Weather Forecasting Using Digital Image Processing

Mrs. G. Gowri Pushpa
Department of Computer Science and Engineering
ANITS, Vishakhapatnam, Andhra Pradesh, India
gowripushpa11@anits.edu.in

Hemaja Patoju
Department of Computer Science and Engineering
ANITS, Vishakhapatnam, Andhra Pradesh, India
phemaja@gmail.com

G. Sai Charan
Department of Computer Science and Engineering
ANITS, Vishakhapatnam, Andhra Pradesh, India
charan4899@gmail.com

M. Kali Charan
Department of Computer Science and Engineering
ANITS, Vishakhapatnam, Andhra Pradesh, India
Marpukalicharan@gmail.com

Pranita Jagtap
Department of Computer Science and Engineering
ANITS, Vishakhapatnam, Andhra Pradesh, India
2494pranita@gmail.com

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.

Keywords - Weather forecasting, normalization, clustering, and cloud mask algorithm.

I. INTRODUCTION

Weather forecasting [1] 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 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 the past, people used barometric pressure, current weather conditions, sky condition to predict whereas now there are many computer based models that consider the atmospheric factors to predict the weather. These methods are not accurate and the reason is due to the chaotic nature of the atmosphere as it keeps on changing. Even predicting weather for a longer period of time will not be accurate that is why most of the current forecasting [1] models predict weather

only for a couple of days not more than 10. The accuracy gets reduced with increase in time. We researched some of the papers and cloud types for this paper. They are

Machine learning applied to weather forecasting:

In this paper, details of weather for the past 2 days are considered. Those details are considered as input and performing linear regression and variation of functional regression, output is obtained. The output is weather for next 10 days. Generally the classification of weather gives 9 classes: clear, scattered clouds, partly cloudy, snow, thunder storm, rain, overcast, fog, mostly cloudy [2]. The dataset considered classified all those into 3 classes: moderate cloudy, very cloudy, 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 low biased with high variance model whereas functional is exactly opposite to it. 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.

Analysis on various techniques for weather forecasting:

1) Support Vector Machines: To predict the maximum temperature of a required location Support Vector Regression (SVR) is used. It performs better than MLP which is trained with back propagation algorithms as it minimizes the upper bound on generalization error. By selecting proper parameters it can replace neural networks based models for applications of weather prediction.

2) Time Series Analysis for Weather Forecasting: Data groups and data variables in the specified time are captured by Time Series Analysis. By comparing actual and predicted values of temperature, the forecasting reliability was evaluated. The results show that important tool for temperature forecasting is network.

3) Prediction of Weather by using Back Propagation Algorithm: Wind, humidity, rainfall and temperature are the parameters recorded using sensors. Using these sensors weather forecasting and processing information is transferred [3]. It classifies compares and predicts the change in other weather parameters by changing any one parameter value that those sensors recorded. A 3 layered neural networks trained with the existing dataset to develop a relation among the parameters of weather that are non-linear.

4) Fuzzy Logic Based Rainfall Prediction model: Two components are made in a developed fuzzy logic model where one is knowledge based and the other is fuzzy reasoning or decision making. Using fuzzification and defuzzification operations outputs are predicted compared with actual rainfall data [3]. A fuzzy model that is well developed is capable of handling the data that is scattered and shows flexibility in modelling weak input and output variable relationship.

Weather forecasting using data mining research based on cloud computing:

A modern method is developed which is service oriented architecture for the weather information systems that forecasts weather using data mining techniques. The method uses Artificial Neural Network and Decision tree Algorithms and meteorological data collected in specific time [4]. It presents the best results for generating classification rules for the mean weather variables. The model predicts temperature, rainfall and wind speed. Cloud computing reduces the cost of infrastructure and storage as it ensures secure reliable and efficient services for the user.

Cloud image analysis and classification:

Generally clouds are classified on the parameters like temperature, shape, colour, density, spectral clustering analysis. Cirrus, Stratus and Cumulus are the 3 basic types of clouds. Based on their height and texture these clouds are divided into 10 other types [5].

1. Cirrus
2. Cirrocumulus
3. Cirrostratus
4. Altocumulus
5. Altostratus
6. Nimbostratus
7. Stratocumulus
8. Stratus
9. Cumulus
10. Cumulonimbus

The dataset we considered is named “HYTA”. It consists of various images of all types of clouds. We considered 4 clusters for all the 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 only 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 and check the accuracy of the model developed. This comparison will be useful for the future development of the model.

II. PROPOSED ALGORITHM

Weather forecasting can also be done by using satellite images but acquiring the satellite images is more difficult and would even cost high. Even predicting using the satellite images needs more technology. So, we are using digital image processing techniques which process the images of the sky like normalization, cloud masking algorithm and k-mean algorithm.

Every system should be divided into modules for better understanding and execution. Dividing into modules helps the programmer and client to work and use the system efficiently, respectively. If any system is not divided into modules and worked as a whole, then there came a numerous errors. Even we find difficulty in correcting those errors. It is must and should to divide the total project into modules and work on each and every module independently to get effective results. Our total system is divided into three modules namely:

1. Normalization of Image
2. Cloud masking algorithm
3. K-means clustering.

2.1. Normalization of Image:

Step 1: Pixel values for each and every pixel are considered. Pixel value consists of red, blue and green colour's values. These values are extracted from the image with the help of pre-defined libraries in python.
Step 2: Now with the help of these pixels, we must change the intensity range of the pixels to [0,1] and increase the intensity to get a clear distinction between the clouds and the sky. Hence the digital picture is normalized [6].
Step 3: The input image can be of any digital image with the extension .jpg, .jpeg, .png.
Step 4: 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 gray scale image.

Figure 1. Original image of cloud

Figure 2. Image after performing normalization

2.2. Cloud masking Algorithm:

Step 1: After normalization, a mean value is generated by adding all the pixel values and by dividing it by the total no of pixels. With the help of this mean value we differentiate the clouds from the input image.
Step 2: 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 a feature in the next process [7].
Step 3: This process is done for all the images in the dataset so that we get features of all images which will be used to cluster the images into groups. The output of this process is shown in Figure 3.

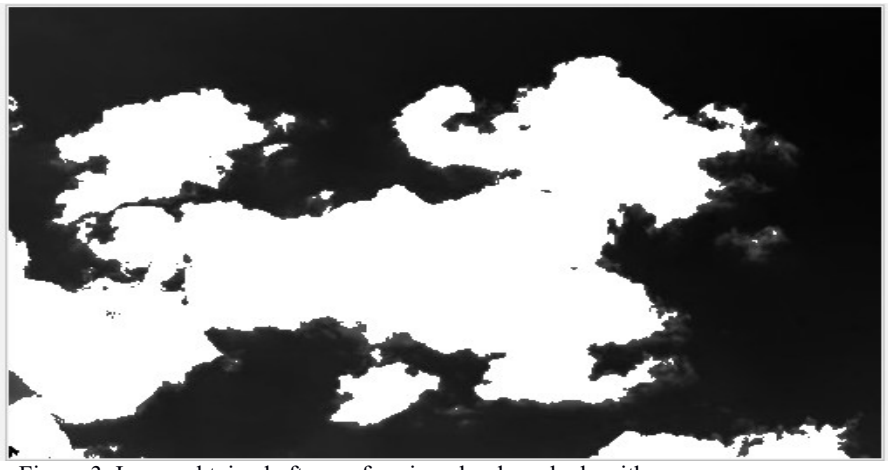

Figure 3. Image obtained after performing cloud mask algorithm

2.3. K-means Clustering 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.

In this the clouds are divided into 10 clusters. The clusters are:

1. Cirrostratus
2. Cirrus
3. Cirrocumulus
4. Altocumulus
5. Altostratus
6. Stratus
7. Stratocumulus
8. Nimbostratus
9. Cumulonimbus
10. Cumulus

Among these the cirrostratus, altostratus, cirrocumulus, cirrus denotes sunny day. Nimbostratus and Cumulonimbus denotes 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.

ALGORITHM:

1. First we initialize k points, called means, randomly.
2. 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.
3. We repeat the process for a given number of iterations and at the end, we have our clusters.

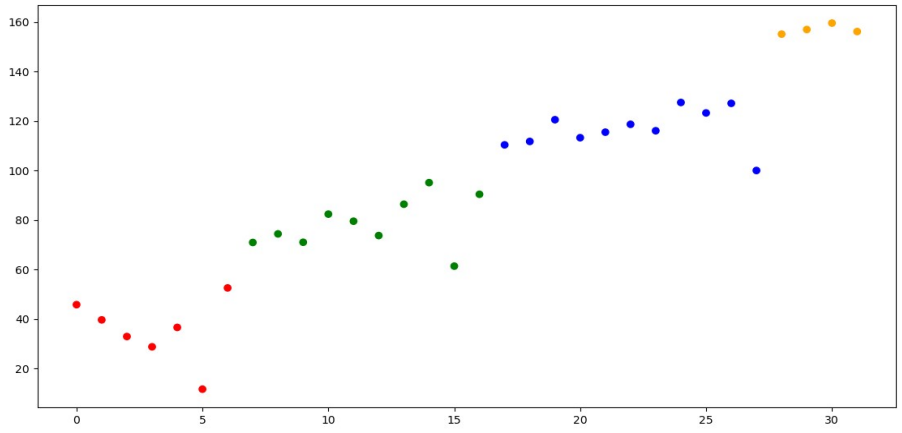

Figure 4. Scatterplot [10] graph showing different clusters formed after training the dataset

III. EXPERIMENT AND RESULTS

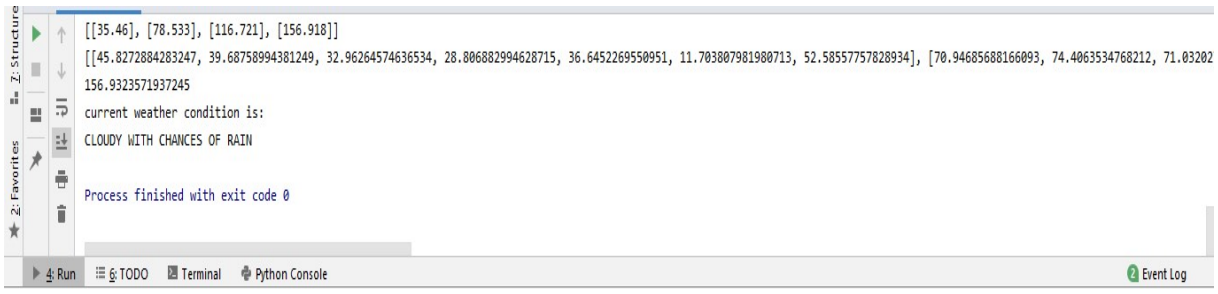

Figure 5. The generated output in the form of text for the given input image

The first line in figure 5 shows four different centroids of the obtained randomly from the HYTA dataset. And the second line shows 4 different lists. Each list contains the points generated after performing cloud masking algorithm to each image in the 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.

IV.CONCLUSION

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 of 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.

This paper can be extended to get weather forecast for the next few days; we can modify our system by using different algorithms and use that as an extension to our current project.

REFERENCES

[1] Weather Forecasting using Satellite Image Processing and Artificial Neural Networks by Nilay S. Kapadia , Urmil Parikh in IJCSIS vol 14,No.11, Nov 2016.

[2] Machine Learning Applied to Weather Forecasting Mark Holmstrom, Dylan Liu, Christopher Vo Stanford University, December 15, 2016.

[3] ANALYSIS ON THE WEATHER FORECASTING AND TECHNIQUES Janani.B, Priyanka Sebastian, Jan 2014.

[4] The Weather Forecast Using Data Mining Research Based on Cloud Computing. ZhanJie Wang and A. B. M. Mazharul Mujib,2017.

[5] Cloud image analysis for rainfall prediction: A survey by Minakshi Gogoi and Gitanjali Devi, Advanced Research in EEE, Oct-Dec 2015.

[6] https://www.mathworks.com/matlabcentral/answers/422493-how-to-do-normalized-blue-red-ratio-operation-of-an-image

[7] https://medium.com/sentinel-hub/improving-cloud-detection-with-machine-learning-c09dc5d7cf13

[8] https://www.geeksforgeeks.org/k-means-clustering-introduction/

[9] https://scied.ucar.edu/learning-zone/clouds/how-clouds-form

[10] https://www.science-emergence.com/Articles/How-to-create-a-scatter-plot-with-several-colors-in-matplotlib-/

[11] https://www.universityworldnews.com/post.php?story=20120620192857538