Weather Classification Model Performance: Using CNN, Keras-TensorFlow

Ashish Sharma [1*], and *Zaid* Saad Ismail [2]

[1]Department of Information Technology, College of Engineering & Computer Science, Lebanese French University, Erbil, Kurdistan, Iraq
[2]Department of Accounting, College of Administrative and Financial Science, Knowledge University, Erbil, Kurdistan, Iraq

Abstract. Nowadays, automation is at its peak. The article provides a base to examine the weather through the machine without human intervention. This study offers a thorough classification model to forecast a weather type. Here, the model facilitates defining the best results for the weather prediction model to any climatic zones and categorizes the climate into four types: cloud, rain, shine, and sunrise. This model designs and reveals using convolution neural networks (CNN) algorithms with Keras framework and TensorFlow library. For practical implementations, use the images dataset available from the kaggle.com website. As a result, this research presents the performance of the designed and developed model. It shows accuracy, validation accuracy, losses, and validation losses approximately 94%, 92%, 18%, and 22%, respectively.

Index Terms: Climate, Convolution Neural Network (CNN), Keras, TensorFlow, Weather Classification Model.

1 Introduction

This article represents the weather classifications model without human intervention. The paper introduces an intensified climate foresight model that employs the notion of data mining in weather prediction. The model can divine any characteristic feasible in the dataset while other data mining-based algorithms produce particular traits. It is computationally valuable and executes appropriately to mobile devices like the Android ecosystem [1]. It uncovers the distinct climatic conditions and separates the cloudy, rainy, shiny, and foggy patterns. Here, the genetic classifications comprise based on the geographic determinants of the environment and the surface energy resources, plus concerns of the air-mass review [2, 6].

It introduces a system that supports a regional weather type class procedure to aid interpretation in city atmosphere considerations. Here, this helps analyze the change in particular environmental conditions from hot to cold or cold to hot, high to low, or low to

* Corresponding author: ashish.sharma@lfu.edu.krd

increased wind speed. As it happens frequently, it intends to define future observations appropriately [3]. It proves climate change peculiarity technically on the baseline high-resolution climatic projections. Both periods measure confidence levels of the assemblage range, presenting relevant implications regarding the authenticity, like the accuracy of the classifications [4].

This system fulfills probability density functions (PDF) to perceive and simulate temperature and precipitation according to zones for various climates. The research examines temperature and precipitation twists above current weather results and identifies systematic errors of the model to acquire better performance. This system comprises many models from different regions to achieve accuracy with results [5].

Machine learning techniques are worthwhile in developing real-time applications, including computer vision, image processing, speech recognition, text processing, Etc. Moreover, it provides beautiful results for classifying text data according to distinct climatic conditions like temperature and humidity [7].

A convolution neural network (CNN) is an artificial neural network (ANN) specified for image pattern detection. It is suitable for image analysis to identify the features of an image, called feature maps. CNN is composed of hidden layers are known as convolutional (ConV) layers. These ConV layers accept input, transform input into some mode, and convert it to output. Now the same production is information to the next layer. This modification is a convolution operation that facilitates detecting patterns from images precisely and accurately. Each convolution layer needs to specify the number of filters. The filters play an essential role in detecting edges called "edge filters" and detecting objects called "object filters." So, the CNN technique is adopted for this proposed work to identify the pattern of the weather images dataset [13].

This work investigates the performance of the designed and developed deep learning model by using CNN architecture with Keras framework and TensorFlow system-defined libraries for Classification on the original Kaggle.com website weather image dataset. Finally, these deep learning algorithms work for two primary responsibilities: Classification and predicting climate situations.

2 Literature review

The author studied Kappen's climate classification schemes. It facilitates regarding analysis of weather changes according to regions and climatic cycle time. Here due to doubts among the classification models to identify variation in the climate zone with measurements and forecasts until outcomes are unclear [12].

The scholars described pressure Laplacian and the geopotential Laplacian algorithms for climatic classifications and represented the output with the help of the regional climate models (RCM) simulation model. Any severe environmental conditions like dryness, floods, high-speed wind, heavy snowfall, Etc., harm biodiversity [11].

The researchers described that global weather types understanding is a great challenge. Manually categorizing different weather types as per temperature and precipitation is difficult to confirm for different longitude, latitude, altitude, Etc. It is an arduous task due to the availability of data resources according to areas. Consequently, mentioned above, severe climatic conditions support further research in weather classifications [8, 10, 11].

The authors compared several high performable machine learning models on multi-class datasets. The various systems developed with different algorithms such as decision tree CART, gradient boosting, KNN, linear regression, lasso, ridge, MLP, deep learning, SVM, and random forest achieve accuracy for weather classification. These models are showing some overfitting. As per the testing phase, the linear regression technique is highly performable for a prediction model. Furthermore, the k- nearest neighbors algorithm is lower in performance [7].

The further authors explained the spatial augmentation of a gridded weather typing classification (GWTC) design to a global rule. Several methods apply for making the system robust in terms of weather classification to recognize the severe climatic conditions [6].

The scholars outlined new weather datasets correlated with the present experimental setup to simulations of REMO from the CORDEX structure. These simulators tend to analyze the climate classifications. The REMO produces relatively large yearly precipitation and temperature inclinations in dry and arctic zones and low skill. The biased outcome always represents an uncertainty to measurements of any weather data result [5].

This research focuses on biodiversity problem-solving strategies. The author studied for one (01) KM region on the map and investigated weather changes. The results presented with higher accuracy and climatic conditions are suitable for vegetation, spices [2, 4].

Moreover, the intellectuals examine the weather classification through the local weather type (LWT) approach. This method is getting direct based on atmospheric data and statistical techniques. It facilitates an analysis of the atmospheric parameters near the ground. It is a proven procedure to scrutinize future problems before time [3].

The researchers observed the difference between k-nearest neighbors' (KNN) and modified KNN (MKNN) for weather classification. Here it shows that the MKNN provides more reliable results than other datamining projection ways [1].

3 Methodology

3.1 Algorithm for Weather Classification Model

This algorithm accomplishes ten steps process to design and develop a weather classification model. These are the following steps:

[1]. Initiate a model
[2]. Import all of the required system defined libraries
[3]. Preparation of image dataset for training the machine and validating the machine model
[4]. Show graphically for classifying the training and validating data into distinct classes
[5]. Build a model using CNN
[6]. Compile a model
[7]. Train a model as per the training dataset
[8]. Validate a model as per the validation dataset
[9]. Display the performance of an implemented model graphically
[10]. Finish the process

3.2 Pseudo-code of the proposed algorithm

The pseudo-code of the proposed algorithm for a weather classification model, such as:

Step-1. *Initiate a model.*

Step-2. *To import all the required inbuilt libraries, such as :*
TensorFlow, Keras, NumPy, Matplotlib

Step-3. *To prepare data for training the model and validating the model.*

- *Set path for image data collection*
- *Classify the image data into four classes: Cloud, Rain, Shine, and Sunrise.*

- *Divide image data into distinct batches, i.e., TRAINING_DIR and VALIDATION_DIR.*

Step-4. *Display bar-graphs*

- *Plot bar-graph for Classification of the training dataset.*
- *Plot bar-graph for Classification of the validating dataset.*

Step-5. *Build a new Weather Classification model by CNN.*

- *Design a sequential model with convolution 2D filters, kernel-size, activation function 'relu,' and padding function 'same'; maximum pooling size of 2D images; add a dense layer with activation function' softmax.'*
- *Summarize the model.*

Step-6. *Compile a new Weather Classification model.*

Step-7. *Train a new Weather Classification model as per the collected images dataset.*

Step-8. *Validate a new Weather Classification model as per the collected images dataset.*

Step-8. *Plot the line graphs to show the performance of 'Training and Validation Accuracy' and 'Training and Validation Loss.'*

Step-9. *End.*

4 Results and discussion

The proposed implemented pseudo-code outcome for the weather classification model is step-by-step performance. The resultant of the proposed model contains as:

4.1 Prepare data

The initial phase prepares the image dataset for further processing with the implementing model. Here, the data details are as follows. See Figure-1 to clarify the details of each type of image for Train and Test (Validation) Datasets.

```
Train set Cloud 240
Train set Rain 173
Train set Shine 203
Train set Sunrise 287

Test set Cloud 60
Test set Rain 42
Test set Shine 50
Test set Sunrise 70
```

Fig. 1. Prepare dataset.

4.2 Exploratory data analysis

It categorizes the data into four classes: Cloud, Rain, Shine, and Sunrise. See Figure-2.

```
['Rain', 'Shine',
'Sunrise', 'Cloud']
```

Fig. 2. Exploratory data analysis.

4.3 Classification of training and validation datasets

After preparing the dataset, it classifies into distinct parts as per the model's requirement. This practice helps in training and then validating the proposed weather classification model. Here, it represents the data in a bar graph as per class. Each bar shows different data sets like cloud image data, rain image data, shining image data, and sunrise image data for training and validation datasets. It shows in Figure-1 for Classification of Training Dataset and Classification of Validation Dataset, respectively.

4.4 Implementation of weather classification deep learning CNN model

This proposed sequential weather classification model architecture comprises eighteen layers: nine Convolution, five MaxPooling, one Flatten, and three Dense layers. This designed model extracts all of the features from dataset images. Accordingly, train the model as per extracted features. It is more and more effective with cleaned data. Thus, it defines that provide the best clean images data to model and get high performing results. Hereafter training, it expresses, as a result, all of the parameters that are 10,411,108 trained. And no single parameter without training. It shows the efficiency of the model.

4.5 Compile and train the model

After developing the model, compile the model to check all types of syntax errors. Now train the model with the collected clean dataset by following thirty (30) epochs or repetitions to get better results.

4.6 Analysis of the model's performance

This step displays the outcome in 'Training and Validation Accuracy' and 'Training and Validation Loss.' It shows accuracy, validation accuracy, losses, and validation losses approximately 94%, 92%, 18%, and 22%, respectively, as look in Figure-3 and Figure-4. It is pretty explained here is no overfitting in the line graphs. All four curves show similarities for training and validation of an implemented model. It is a great performing model.

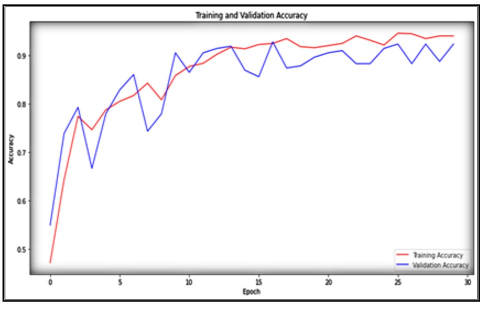

Fig. 3. Training and validation accuracy for the weather classification model.

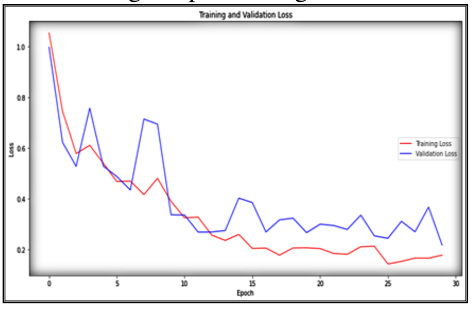

Fig. 4. Training and validation loss for the weather classification model.

4.7 Confusion matrix of the model's performance

This confusion matrix shows the results for total dataset images of 903 are [[57, 48, 73, 62], [37, 36, 42, 58], [54, 45, 47, 57], [51, 72, 66, 98]]), such as correct prediction for cloud, rain, shine, and sunrise are 57 out of 240, 36 out of 173, 47 out of 203 and finally 98 out of 287, respectively, look in Figure-5.

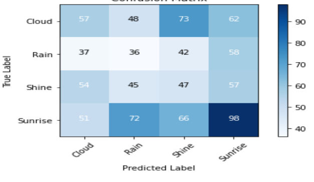

Fig. 5. Confusion matrix

5 Conclusion and future work

This work concludes the implemented weather classification model with its performance. After following all of the processing steps, it shows incredible results. It here implemented model trained through collected dataset from Kaggle.com website. This image dataset is a collection of clean data. It finds 94% training accuracy, 92% validation accuracy, 18% training losses, and 22% validation losses. As well as no overfitting or underfitting notice during the training and validating stages. So, it concludes that the model performance is highly acceptable.

In the future, this model will extend with more attributes like humidity, precipitation, air pressure, atmospheric pressure, temperature, solar radiations, Etc. and then integrate with an early warning system. It will provide the facility of finding natural disasters before time plus support to predict global warming in real-time.

References

1. M. Abrar, A. Sim, D. Shah, S. Khusro, A. Salam, *Weather Prediction using Classification*, Science International. **26**, 2217-2223, https://www.researchgate.net/publication/272482887_Weather_Prediction_using_Classification, (2014)
2. A. John Arnfield, *Climate classification*, Encyclopedia Britannica, https://www.britannica.com/topic/classification-1703397, (17 Mar. 2016)
3. J. Hidalgo, R. Jougla, *On the use of local weather types classification to improve climate understanding: An application on the urban climate of Toulouse*, PLoS ONE 13**(12)**: e0208138. https://doi.org/10.1371/journal.pone.0208138, (2018)
4. H. Beck, N. Zimmermann, T. McVicar, et al., *Present and future Köppen-Geiger climate classification maps at 1-km resolution*, Sci Data **5**, 180214. https://doi.org/10.1038/sdata.2018.214, (2018)
5. Armelle R. Remedio, C. Teichmann, et al. *Evaluation of New CORDEX Simulations Using an Updated Köppen–Trewartha Climate Classification,* Atmosphere 10, no. **11**: 726, https://doi.org/10.3390/atmos10110726, (2019)
6. CC. Lee, *The gridded weather typing classification version 2: A global-scale expansion*, Int J Climatol. 2020, **40**: 1178– 1196. https://doi.org/10.1002/joc.6263, (2020)
7. I. Gad, H. Doreswamy*, A comparative study of prediction and classification models on NCDC weather data*, International Journal of Computers and Applications. 1-12, 10.1080/1206212X.2020.1766769, (2020)
8. A. Gardner, I. Maclean, K. Gaston, *A new system to classify global climate zones based on plant physiology and using high temporal resolution climate data*, Journal of Biogeography. 47. 10.1111/jbi.13927, (2020)
9. K. Piotrowicz, D. Ciaranek*, A selection of weather type classification systems and examples of their application*. Theoretical and Applied Climatology. 140. 10.1007/s00704-020-03118-2, (2020)
10. Y. Lai, N. Tanaka, S. Im, K. Kuraji, et al. *Climate classification of Asian university forests under current and future climate*, Journal of Forest Research. **25**. 1-11. 10.1080/13416979.2020.1759898, (2020)
11. A. Belušić Vozila, M. Telišman Prtenjak, I. Güttler, *A Weather-Type Classification and Its Application to Near-Surface Wind Climate Change Projections over the Adriatic Region*, Atmosphere, 2021; 12**(8)**:948. https://doi.org/10.3390/atmos12080948, (2021)
12. D. Cui, S. Liang, D. Wang, *Observed and projected changes in global climate zones based on Köppen climate classification*, WIREs Climate Change, 12. 10.1002/wcc.701, (2021)
13. Ketkar N. Convolutional Neural Networks. In: Deep Learning with Python. Apress, Berkeley, CA. https://doi.org/10.1007/978-1-4842-2766-4_5, (2017)