Comparative analytical study to evaluate the precision and accuracy of weather prediction using numerical data against image classification

Murali, Padma Priya
padmapriyamurali@my.unt.edu

Medam, Rohith
chowdeshwarrohithmedam@my.unt.edu

Burgula,Jayanth Prudhivi
burgulajayanthprudhivi@my.unt.edu

Rachamallu,Jahnavi
jahnavirachamallu@my.unt.edu

February 2023

1 Introduction

Weather is the condition of the atmosphere at a given point in time, influenced by its fundamentals like temperature, air pressure, wind, humidity, etc. Weather is extremely salient and significantly impacts not only humans but also the fauna and flora. The weather indirectly affects its capacity to sustain life and support habitats or bolster the growth of vegetation. Weather also substantially influences the socio-economic conditions of a region, by determining the amount of rain, wind, heat, humidity, etc. it receives. Changes in these elements can affect the availability of water, conditions of soils, the health of the public, working of industries, etc [4]. Numerous other aspects

of our everyday activities and output are also impacted by the weather [5], including what we dress, how we move, how we use solar technology, and more. This makes having a good understanding of weather conditions, and the knowledge and ability to efficiently predict them of the utmost importance.

In the last decade alone, on average there were over 470 weather-related fatalities in the United States [3], all of which could have been mitigated through efficient weather prediction. Predicting weather conditions efficiently can aid sectors like tourism, transportation, agriculture, etc. in turn improving the lives of people. So, what exactly is weather prediction? As the name suggests, it is the process of effectively predicting or forecasting weather conditions for a given region through the application of technology and science. One such application is utilizing specialized hardware which when installed in an area can monitor and forecast weather. These hardware devices are typically expensive and require professional maintenance, thus narrowing their scope of application. So, an alternative would be to use computer vision, a subset of artificial intelligence.

For this project, computer vision is implemented through a machine learning model that makes use of a convolutional neural network (CNN) and deep learning. When large amounts of digital images for different weather conditions are fed through the model, the images are deconstructed into pixels and significant features are extracted. This is repeated over multiple iterations until the model is able to differentiate between images and is able to make predictions [1]. To achieve a contrast against the performance of predicting weather through image classification, a k-nearest neighbors (KNN) model will be implemented on quantitative data that encompasses readings of temperature, humidity, precipitation, etc. to predict a weather condition for a given set of atmospheric readings. The performance of the aforementioned models will be compared and critiqued using their prediction accuracies. Working on and understanding this project and its intent can help establish solid grounds in terms of the type of data required for making weather predictions as they have a lot of use cases. Referring to previous works for this project has helped us to strengthen our grip over the domains of machine learning, in specific deep learning and data prepossessing, both of which are essential to the field of data science. This project will challenge us in terms of our coding as well as reasoning skills. The motivation to take up this topic was

our fascination with the fact that the implementation of this project can be augmented to assist the concept of cloud seeding.

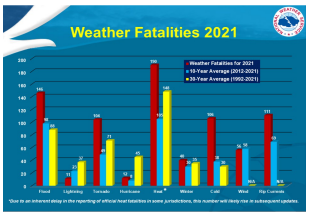

Figure 1: Weather Related Fatality and Injury Statistics

2 Problem Statement

It is challenging to determine the weather from just one image, considering there are more than just 4 weather conditions. With one picture, various weather patterns could be notified immediately. Hence it tends to be challenging to tell whether an image fits into a specific weather class. Deciding the exact weather from a photo would be the fundamental objective of this examination. To settle this issue, we endeavor the first multi-label classification task way to deal with weather acknowledgment, which involves naming a picture with many labels as per the shown weather patterns. A CNN-RNN-based multi-label classification strategy is used. To separate the most related visual data, a channel-wise consideration model is added to the convolutional brain organization (CNN). The Recurrent Neural Network (RNN)goes above and beyond in handling the data and uncovers the connections between the different weather conditions classes. The last stage is a bit-by-bit forecast of the weather condition names. Additionally, we make two datasets for the objective of recognizing the weather and examining the associations between different weather situations.

Proposed Guide: Dr.Stephen Wheeler, College of Information, University of North Texas

3 Related Work

Our first instinct after choosing a research topic was to explore the internet for pertinent papers in order to determine their study constraints. We

looked for ways to overcome these constraints or come up with new hypotheses. The performance of identical models on two different types of datasets is the focus of our research. We specifically aim to establish whether photos provide a more thorough understanding of climatic conditions and result in more precise forecasts. If our theory is correct, we must look into practical methods for putting image capture and analysis into practice.

A novel deep-learning model for multi-label weather categorization that combines the benefits of CNNs and RNNs is presented by [7]. The suggested architecture has potential applications in weather monitoring and forecasting and offers cutting-edge performance on a sizable weather recognition dataset. The application of CNNs and ANNs in weather forecasting is thoroughly reviewed by [2], who also note the possible advantages and difficulties of these techniques. The article can be a helpful resource for academics and professionals interested in researching the application of deep-learning techniques in the field of weather forecasting [6]provided a thorough analysis of the suggested method on a benchmark dataset of weather image data and demonstrate that it outperforms other cutting-edge approaches in terms of accuracy and effectiveness. The authors also provide a thorough analysis of the findings and go through the advantages and disadvantages of the suggested strategy. Through our study we are trying to establish which method of weather prediction would be more accurate, keeping aside computational challenge.

4 Objectives

We have two primary objectives:

1. Examining the reliability of weather predictions: The main goal of the study is to evaluate the reliability of weather predictions made using numerical data and picture categorization. This would entail contrasting the forecasted and actual weather for a specific time frame.

2. Comparing the accuracy of numerical data and picture classification for forecasting the weather: This study will contrast the accuracy of numerical data and image classification for forecasting the weather. This would involve evaluating the level of precision and information provided by each method.

5 Research Design and Methodology

This study compares the prediction accuracy for both picture and numerical data using a variety of machine learning and deep learning models, with a focus on deep learning. Although it is obvious that numerical data normally requires fewer computing resources to analyze, our analysis will necessitate that we make certain assumptions about the data as we write our code. For the data collection part of our research, we used an image dataset released by Mendeley data. These images were collected and sourced by the University of South Africa. When an issue or query involves visual patterns or traits, such as identifying faces, items in a scene, or anomalies in medical scans, photos can be especially helpful. Using computer vision or other machine learning approaches, images can be processed to reveal a plethora of information that may be challenging to define or measure using conventional numerical methods. In our code, we will deploy KNN, CNN, RNN, and multiple other models. When an issue or query involves visual patterns or traits, such as identifying faces, items in a scene, or anomalies in medical scans, photos can be especially helpful. Using computer vision or other machine learning approaches, images can be processed to reveal a plethora of information that may be challenging to define or measure using conventional numerical methods.

6 About the data

We will be using two datasets for our research study. The first dataset would consist of 1125 images classified into two 4 categories -Sunrise, Shine, Rain, and Cloudy. Although we are at the early stages of building our model, if we are able to establish a robust model, more categories can be added by diving deeper into the different types of weather conditions. This study has its fair share of complications. Starting with the dataset itself. Both our datasets are unbalanced within themselves and there is a mismatch in dataset size with each other. Both are datasets are labeled, but that may not always be the case in real life as this is a controlled study. We made some assumptions for our study as mentioned below :

1. We are going to assume that our dataset is balanced, meaning there are an equal number of images representing each category. Data augmentation is an option to equalize the categories.

2. They were retrieved from the same geolocation on the same day at that very instant.

3. We will manipulate the larger dataset to have the same amount of data as our image dataset.

7 Contributions

1. Padma Priya Murali (Team lead) - Involved in creating hypotheses, building research methodology, and acquiring datasets.

2. Rohith Medam - Involved in proposal drafting, introduction, and project brainstorming.

3. Jayanth Prudhivi Burgula - In charge of proposal drafting, reading related work and finding their limitations, and defining project objectives.

4. Jahnavi Rachamallu - Developed problem statement and researched hypothesis.

   All four of us will equally contribute towards the coding aspect of the project

8 Conclusion

Our interest in cloud seeding research served as the fundamental motivation for our undertaking. In addition, the weather has a substantial impact on many daily activities, and since these operations have high costs, accurate forecasting systems are essential. We were motivated to write this study report by the notion that showing a picture of an apple is more effective than discussing it. Training the models on appropriate cloud composition and texture is essential before sending out drones to spray salt into the air. The goal of this study was to persuade data specialists in this discipline to contrast several dataset types that address the same problem. If it is determined that images are a superior type of data, we must begin thinking about how to increase hardware capacity without going over the cost-effective threshold.

References

[1] The International Business Machines Corporation. What is computer vision? https://www.ibm.com/topics/computer-vision. Accessed: 2023-02-15.

[2] Shahab Kareem, Zhala J. Hamad, and Shavan Askar. An evaluation of cnn and ann in prediction weather forecasting: A review. *Sustainable Engineering and Innovation*, 3(2):148–159, 2021.

[3] NOAA's National Weather Service. Weather related fatality and injury statistics. https://www.weather.gov/hazstat/. Accessed: 2023-02-15.

[4] Santa Barbara The University of California. Ucbs science line. https://scienceline.ucsb.edu/getkey.php?key=4580. Accessed: 2023-02-15.

[5] Jingming Xia, Dawei Xuan, Ling Ling Tan, and Luping Xing. ResNet15: Weather Recognition on Traffic Road with Deep Convolutional Neural Network. *Advances in Meteorology*, 2020:1–11, 5 2020.

[6] H. Xiao, F. Zhang, Z. Shen, K. Wu, and J. Zhang. Classification of Weather Phenomenon From Images by Using Deep Convolutional Neural Network. 4 2021.

[7] Bin Zhao, Xuelong Li, Xiaoqiang Lu, and Zhigang Wang. A cnn–rnn architecture for multi-label weather recognition. *Neurocomputing*, 322:47–57, 2018.