WEATHER PREDICTION USING CLOUD IMAGES

A Seminar Report

Submitted to the APJ Abdul Kalam Technological University

in partial fulfillment of requirements for the award of degree

Bachelor of Technology

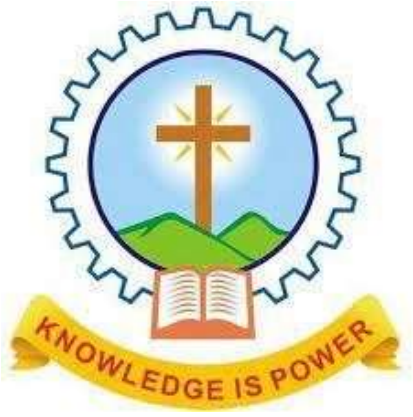
in

Electronics and CoMMunication Engineering

by

AHAMMEDUNNY NAVAS

MAC19EC007



DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

MAR ATHANASIUS COLLEGE OF ENGINEERING, KOTHAMANGALAM

KERALA

JANUARY 2023

DEPT. OF ELECTRONICS & COMMUNICATION ENGINEERING

MAR ATHANASIUS COLLEGE OF ENGINEERING

KOTHAMANGALAM 2019 - 23

CERTIFICATE

This is to certify that the report entitled **WEATHER PREDICTION USING CLOUD IMAGES** submitted by **AHAMMEDUNNY NAVAS** (MAC19EC007), to the APJ Abdul Kalam Technological University in partial fulfillment of the M.Tech. degree in Electronics and Communication Engineering is a bonafide record of the seminar work carried out by him under our guidance and supervision. This report in any form has not been submitted to any other University or Institute for any purpose.

Prof.Belma Joseph
(Project Guide)
Assistant Professor
Dept.of ECE
Mar Athanasius College of Engineering
Kothamangalam

Dr.Thomas George
(Head of Department)
Associate Professor
Dept.of ECE
Mar Athanasius College of Engineering
Kothamangalam

Prof. Arun K L
(Project Coordinator)
Assistant Professor
Dept.of ECE
Mar Athanasius College of Engineering
Kothamangalam

DECLARATION

I AHAMMEDUNNY NAVAS hereby declare that the project report **WEATHER PREDICTION USING CLOUD IMAGES**, submitted for partial fulfillment of the requirements for the award of degree of Master of Technology of the APJ Abdul Kalam Technological University, Kerala is a bonafide work done by me under supervision of Prof.Belma Joseph .

This submission represents my ideas in my own words and where ideas or words of others have been included, I have adequately and accurately cited and referenced the original sources.

I also declare that I have adhered to ethics of academic honesty and integrity and have not misrepresented or fabricated any data or idea or fact or source in my submission. I understand that any violation of the above will be a cause for disciplinary action by the institute and/or the University and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been obtained. This report has not been previously formed the basis for the award of any degree, diploma or similar title of any other University.

Kothamangalam

25-11-2023

AHAMMEDUNNY NAVAS

Abstract

To predict the conditions of the atmosphere for a given location Weather Forecasting is used. It is the application of science and technology. Weather forecast is more helpful for people as it predicts how the future weather is going to be and people may plan accordingly. Farmers will be the most beneficial one's as they may know the rainfall prediction and grow crops accordingly. The weather forecast can be done in many ways like using the previous data or analyzing the current clouds. The authors predict the weather using the status of the clouds. The author used methodologies like Normalization, Clustering, and Cloud mask algorithm to predict the weather more accurately. Normalization is done using RGB values of each pixel. In many fields of research and in industrial and military applications Digital-image processing has become economical.

Acknowledgement

I take this opportunity to express my deepest sense of gratitude and sincere thanks to everyone who helped me to complete this work successfully. I express my sincere thanks to **Dr.Thomas George**, Head of Department, Electronics and Communication Engineering, Mar Athanasius College of Engineering Kothamangalam for providing me with all the necessary facilities and support.

I would like to express my sincere gratitude to the Prof. Arun K L , Department of Electronics and Communication Engineering, Mar Athanasius College of Engineering Kothamangalam for the support and co-operation.

I would like to express my sincere gratitude to the , Department of Electronics and Communication Engineering, Mar Athanasius College of Engineering Kothamangalam for the support and co-operation.

I would like to place on record my sincere gratitude to my project guide Prof.Belma Joseph, Assistant Professor, Electronics and Communication Engineering, Mar Athanasius College of Engineering for the guidance and mentorship throughout the course.

Finally I thank my family, and friends who contributed to the successful fulfilment of this seminar work.

AHAMMEDUNNY NAVAS

Contents

List of Figures

Chapter 1

Introduction

Weather forecasting means predicting the weather and telling how the weather changes with change in time. Change in weather occurs due to movement or transfer of energy. Many meteorological patterns and features like anticyclones, depressions, thunderstorms, hurricanes, and tornadoes occur due to the physical transfer of heat and moisture by convective processes. Clouds are formed by the evaporation of water vapor. As the water cycle keeps on evolving the water content in the clouds increases which in turn leads to precipitation. This is how the convective process happens and also the change in weather. Many factors like temperature, rainfall, pressure, humidity, sunshine, wind, and cloudiness are considered for predicting the weather. It is also possible to identify the different types of clouds associated with different patterns of weather. These patterns of weather help in predicting the weather forecast. In this paper, details of the weather for the past 2 days are considered. Those details are considered as input and performing linear regression and variation of functional regression, the output is obtained. The output is the weather for the next 10 days. Generally, the classification of weather gives 9 classes: clear, scattered clouds, partly cloudy, snow, thunderstorm, rain, overcast, fog, and mostly cloudy. The dataset considered classified all those into 3 classes: moderate cloudy, very cloudy, and precipitation. Linear regression is slowly biased with a high variance model whereas functional is exactly the opposite to it. The collection of more data can improve the linear regression model [2]. Hence the author suggests considering 4 to 5 days of data as input to the model.

1.1 Motivation

Seasons and nature play a major role in agriculture and farming. When it comes to the farming of various fruits, vegetables, and pulses, the temperature is extremely important. Farmers didn't have a better understanding of weather forecasts before, so they had to rely on estimates to do their jobs. They do, however, sometimes suffer losses as a result of inaccurate weather forecasts. Farmers will now get all of their forecasts on their smartphones, thanks to advances in technology and the use of unique weather forecasting mechanisms. Of course, education in this area is critical, but the majority of the farmer community at this point understands the fundamentals, making it simple for them to use the features.

1.2 Objective

People have attempted to predict the weather for centuries. The Babylonians used cloud formations and astrology to forecast conditions in 650 BCE. Aristotle's Meteorologica, written about 350 BCE, identified weather patterns. Theophrastus also compiled the Book of Signs, a book on weather forecasts. Weather prediction lore in China dates back to at least 300 BCE, around the same period as ancient Indian astronomy developed weather-prediction methods. Observed cycles of events, also known as pattern recognition, are used in ancient weather forecasting methods. It has been noted, for example, that if the sunset was especially red, the next day was usually sunny. This knowledge was passed down over the years, resulting in weather lore. However, not all of these forecasts are accurate, and many of them have been proven to be unreliable after being subjected to stringent statistical research.

Chapter 2

Literature Review

[1] The dataset we considered is named “HYTA”. It consists of various images of all types of clouds. We considered 4 clusters for all types of clouds namely: clear sky, sunny, cloudy and sunny, rainy. For each type of cluster, this HYTA dataset consists of nearly 8 to 10 images. Every image consists of the only a plain sky with respective clouds and no other objects like buildings, trees, and poles. In some images, the sun might appear along with the sky and clouds. Along with this standard dataset we considered 4 different datasets. Every data set consists of more than 10 photos for each type of cloud. These datasets are considered to compare the outputs obtained anprovenk the accuracy of the model developed. This comparison will be useful for the future development of the model. In image preprocessing, image data recorded by sensors on a satellite restrain errors related to geometry and brightness values of the pixels. These errors are corrected using appropriate ematical models which are either definite or statistical models. Image enhancement is the modification of an image by changing the pixel brightness values to improve its visual impact.

[2] Image enhancement involves a collection of techniques that are used to improve the visual appearance of an image or to convert the image to a form that is better suited for a human or machine interpretation. Some images (eg. Over water bodies, deserts, dense forests, snow, clouds, and under hazy conditions over heterogeneous regions) are homogeneous i.e., they do not have much change in their levels. In terms of histogram representation, they are characterized as the occurrence of very narrow peaks. The homogeneity can also be due to the incorrect illumination of the

scene. Ultimately the images hence obtained are not easily interpretable due to poor human perceptibility. This is because there exists only a narrow range of gray levels in the image having provision for a wider range of gray levels. The contrast stretching methods are designed exclusively for frequently encountered situations

[3] In this paper, details of the weather for the past 2 days are considered. Those details are considered as input and performing linear regression and variation of functional regression, the output is obtained. The output is the weather for the next 10 days. Generally, the classification of weather gives 9 classes: clear, scattered clouds, partly cloudy, snow, thunderstorm, rain, overcast, fog, and mostly cloudy. The dataset considered, classified all those into 3 classes: moderate cloudy, very cloudy, and precipitation. The least mean square error for the linear regression and variation on functional regression is calculated and learning curves are drawn in this paper. Linear regression is lowly biased with a high variance model whereas functional is exactly the opposite to it. The collection of more data can improve the linear regression model. Hence the author suggests to consider 4 to 5 days of data as input to the model. Both linear regression and functional regression were outperformed by professional weather forecasting services, although the discrepancy in their performance decreased significantly for later days, indicating that over longer periods of time, our models may outperform professional ones. Linear regression proved to be a low-bias, high-variance model whereas functional regression proved to be a high-bias, low-variance model. Linear regression is inherently a high-variance model as it is unstable to outliers, so one way to improve the linear regression model is by collecting more data.

[4] Machine learning, on the contrary, is relatively robust to perturbations and doesn't require a complete understanding of the physical processes that govern the atmosphere. Therefore, machine learning may represent a viable alternative to physical models in weather forecasting. Two machine learning algorithms were implemented: linear regression and a variety of functional regression. A corpus of historical weather data for Stanford, CA was obtained and used to train these algorithms. The input to these algorithms was the weather data of the past two days, which include the maximum temperature, minimum temperature[3], mean humidity, mean atmospheric pressure,

and weather classification for each day. The output was then the maximum and minimum temperatures for each of the next seven days. The segmentation results are provided quickly and with potentially enough accuracy to be integrated into a complete automated weather interpretation system or for cloud cover estimation. Furthermore, in the case of the neural network model it can be successfully concluded from the above results that this unique combination of NAR and NARX neural network produces a positive result and the prediction is accurate to a good extent although there is always there lies a vast possibility of an error currently known as weather phenomenon[2] are such that the features required to be incorporated to create an extremely efficient model are very high, varied and in many cases incalculable. Although efforts in this area will always develop the scientific community as well as the world.

Chapter 3

Methodology

3.1 Image Data Collection

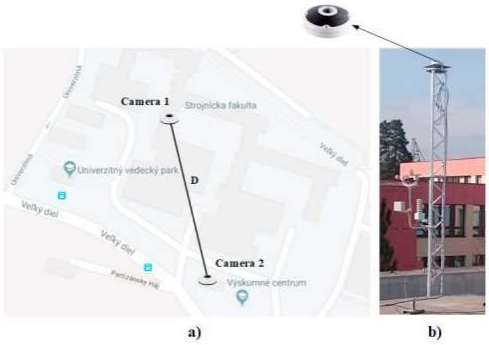


Fig. 1. Camera system; (a) placement of the cameras; (b) a platform with the installed weather station and single IP camera

Figure 3.1: Fish Eye Camera For Image Data Collection

The needed system for collecting sky images and meteorological data consists of two main parts (all-sky image recording, and weather sensor array). For obtaining all-sky images, it is possible to use a fisheye camera. Such a camera has a lens that provides a 360° panorama view and a 180° vertical view. The camera is connected to a server through an Ethernet with an option of wireless communication. Also, it is set to acquire images in the form of video sequences continuously throughout the day. The captured video will be compressed for acquiring less storage and this data will be used for further processing.

3.2 Normalaization Of Image

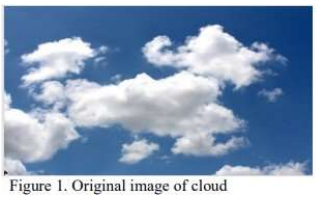


Figure 1. Original image of cloud

Figure 2. Image after performing normalization

Figure 3.2: Fish Eye Camera For Image Data Collection

Pixel values for each and every pixel are considered. Pixel value consists of red, blue, and green color values. These values are extracted from the image with the help of pre-defined libraries in python.

Now with the help of these pixels, we must change the intensity range of the pixelsto [0,1] and increase the intensity to get a clear distinction between the clouds and the sky. Hence the digital picture is normalized. : The input image can be any digital image with the extension .jpg, .jpeg, .png. The output of this module would be a normalized image of the given digital image which seems likely to be a black-and-white or grayscale image.

We used different formulas to get the image normalized using the red, blThe needed green values of each pixel. And the threshold value is generated by taking the mean of all the pixel values. The input image can be of any digital image with the extension .jpg, .jpeg, .png. And the size of the image must vary between 20kb to 20mb. The output of this module would be a normalized image of the given digital image which seems likely to be a black-and-white or grayscale scale image.

3.3 Cloud Masking Algorithm

After normalization, a mean value is generated by adding all the pixel valuesand by dividing it by the total no of pixels. With the help of this mean value, we differentiate the clouds from the input image.

Now we have to extract the feature of the cloud part by again finding the mean value of the cloud area which will be used as a feature in the next process. This process is done for all the images in the dataset so that we get the features of all images which will be used to cluster the images into groups.

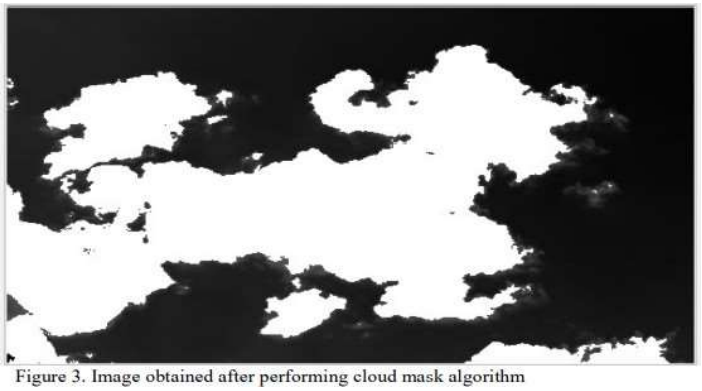


Figure 3. Image obtained after performing cloud mask algorithm

Figure 3.3: Fish Eye Camera For Image Data Collection

The data set is considered and normalization[9] is done for each and every image in it. Then after getting the clouds separated from the image, the mean point is derived from each cloud. Based on those mean points, clouds are separated. Comparing the input image threshold value and the mean value of the cloud, would be pushed into that category of the cloud for which its values coincide.

3.4 K-means Clustsering Algorithm

We considered clustering because for classification there would be less no of classes. But we considered ten types and hence we considered clustering rather than classification [8]. Here we considered the clusters of the clouds as we would divide the image based on the cloud mean point. Among these, the cirrostratus, altostratus, cirrocumulus, and cirrus denotes sunny day. Nimbostratus and Cumulonimbus denotes a rainy day. And cumulus, stratus, altocumulus, and stratocumulus denotes cloudy day. All the classification can be done depending on the mean threshold value. From the dataset, after applying normalization and cloud masking algorithm we can get a threshold value for each and every cloud cluster. Based on that value i.e. the threshold value, the input cloud image is classified into a cluster [9]. Then based on the cluster we can forecast the weather

First, we initialize k points, called means, randomly. We categorize each item to its closest mean and we update the mean's coordinates, which are the averages of the items categorized in that mean so far. We repeat the process for a given number of

iterations and at the end, we have our clusters.

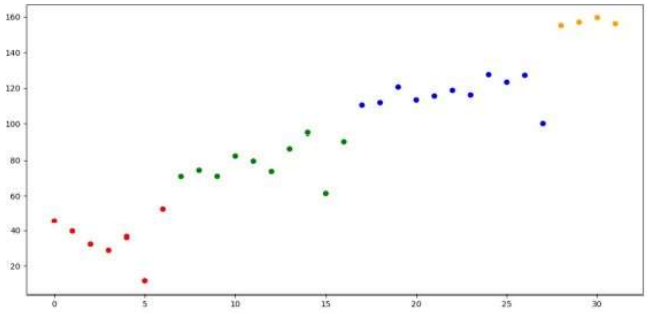


Figure 4. Scatterplot [10] graph showing different clusters formed after training the dataset

Figure 3.4: Fish Eye Camera For Image Data Collection

3.5 Matlab Result

III. EXPERIMENT AND RESULTS

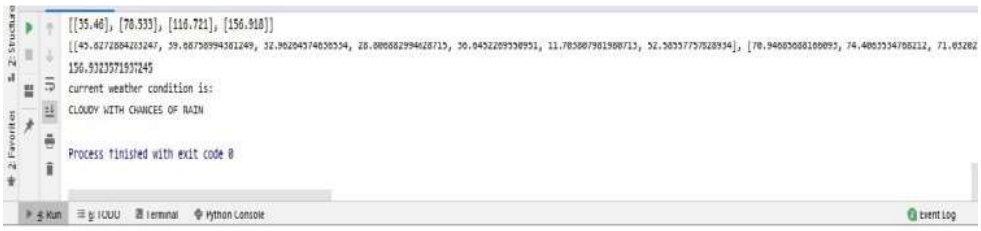


Figure 5. The generated output in the form of text for the given input image

Figure 3.5: Fish Eye Camera For Image Data Collection

The first line in figure 5 shows four different centroids of obtained randomly fromthe HYTA dataset. And the second line shows 4 different lists. Each list contains the points generated after performing a cloud masking algorithm for each image inthe dataset. Every list consists of the points that are corresponded to that respective centroid. The third line is the mean point calculated out of the k means clustering algorithm. At last, it shows the result of the input image i.e. Figure 1 in the text format.

3.6 Software requirements

[1].Python : Python is an interpreted high level andgeneral-purposee programing language created by Guido van Rossum and first released in 1991, Python's design philosophy emphasizes code readability with its notable use of significant whitespace.Its language constructs and object-oriented approach aim to help programmers write clear, logical code for small and large-scale projects. Python is dynamically typed andgarbage-collected. It supports multiple programming paradigms, including procedural,object-oriented, and functional programming.

[2]. PyCharm: PyCharm is the only integrated development environment (IDE) used in computer programming, specifically for the Python language. It is developed by the Czech company JetBrains. It provides code analysis, a graphical debugger, an integrated unit testertegration with version control systems (VCSes), and supports web development with Django as well as Data Science with Anaconda. PyCharm is cross-platform, with Windows, macOS, and Linux versions. The Community Edition is released under the Apache License and there is also Professional Edition with extra features – released under a proprietary license

[3]. OpenCV: OpenCV (Open source computer vision) is a library of programming functions mainly aimed at real-time computer vision. Originally developed by Intel, it was later supported by Willow Garage and then Itseez (which was later acquired by Intel). The library is cross-platform and free for use under the open- 50 source BSD license. OpenCV supports some models from deep learning frameworks like TensorFlow, Torch, and PyTorch (after converting to an ONNX model) and Caffe according to a defined list of supported layers. It promotes OpenVisionCapsules. which is a portable format, compatible with all other formats.

[4].NumPy: NumPy is a library for the Python programming language, that addssupport for large, multi-dimensional arrays and matrices, along with a large collectionof high-level mathematical functions to operate on these arrays. The ancestor of NumPy, Numeric, was originally created by Jim Hugunin with contributions from several other developers. In 2005, Travis Oliphant created NumPy by incorporating features of the competing Numarray into Numeric, with extensive modifications. NumPy is open-source software and has many contributors.

3.7 Hardware Requirements

[1]Processor: 64-bit, quad-core, 2.5 GHz minimum per core

[2]Ram: 4 GB or more

[3] Hard disk: 20 GB of available space or more.

[4] Display: Dual XGA (1024 x 768) or higher resolution monitors

Chapter

Weather Prediction In Kochi

Figure 4.1: Fish Eye Camera For Image Data Collection

Working on the data obtained by the wind profiles in the atmospheric region. 690 antennae arranged in a rectangular grid. Working on 205Mhz frequency range. Each antenna consists of a transmitter and receiver module and all these are connected to the processor for the prediction of weather. Cochin University of Science and Technology launched its initiative utilizing the data generated using the facility at the site of its Rs. 20 crore Stratosphere Troposphere (ST) Radar.

Chapter

Google Weather Prediction

.

Using radar images, Google treats this as a computer vision problem. They use a "data-driven physics-free approach," which means they are not using atmospheric conditions and physics to predict the weather. Instead, they treat weather prediction as an image-to-image translation problem. One where image analysis of radar images and the use of convolutional neural networks (CNNs) can be utilized to predict the weather. Currently, the National Oceanic and Atmospheric Administration (NOAA) collects around 100 terabytes of data per day. This data is fed into supercomputers that provide 1 to 10 day forecasts through numerical computation of several physical processes such as atmospheric dynamics, thermal radiation, vegetation, lake and ocean effects, etc.

A convolutional neural network (CNN) is made up of a linear sequence of layers. Each layer can do something different. One layer may decrease the resolution of

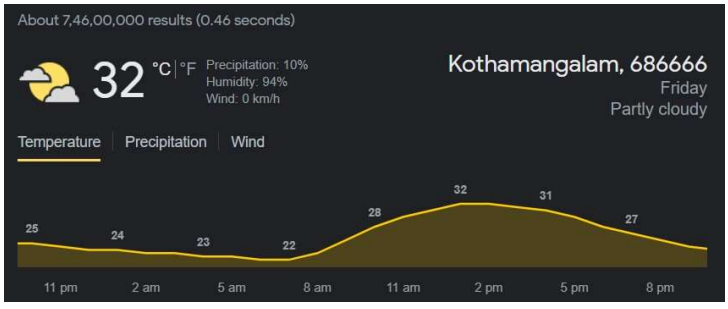


Figure 5.1: Schematic Model of Motor Driver section

an image, another layer may change the number of channels in that image. Upon going through each layer, the original image is transformed into a new output image. With each layer, the image is also convolved using a set of convolutional filters. Convolution is defined as a function derived from two given functions by integration which expresses how the shape of one is modified by the other.

Chapter 6

Conclusion and Future Work

Generally, all the other weather forecasting applications and sources would give the weather report of that particular area. Using the GPS, the location would be tracked and using satellite information, weather conditions would be given at that place. The types of clouds the author considered can deliberately give an accurate condition of the weather. For now, the model can give the weather condition at that point of time.

6.1 FUTURE WORK

Generally, any weather forecasting applications and sources would give the weather report of a particular area with the help of GPS [11] or using satellite information. Our model can give the weather condition at any point in time for any place with the help of the current cloud image at that place. In future, this model can be developed as to predict the weather for the next few hours based on the image with the help of cloud analysis.

References

[1] Kapadia Nilay S, Urmil Parikh in Nov 2016, Weather Forecasting using Satellite Image Processing and Artificial neural networks. IJCSIS vol 14, No.11,pp. 3724-3738, May 2017, doi: 10.1109/TVT.2016.2611655.

[2] MinakshiGogoi and Gitanjali Devi (Oct-Dec 2015), Cloud image analysis for rainfall prediction, Advanced Research in EEE.

[3] Calbo, J., and J. Sabburg, 2008: Feature extraction from whole-sky ground-based images for cloud-type reingnition. J. Atmos. Oceanic Technol., 25, 3–14

[4] J.https://www.science-emergence.com/Articles/How-to-create-a-scatter-plot-withseveral-colors-in-matplotlib-/