. 10. (b) Feature Map of last three Convolutional layers of Inception V3 trained model.

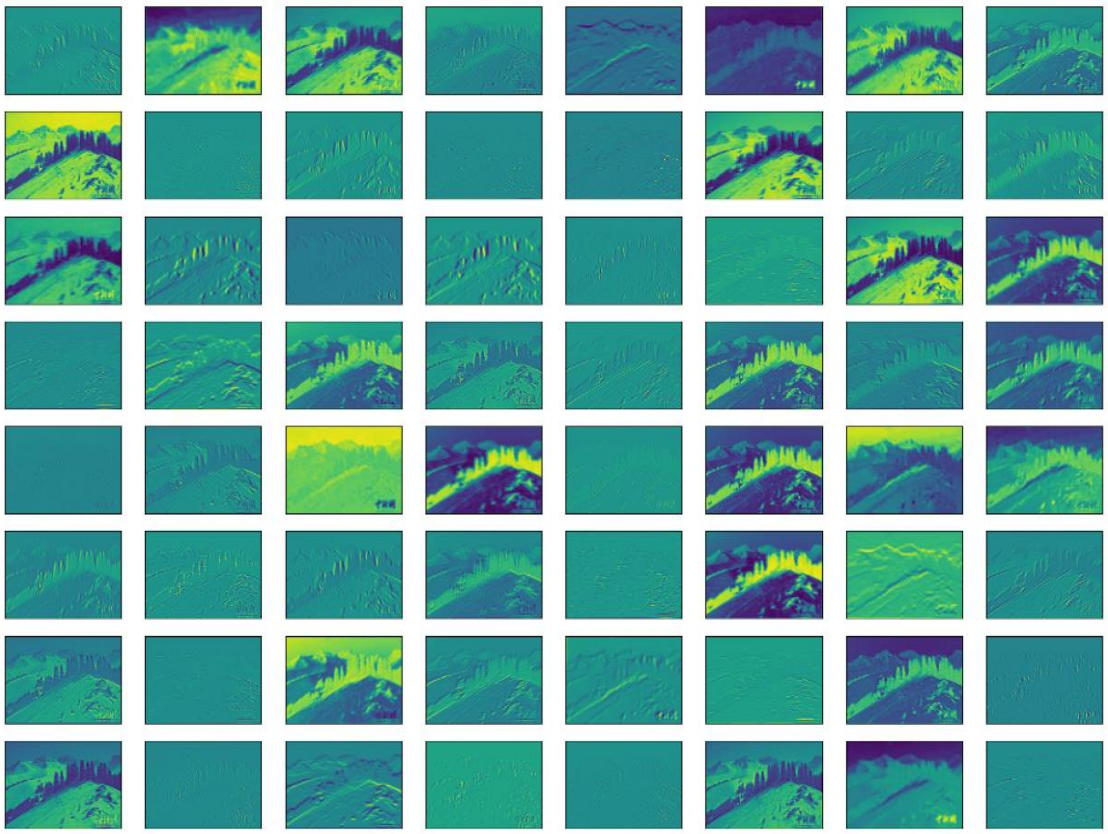
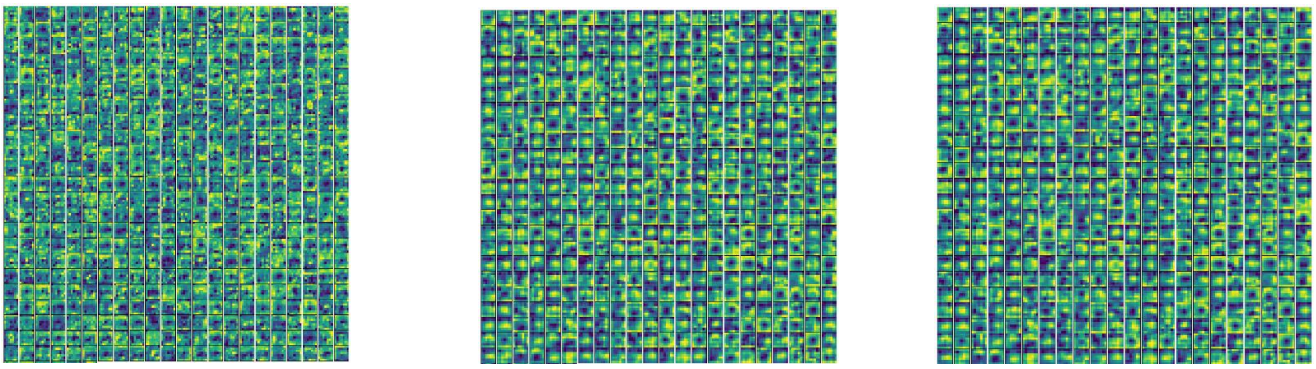


Fig. 11. (a) Feature Map of 1st Convolutional layer of Rest Net 50 trained model.



(b)

Fig. 11. (b) Feature Map of last three Convolutional layers of Rest Net 50 trained model.

4.4 Fine-tuned parameters

While training the model, several hyperparameters and configurations are defined to guide the learning process effectively. The learning rate, set at a value of 0.0001, plays a crucial role in determining the step size of weight updates during training, affecting the model's convergence and stability. A batch size of 16 specifies the number of training samples processed in each iteration, balancing computational efficiency. The training process spans a maximum of 100 epochs, allowing the model to iteratively refine its parameters through multiple passes over the dataset. The optimizer chosen is Adam, a popular optimization algorithm that dynamically adapts the learning rate during training to accelerate convergence. Finally, the categorical cross-entropy loss function is employed to quantify the dissimilarity between predicted and true class probabilities, guiding the model towards more accurate classification results. These carefully selected hyperparameters and configurations collectively contribute to the successful training and performance of the deep learning model as shown in Table 3.

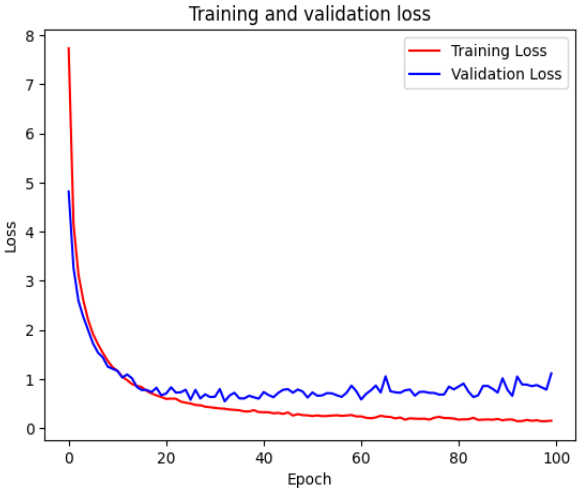
CHAPTER 5

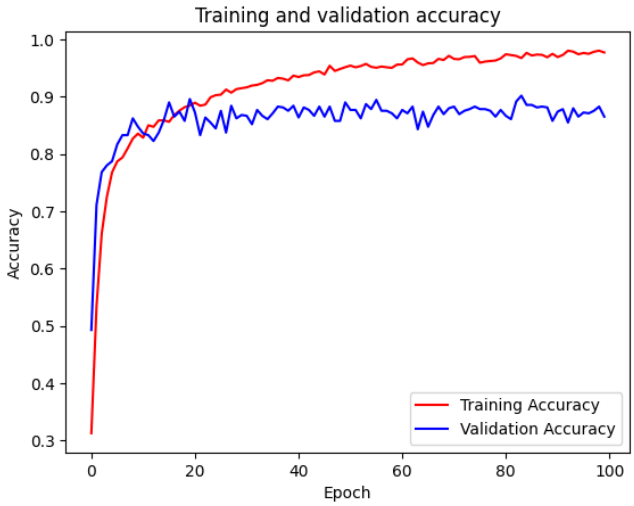
Results And Analysis

5.1 VGG-16 Results

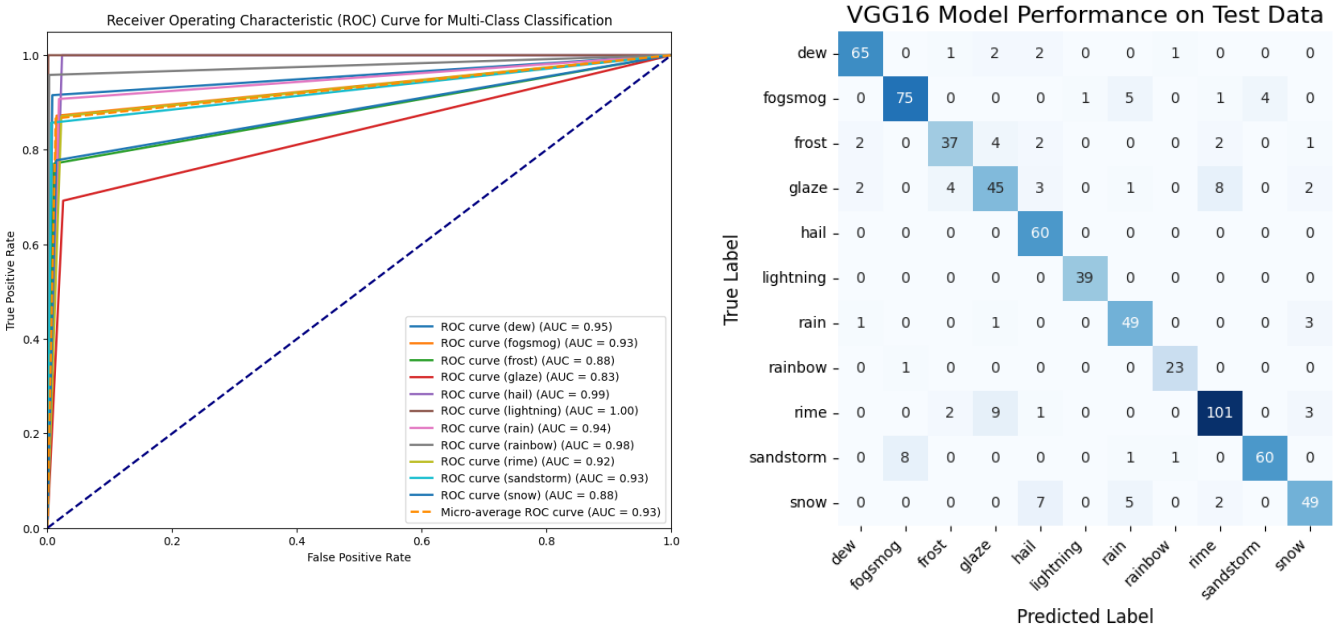
Table.3. Performance scores of fine-tuned VGG16 Model

classes	Precision		Recall		F1-Score	
	Avg.	Std.	Avg.	Std.	Avg.	Std.
Dew	0.946	0.018	0.958	0.038	0.96	0.018
Fog/Smog	0.912	0.031	0.916	0.040	0.918	0.030
Frost	0.862	0.092	0.784	0.079	0.814	0.044
Glaze	0.802	0.053	0.784	0.079	0.792	0.061
Hail	0.926	0.072	0.962	0.047	0.942	0.036
Lightning	0.976	0.013	1.0	0	0.992	0.004
Rain	0.868	0.051	0.890	0.024	0.876	0.026
Rainbow	0.968	0.044	0.984	0.022	0.976	0.026
Rime	0.866	0.027	0.912	0.029	0.890	0.016
Sandstorm	0.964	0.023	0.888	0.052	0.924	0.023
Snow	0.830	0.087	0.788	0.048	0.798	0.045





(a)



(b)

Fig. 12. Accuracy and Loss Curves for a simulation of (a) VGG16 , (b) ROC curve and Confusion matric for VGG-16 Model without augmentation

The VGG16 deep learning Model fits well in the WEAPD Dataset and was able to reach a classification accuracy of nearly 90% with the regularization parameters. But the gap between the training and validation accuracy remains in the model which leads to poor generalization of the weather events.

5.2 VGG-16 With Data Augmentations Results

Table.4. Performance scores with augmentation

Classes	Precision		Recall		F1-Score	
	Avg.	Std.	Avg.	Std.	Avg.	Std.
Dew	0.960	0.010	0.970	0	0.967	0.006
Fog/Smog	0.933	0.025	0.897	0.038	0.913	0.032
Frost	0.883	0.040	0.780	0.104	0.823	0.042
Glaze	0.833	0.029	0.833	0.057	0.830	0.026
Hail	0.937	0.057	0.947	0.058	0.943	0.006
Lightning	0.970	0	1.0	0	0.990	0
Rain	0.903	0.012	0.897	0.012	0.900	0.010
Rainbow	1.0	0	1.0	0	1.0	0
Rime	0.887	0.038	0.937	0.012	0.910	0.0.17
Sandstorm	0.943	0.015	0.943	0.015	0.943	0.006
Snow	0.860	0.050	0.853	0.059	0.857	0.015

Class-wise results (average and standard deviation) for precision, recall and F1-score, for without augmentation & static learning rate and with augmentation & dynamic learning rate, have been noted down in Table 4 and Table 5, respectively. The corresponding average accuracy values are 89.2% and 91.14%. Table 9 shows the comparison of the proposed approach with other popular methods. The best accuracy value obtained with augmentation and dynamic learning rate is 92.10% showing the effectiveness of the proposed method compared to others.

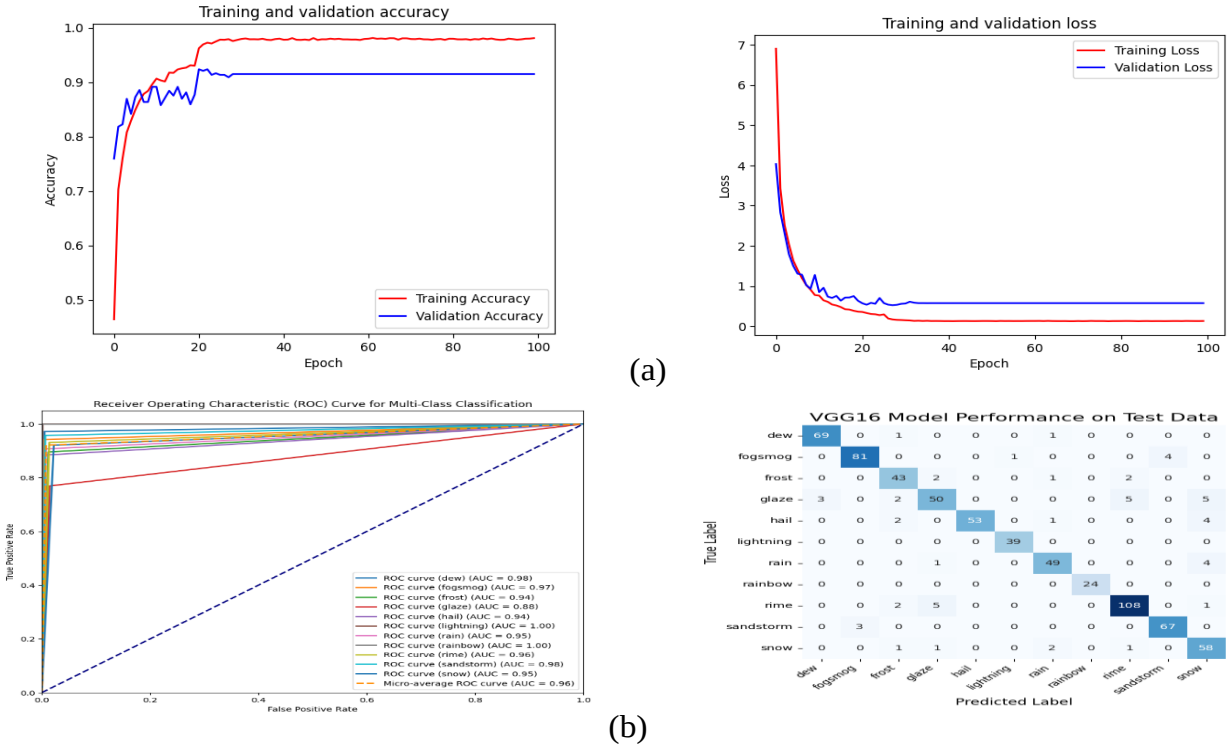


Fig. 13. (a) Accuracy and Loss Curves for a simulation of VGG16, (b) ROC curve and Confusion matric for VGG-16 Model with augmented weather data.

5.3 Ensemble Results

Table.5. Performance scores with Maximum Ensemble having 92.02%

Classes	Precision		Recall		F1-Score	
	Avg.	Std.	Avg.	Std.	Avg.	Std.
Dew	0.98	0.008	0.97	0.008	0.98	0.007
Fog/Smog	0.92	0.023	0.95	0.005	0.94	0.011
Frost	0.86	0.018	0.87	0.024	0.86	0.016
Glaze	0.83	0.035	0.85	0.016	0.84	0.016
Hail	0.95	0.016	0.96	0.008	0.96	0.007
Lightning	1.00	0.00	1.00	0.00	1.00	0.00
Rain	0.93	0.024	0.91	0.021	0.92	0.015
Rainbow	1.00	0.00	1.00	0.00	1.00	0.00
Rime	0.90	0.018	0.91	0.004	0.90	0.013
Sandstorm	0.94	0.012	0.93	0.004	0.93	0.005
Snow	0.85	0.034	0.80	0.043	0.82	0.033

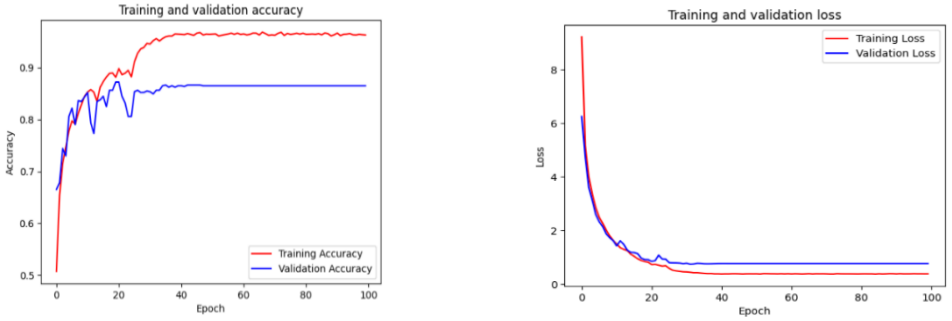
Table.6. Performance scores with GHE ensemble having average 92.75% accuracy

Classes	Precision		Recall		F1-Score	
	Avg.	Std.	Avg.	Std.	Avg.	Std.
Dew	0.99	0.00	0.98	0.010	0.98	0.005
Fog/Smog	0.94	0.008	0.95	0.013	0.95	0.005
Frost	0.88	0.032	0.86	0.026	0.87	0.019
Glaze	0.83	0.033	0.87	0.026	0.85	0.028
Hail	0.95	0.014	0.96	0.010	0.95	0.008
Lightning	1.00	0.00	1.00	0.00	1.00	0.00
Rain	0.95	0.033	0.92	0.014	0.93	0.023
Rainbow	1.00	0.00	1.00	0.00	1.00	0.00
Rime	0.91	0.015	0.90	0.010	0.91	0.010
Sandstorm	0.94	0.011	0.92	0.010	0.93	0.005
Snow	0.86	0.021	0.85	0.066	0.86	0.038

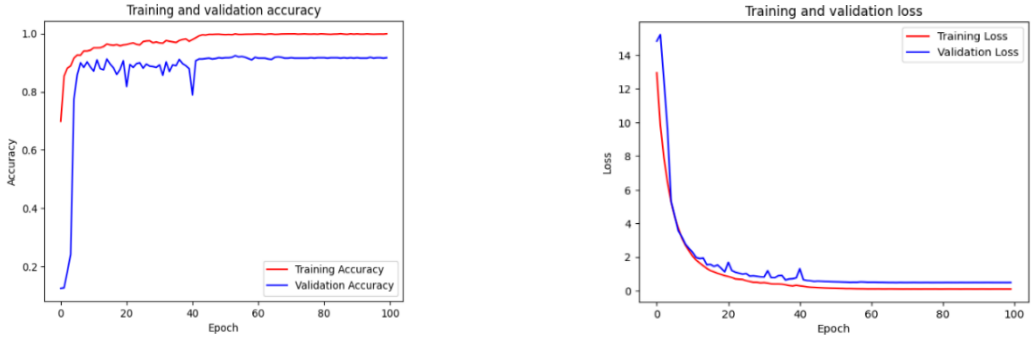
Table.7. Performance scores with Average Ensemble having 92.52%

Classes	Precision		Recall		F1-Score	
	Avg.	Std.	Avg.	Std.	Avg.	Std.
Dew	0.99	0.00	0.98	0.008	0.98	0.004
Fog/Smog	0.93	0.008	0.95	0.013	0.94	0.005
Frost	0.86	0.038	0.85	0.021	0.86	0.02
Glaze	0.84	0.023	0.83	0.028	0.84	0.023
Hail	0.94	0.014	0.97	0.012	0.95	0.008
Lightning	1.00	0.00	1.00	0.00	1.00	0.00
Rain	0.96	0.029	0.90	0.010	0.93	0.011
Rainbow	1.00	0.00	1.00	0.00	1.00	0.00
Rime	0.91	0.024	0.91	0.00	0.91	0.008
Sandstorm	0.93	0.008	0.91	0.008	0.92	0.005
Snow	0.85	0.018	0.90	0.020	0.87	0.004

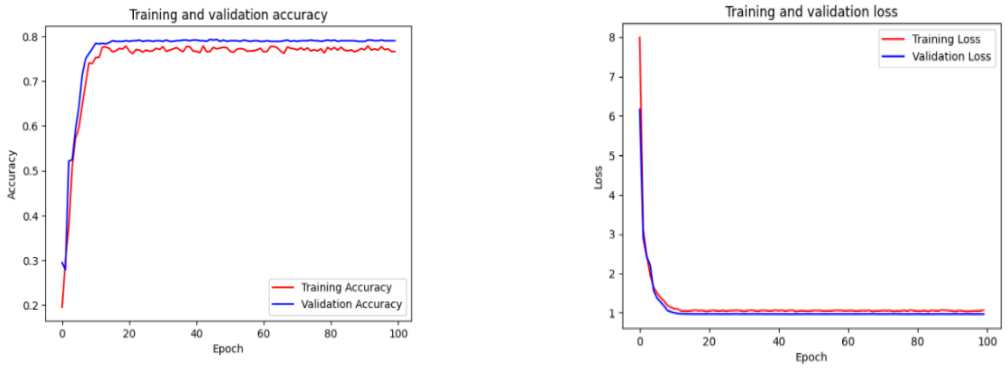
Below are the model metrics – Model accuracy and loss after running fine tuning VGG16 Model on the Training set



(a)



(b)



(c)

Fig. 14. Accuracy and Loss Curves for a simulation of VGG16 (a), RestNet50 (b) and (c) Inception V3

Variations of accuracy and loss values over epochs for both training and validation image sets by tuning hyperparameters on VGG16 architecture with data augmentation and dynamic learning rate for the proposed framework have been shown in Fig 2. From Fig. 2 it is seen that accuracy consistently increases in each epoch till 30 epochs thereafter reaching a stable state. Its robust performance is marked by reduced generalization gaps, ensuring steadfast convergence.

From the Table 6, Table 7 and Table 8 several insights can be drawn from the evaluation metrics for maximum, average and GHE (Geometric-Harmonic Ensemble) mean ensemble models. Firstly, across all models, classes such as “Dew”,” Fog/Smog”, “Hail” and “Lightning” consistently exhibit high precision, recall, f1-score, and accuracy, indicating robust performance in their classification. On the other hand, classes like “Snow” and “Glaze” demonstrate slightly lower performance metrics compared to others, particularly in terms of precision and recall, suggesting potential areas for improvement. Additionally, the ensemble methods show comparable performance to the maximum model, indicating the effectiveness of ensemble techniques in improving classification accuracy.

Class-wise results (average and standard deviation) for precision, recall and F1-score, for without augmentation & static learning rate and with augmentation & dynamic learning rate, have been noted down in Table 4 and Table 5, respectively. The corresponding average accuracy values are 89.2% and 91.14%. Table 9 shows the comparison of the proposed approach with other popular methods. The best accuracy value obtained with augmentation and dynamic learning rate is 92.10% showing the effectiveness of the proposed method compared to others

Loss and Accuracy Curve Plotting: Loss and accuracy curves are plotted for each model during training/validation. Loss curve typically depicts the decrease in model error over epochs, while the accuracy curve shows the increase in model performance over epochs.

Observation: Among the three models, RestNet50 demonstrates the highest accuracy and potentially the lowest loss. This observation suggests that RestNet50 performs the best on the given dataset compared to VGG16 and Inception V3

Ensemble Method: After identifying RestNest50 as the best-performing model, an ensemble approach is employed. Ensemble learning combines predictions from multiple models to improve overall performance. In this case, predictions from VGG16, RestNet50, and Inception V3 are combined to leverage the strengths of each model.

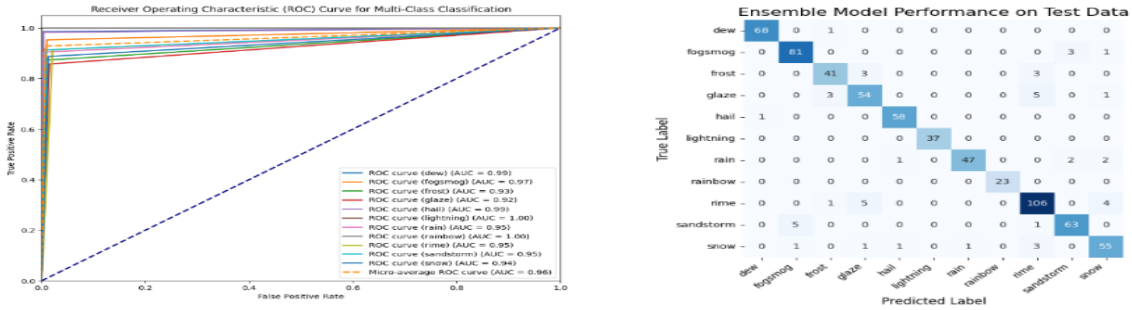


Fig 15: The ROC curve and confusion matrix of the proposed Geometric – Harmonic Ensemble Model on the testing set.

The ROC curve visually shows the model's strong ability to distinguish between positive and negative cases. The steep slope indicates high sensitivity, accurately identifying positive instances. Diagonal elements represent correct classifications, showcasing overall accuracy, while off-diagonal elements highlight areas for improvement. In short, the ROC curves demonstrate the model's effective discrimination and performance at different decision thresholds.

5.4 Comparison Analysis

In comparing the VGG16 model, VGG-16 with augmentation, and the ensemble of VGG16, Inception V3, and ResNet50, several factors come into play. VGG-16, with its simplistic architecture, performs well but can be computationally expensive due to its large number of parameters. Augmenting the data can enhance VGG-16's performance by improving its robustness to variations in input images, potentially

mitigating overfitting. On the other hand, ensembling different architectures offers a robust solution by leveraging the strengths of each model and potentially mitigating individual weaknesses, leading to improved generalization and reduced overfitting. However, ensemble methods demand additional computational resources for training and inference and require careful selection of combination strategies. Ultimately, the choice among these approaches depends on factors such as available computational resources, dataset characteristics, and the desired trade-off between model complexity and performance.

5.5 Comparison With Other State Of The Arts Models

Table 8: Comparison with state-of-the-art methods

Paper	Method	Accuracy (%)
[17]	DenseNet201	83.00
[17]	MobileNet	81.00
[17]	InceptionResNetv2	73.00
[17]	Xception	76.00
[18]	Mete-CNN	92.68
[19]	ViT	87.69
--	**Proposed (GHE)**	**92.82**

Based on the findings presented in Table 9, we can observe the performance of several existing models alongside a newly proposed one called GHE. Among the existing models, DenseNet201 achieves an accuracy of 83%, MobileNet reaches 81%, InceptionResNetv2 scores 73%, Xception demonstrates a notably higher accuracy of 92.68%, and a model named Mete-CNN achieves an accuracy of 87.69%. However, the GHE model surpasses all these existing methods, achieving the highest accuracy of 92.82%.

This suggests that the GHE model exhibits superior performance compared to the current state-of-the-art models in the evaluated task. The results indicate that the innovative approach incorporated into the GHE model could potentially lead to advancements in both accuracy and efficacy over established methods in this domain. In simpler terms, the GHE model performs better than existing models, showing promise for improving accuracy and effectiveness in this particular task.

5.6 Discussion

The geometric harmonic mean (GHM) emerges as an exceptional tool for amalgamating predictions in machine learning endeavors due to its array of beneficial properties. To begin with, its assurance of non-negativity is pivotal, particularly in tasks where predictions or scores must inherently maintain positive values. This feature preserves the interpretability and validity of the ensemble's outcomes, ensuring they remain meaningful and applicable within the context of the problem at hand. Additionally, the scale invariance property of GHM proves

indispensable in scenarios where input values originate from disparate scales across models. By according equal weight to each model's predictions regardless of their magnitude, the ensemble sustains a sense of fairness and consistency in its decision-making process.

Moreover, GHM's resilience to outliers, although not strictly mathematical, is immensely valuable. Its diminished susceptibility to extreme values shields against result distortion, bolstering the reliability of the ensemble's conclusions. Furthermore, the multiplicative nature of GHM preserves the inherent multiplicative relationships between probabilities or likelihoods. This facet is especially advantageous when combining predictions from diverse models like VGG16, ResNet50, and Inception V3, as it ensures the ensemble retains the integrity of these relationships. Consequently, by harnessing these properties, GHM furnishes a robust, balanced, and mathematically grounded methodology for merging predictions. This approach guarantees the integrity and efficacy of the ensemble in diverse machine learning applications, cementing its status as a cornerstone technique in the field.

CHAPTER 6

Conclusion and Future Scope

6.1 Conclusions Of the Present Work

The present work utilized the pre-trained deep neural networks with dynamic learning rate for weather image classification using WEAPD and achieved an impressive accuracy of 92.82% on holdout test images. The application of transfer learning helps facilitate the development of a resilient and potent classifier. Incorporating the Geometric-Harmonic Ensemble (GHE) technique into our methodology further enhances the accuracy and robustness of weather image classification. GHE combines the strengths of geometric mean (GM) and harmonic mean (HM) by averaging their results. The GM emphasizes consensus among models, while the HM gives more weight to smaller values, ensuring that no single model dominates the ensemble. By striking a balance between consensus and diversity, GHE effectively improves classification performance.

6.2 Scope For Future Works

While the proposed methodology demonstrates commendable strengths in weather image classification, it's crucial to recognize inherent limitations. Relying on a pre-trained VGG16 model may present challenges in capturing nuanced features unique to diverse weather phenomena. Introducing a self-attention mechanism, although beneficial for feature extraction, may escalate computational complexity. Future exploration should consider advanced architectures tailored for meteorological imaging, along with efficient attention mechanisms. Robust validation on varied datasets and collaboration with meteorological experts would bolster the model's adaptability. Addressing these considerations holds promise for refining weather image classification methodologies, enhancing accuracy for practical applications in meteorology.

Bibliography

[1] Stabinger, Sebastian, and Antonio Rodriguez-Sanchez. "Evaluation of deep learning on an abstract image classification dataset." Proceedings of the IEEE international conference on computer vision workshops. 2017.

[2] S. B., A. Lesle A., B. Diwakar, K. R. and G. M., "Evaluating Performance of Deep Learning Architectures for Image Classification," 2020 5th International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India, 2020.

[3] A. Alem and S. Kumar, "Deep Learning Models Performance Evaluations for Remote Sensed Image Classification," in IEEE Access, vol. 10, pp. 111784-111793, 2022

[4] Ezat, W. A., M. M. Dessouky, and N. A. Ismail. "Multi-class image classification using deep learning algorithm." Journal of Physics: Conference Series. Vol. 1447. No. 1. IOP Publishing, 2020.

[5] Perez, Luis, and Jason Wang. "The effectiveness of data augmentation in image classification using deep learning." arXiv preprint arXiv:1712.04621 (2017).

[6] A. ElAdel, R. Ejbali, M. Zaied and C. Ben Amar, "Deep learning with shallow architecture for image classification," 2015 International Conference on High Performance Computing & Simulation (HPCS), Amsterdam, Netherlands, 2015, pp. 408-412.

[7] Zhang, Zheng, and Huadong Ma. "Multi-class weather classification on single images." 2015 IEEE International Conference on Image Processing (ICIP). IEEE, 2015.

[8] Zhu, Ziqi, et al. "Extreme weather recognition using convolutional neural networks." 2016 IEEE International Symposium on Multimedia (ISM). IEEE, 2016.

[9] Z. Li, Y. Jin, Y. Li, Z. Lin and S. Wang, "Imbalanced Adversarial Learning for Weather Image Generation and Classification," 2018 14th IEEE International Conference on Signal Processing (ICSP), Beijing, China, 2018, pp. 1093-1097.

[10] Toğaçar, Mesut, Burhan Ergen, and Zafer Cömert. "Detection of weather images by using spiking neural networks of deep learning models." Neural Computing and Applications 33.11 (2021).

[11] Mashudi, N. A., Ahmad, N., Sam, S. M., Mohamed, N., & Ahmad, R. (2022, November). Deep Learning Approaches for Weather Image Recognition in Agriculture. In 2022 IEEE Symposium on Future Telecommunication Technologies (SOFTT) (pp. 72-77). IEEE.

[12] Xiao, H., Zhang, F., Shen, Z., Wu, K., & Zhang, J. (2021). Classification of weather phenomenon from images by using deep convolutional neural network. Earth and Space Science, 8(5), e2020EA001604.

[13] Li, J., & Luo, X. (2023). A Study of Weather-Image Classification Combining VIT and a Dual EnhancedAttention Module. Electronics, 12(5), 1213.

[14] Hameed, Z.; Zahia, S.; Garcia-Zapirain, B.; Javier Aguirre, J.; María Vanegas, A. Breast Cancer Histopathology Image Classification Using an Ensemble of Deep Learning Models. Sensors 2020, 20, 4373.

[15] Pan, Q.; Gao, M.; Wu, P.; Yan, J.; AbdelRahman, M.A.E. Image Classification of Wheat Rust Based on Ensemble Learning. Sensors 2022.

[16] Xue, Dan, et al. "An application of transfer learning and ensemble learning techniques for cervical histopathology image classification." IEEE Access 8 (2020): 104603-104618.

[17] Lei, T.C.; Wan, S.; Wu, S.-C.; Wang, H.-P. A New Approach of Ensemble Learning Technique to Resolve the Uncertainties of Paddy Area through Image Classification. Remote Sens. 2020, 12, 3666.

[18] He, Ziping, et al. "Semi-Supervised anchor graph ensemble for large-scale hyperspectral image classification." International Journal of Remote Sensing 43.5 (2022): 1894-1918.

[19] Deo, Bhaswati Singha, et al. "An ensemble deep learning model with empirical wavelet transform feature for oral cancer histopathological image classification." International Journal of Data Science and Analytics (2024)

[20] Xiao, H. (2021). Weather phenomenon database(WEAPD)(V1). Harvard Dataverse.https://doi.org/10.7910/DVN/M8JQCR

[21] Tammina, Srikanth. "Transfer learning using vgg-16 with deep convolutional neural network for classifying images." International Journal of Scientific and Research Publications (IJSRP) 9.10 (2019): 143-150.

[22] Lu, Y. (2019). Artificial Intelligence: A Survey on Evolution, Models, Applications and Future Trends. Journal of Management Analytics, 6(1), 1-29.

[23] Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT press.

[24] Chatterjee, A. and Ghosh, S. (2023). Skip Flow Net: Integrating Skip Connection and Spatial Dropout in Deep Convolutional Neural Network for Improved Pneumonia Detection. In 2023 2nd International Conference on Ambient Intelligence in Health Care (ICAIHC), India, 1-6, doi: 10.1109/ICAIHC59020.2023.10431443.[4] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep Learning. Nature, 521(7553), 436-444.

[25] Ghosh, S. and Chatterjee, A. (2023). Transfer-Ensemble Learning Based Deep Convolutional Neural Networks for Diabetic Retinopathy Classification," In 2023 3rd International Conference on Advancement in Electronics & Communication Engineering (AECE), India, 489-493, doi: 10.1109/AECE59614.2023.10428233.

[26] Dong, S., Wang, P., & Abbas, K. (2021). A Survey on Deep Learning and its Applications. Computer Science Review, 40, 100379.

[27] Dehuri, S., Ghosh, S., & Cho, S. B. (Eds.). (2011). Integration of Swarm Intelligence and Artificial Neural Network (Vol. 78). World Scientific.

[28] Halder, A., Ghosh, A., & Ghosh, S. (2009). Aggregation Pheromone Density Based Pattern Classification. Fundamenta Informaticae, 92(4), 345-362.

[29] Halder, A., Ghosh, S., & Ghosh, A. (2013). Aggregation Pheromone Metaphor for Semi-Supervised Classification. Pattern Recognition, 46(8), 2239-2248.

[30] Ross, T. J. (2004). Fuzzy Logic with Engineering Applications. Wiley.

[31] Simonyan, K., & Zisserman, A. (2014). Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv preprint arXiv:1409.1556.

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

[33] Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the Inception Architecture for Computer Vision. In 22 Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 2818-2826).

[34] Xiao, H. (2021). Weather Phenomenon Database (WEAPD)(V1).Harvard Dataverse. https://doi.org/10.7910/DVN/M8JQCR.

[35] Zhang, Z.; Ma, H. Multi-Class Weather Classification on Single Images. In Proceedings of the 2015 IEEE International Conference on Image Processing (ICIP), Quebec City, QC, Canada, 27–30 September 2015; pp. 4396–4400.

[36] Zhu, Z.; Zhuo, L.; Qu, P.; Zhou, K.; Zhang, J. Extreme Weather Recognition Using Convolutional Neural Networks. In Proceedings of the 2016 IEEE International Symposium on Multimedia (ISM), San Jose, CA, USA, 11–13 December 2016; pp. 621–625.

[37] Mashudi, N. A., Ahmad, N., Sam, S. M., Mohamed, N., & Ahmad, R.(2022, November). Deep Learning Approaches for Weather Image Recognition in Agriculture. In 2022 IEEE Symposium on Future Telecommunication Technologies (SOFTT) (pp. 72-77). IEEE.

[38] Xiao, H., Zhang, F., Shen, Z., Wu, K., & Zhang, J. (2021). Classification of Weather Phenomenon from Images by Using Deep Convolutional Neural Network. Earth and Space Science, 8(5), e2020EA001604.

[39] Li, J., & Luo, X. (2023). A Study of Weather-Image Classification Combining VIT and a Dual Enhanced-Attention Module. Electronics, 12(5), 1213.

[40] Hanan Aljuaid, Nazik Alturki, Najah Alsubaie, Lucia Cavallaro, Antonio Liotta,Computer-aided diagnosis for breast cancer classification using deep neural networks and transfer learning,Computer Methods and Programs in Biomedicine,Volume 223,2022,106951,ISSN 0169-2607.

[41] M. Zeeshan, B. Salam, Q.S.B. Khalid, S. Alam, R. Sayani, Diagnostic accuracy of digital mammography in the detection of breast cancer, Cureus 10 (4).

[42] H. Zabit, Z. Sofia, G. Begonya, J. J. Aguirre, and A.M.Vanegas, "Breast Cancer Histopathology Image Classification Using an Ensemble of Deep Learning Models", Sensors 20, 4373; doi:10.3390/s20164373, 2020.

[43] Laurie, M. A., Zhou, S. R., Islam, M. T., Shkolyar, E., Xing, L., & Liao, J. C. (2023). Bladder Cancer and Artificial Intelligence: Emerging Applications. Urologic Clinics of North America.

[44] Ghoneim, A., Muhammad, G., & Hossain, M. S. (2020). Cervical cancer classification using convolutional neural networks and extreme learning machines. Future Generation Computer Systems, 102, 643–649.

[45] Liu, Z., Yang, C., Huang, J., Liu, S., Zhuo, Y., & Lu, X. (2021). Deep learning framework based on integration of S-Mask R-CNN and Inception-v3 for ultrasound image-aided diagnosis of prostate cancer. Future Generation Computer Systems, 114, 358–367.

[46] Guyon, I. & Elisseeff, A. An introduction to variable and feature selection. J. Mach. Learn. Res. 3, 1157–1182 (2003).

[47] Krizhevsky, A., Sutskever, I. & Hinton, G. E. ImageNet classification with deep convolutional neural networks. In Advances in Neural Information Processing Systems Vol. 25 (eds Pereira, F. et al.) 1097–1105.

[48] A. Krizhevsky, I. Sutskever, G. E. Hinton. ImageNet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems 25, NIPS, Lake Tahoe, USA, pp. 1097–1105, 2012.

[49] Abokaff, L. Classification of Breast Cancer Diagnosis Systems Using Artificial Intelligence Techniques: Survey. SN COMPUT. SCI. 3, 368 (2022).

[50] M. Ahmed, "A combined feature-vector based multiple instances learning convolutional neural network in breast cancer classification from histopathological images," science direct, 2023.