Deep learning for satellite rainfall retrieval using Himawari-8 multiple spectral channels

Jenq-Dar Tsay 1*, Kevin Kao 1, Chun-Chieh Chao 1 and Yu-Cheng Chang 2

1 Meteorology Informatics Business Division, Government and Financial Department, International Integration System, Inc., New Taipei City, Taiwan; k890268@gmail.com (K.K); chunchieh0617@gmail.com (C.C.C)
2 Meteorological Satellite Center, Central Weather Bureau, Taipei City, Taiwan; luu@cwb.gov.tw
* Correspondence: jenqdar.tsay@iisigroup.com

Abstract: Rainfall retrieval using geostationary satellites provides critical means to the monitoring of extreme rainfall events. Using the relatively new Himawari 8 meteorological satellite with three times more channels than its predecessors, the deep learning framework of “convolutional autoencoder” (CAE) was applied to the extraction of cloud and precipitation features. The CAE method was incorporated into the Convolution Neural Network version of the PERSIANN precipitation retrieval that uses GOES satellites. By applying the CAE technique with the addition of Residual Blocks and other modifications of deep learning architecture, the presented derivation of PERSIANN operated at the Central Weather Bureau of Taiwan (referred to as PERSIANN-CWB) expands four extra convolution layers to fully use Himawari 8’s infrared and water vapor channels, while preventing degradation of accuracy caused by the deeper network. The development of PERSIANN-CWB was trained over Taiwan for its diverse weather systems and localized rainfall features, and the evaluation reveals an overall improvement from its CNN counterpart and superior performance over all other rainfall retrievals analyzed. Limitation of this model was found in the derivation of typhoon rainfall, an area requiring further research.

Keywords: satellite rainfall retrieval; deep learning; satellite meteorology

1. Introduction

Satellite-based precipitation estimates consist of two main types: low earth orbit satellites (LEO) with both passive and active advanced instruments, and geostationary satellites (GEO) that provide a broader coverage but has limited capability to retrieve information below the cloud top. Most satellite precipitation retrieval methods adopt algorithms or statistic models that describe the relationship between microwave sounding (MV) and infrared (IR). Interpolation methods were applied to fill the observational gaps left by LEO observation [1-3]. Newer GEO like the Japan Meteorological Agency (JMA) Himawari 8 meteorological satellite (H8 hereafter) function with sixteen multiple-band measurements of visible (VIS) and infrared (IR), providing three times more channels than its predecessors. Those additional channels enable H8 to provide information beyond cloud top and, in doing so, advance the accuracy of GEO rainfall estimation.

Various approaches have been used for building satellite rainfall retrieval models and most of which are based on regression-based methods [2,4] and neural network-based methods. Neural network has been used to retrieve rainfall from GEO infrared channel such as Precipitation Estimations from Remotely Sensed Information Using Artificial Neural Networks (PERSIANN) [5] and PERSIANN Cloud Classification System (PERSIANN-CCS; [6]). Recent advances in machine learning (or deep learning) significantly improve the capability of neural network. A number of studies have shown the advantage of applying the Convolution Neural Network (CNN) on image-based GEO-synthetic radar reflectivity [7,8]. Sadeghi et. al. [9] proposed a CNN version of PERSIANN (PERSIANN-CNN) that applies convolutional autoencoder to retrieve rainfall from GOES’s 10.7 and 6.7 µm channels. A convolutional filter extracts some part of the content from an
input signal or image, while convolutional autoencoder automates the process allowing the model to learn the optimal filters that minimize the reconstruction error.

Given that PERSIANN-CNN has not been applied in the western Pacific-East Asia weather systems, we conducted the ensuing analysis with PERSIANN-CNN to retrieve rainfall from H8’s IR (10.4 µm) and WV (6.9 µm) channels by training it with Taiwan Central Weather Bureau (CWB)’s radar-based Quantitative Precipitation Estimation-Quantitative Precipitation Estimation-Segregation Using Multiple Sensors (QPESUMS) [10]. Taiwan is a subtropical island located between Japan and Philippines and it exhibits a broad range of extreme weather systems both local and synoptic, making it ideal for testing the PERSIANN-CNN rainfall retrieval using H8. This paper describes the improvement steps taken to enhance the performance of PERSIANN-CNN based on the hypothesis that the accuracy of deep neural network can be further improved by two methods: 1) adding the Residual Blocks (ResBlock) [11] to allow information extracted from early layers to be used in the later layers, 2) conducting “feature engineering” by analyzing H8’s seven IR, three WV channels, and difference between channels to estimate physically sound properties such as cloud phase, cloud type, and cloud height in order to retrieve rainfall. The proposed new deep learning model is named PERSIANN-CWB and details of which are described in section 3.

This study is arranged as follows: Radar precipitation (QPESUMS), satellite data, satellite rainfall products, data period and area of rainfall estimation are described in Section 2. The modification of neural network and performance evaluation are presented in Section 3. Results of model prediction and comparison with other satellite rainfall products are presented in Section 4. Some concluding remark are offered in Section 5.

2. Data

2.1. Himawari 8 multiple-spectra channels

Japan’s H8 satellite provides high spatial-temporal resolution multiple channel data over the East Asia-Pacific region since 2014. H8 carries the Advanced Himawari Imager (AHI) that provides sixteen channels across the VIS, IR, and WV bands. It provides full-disk observations every 10 minutes at variable spatial resolutions ranging from 0.5 km to 2 km depending on the channel. VIS channels were not used in this study because of the lack of information during the nighttime.

2.2. Training and comparative datasets

The CWB’s QPESUMS precipitation estimation was developed by NOAA’s National Severe Storm Laboratory (NSSL) and modified by CWB for Taiwan and the East Asian monsoons, with a resolution of 1 km every 10 minutes (https://qpeplus.cwb.gov.tw/) [11]. Precipitation of QPESUMS was used here as the ground-truth rainfall observation for model training. Three operational rainfall datasets were included for the evaluation of the deep learning model: the CWB modification of NOAA’s STAR Satellite Rainfall Estimates Self-Calibrating Multivariate Precipitation Retrieval (CWB SCaMPR) [1], NOAA’s CMORPH real time data (CMORPH RT) [2], and NASA’s Integrated Multi-satellitE Retrievals of GPM (IMERG) V06 data [3]. The SCaMPR algorithm [1], which uses IR data and calibrates with MW rain rate, is currently operated by CWB as a standard satellite product to monitor heavy rainfall events, with a 2 km and 30 minutes resolution. The IMERG-Final V6 (IMERG V06 here after) has a spatial and temporal resolution of 10 km and 30 minutes.

2.3. Study area and analysis period

Given the validation data of QPESUMS, we focused on the area around Taiwan within the range of its weather radar network, which covers 119˚E-122.98˚E and 21˚N-25.98˚N. We analyzed the summer rainy seasons of July-August from 2016 to 2018 and July 2019 as the training data and used August 2019 for the leave-one-out verification to evaluate the performance of the presented deep learning model.

3. Methodology
The PERSIANN-CWB extended from the framework of PERSIANN-CNN with the addition of Residual Blocks [11] to overcome the accuracy degradation problem while enabling a “deeper layer” learning; this was done by expanding four extra convolution belts to handle the input data of H8. The procedures of this modification are explained as follows.

3.1. Feature engineering

“Feature engineering” is one of the methods to improve the performance of machine learning models. In terms of satellite rainfall retrieval, the brightness temperature differences (BTD) between channels contain physical meanings that can increase the “features” for machine learning, such as the indication of convection initialization. For example, Mecikalski et al. [12] used BTDs to infer physical characteristics of cloud, while So and Shih [13] used three BTDs (6.2–11.2 µm, 8.6–11.2 µm, 11.2–12.4 µm) and a difference among BTDs [ΔBTD = (8.6–11.2 µm) - (11.2–12.4 µm)] from AHI to classify precipitating clouds. We adopted a similar approach to extract cloud features from H8 for the purpose of training the model. Other cloud characteristics were retrieved by the Clouds from AVHRR Extended System (CLAVR-x) [14] by using Sun-synchronous orbiting satellites with NOAA’s Enterprise Cloud Algorithms. For H8, CLAVR-x can also use its AHI data to generate cloud mask, cloud type, cloud-top pressure, and cloud-top temperature [15]. Those data were used to train the convolutional autoencoder model used here.

3.2. Neural Network architecture

In satellite meteorology, it is common practice to obtain cloud characteristics by estimating the standard deviation or variation of each pixel against its surrounding pixels from satellite images. CNN uses the similar approach to extract features from satellite images. A convolution layer, the most importance part of CNN, contains a set of independent filters derived from arrays of random parameters that are adjusted during the training phase. The image features are mathematically extracted by a series of dot products between the filters and the images. Each set of the extracted features is then passed to the next computation of the neural network. The parameters of the whole neural network are adjusted and optimized by a back-propagation algorithm, which uses the chain rule of differentiation to pass error value back to each layer for further parameter optimization during the training process.

The main architecture of PERSIANN-CWB (Figure 1) is a convolutional auto-encoder that has the encoder and decoder parts, which are additions to its CNN counterpart. The four extra convolutional layers added in PERSIANN-CWB enable it to handle H8’s multi-channel data and other satellite-specific features (yellow shading parts in Figure 1). As shown in Table 1, 7 IRs, 3 WVIs, 7 BTDs, 1 ΔBTD, several cloud characteristics, and topography from H8 are classified into six groups based on the model tuning trials. These six groups of data were subject to six individual channels of convolutions layers. In Figure 1, the far-left columns of boxes are followed by a sequence of convolution and “upsampling layers” (or decoder layers) which increases image resolution to restore original image size. The process translates the detected patterns into a different, “learned” representation of those patterns to generate the output image.
Figure 1. Neural network diagram of the PERSIANN-CWB. Pink shading indicates Residual Block layers. Yellow shading indicates convolutional layers for input data except near inferred and water vapor channel.

Figure 2. Detail diagram of the ResBlock.

Table 1. Input data in six groups. C is acronym of “channel”.

<table>
<thead>
<tr>
<th>Group number</th>
<th>Group name</th>
<th>Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Near infrared</td>
<td>c07</td>
</tr>
<tr>
<td>2</td>
<td>Water vapor</td>
<td>c08, c09, c10</td>
</tr>
<tr>
<td>3</td>
<td>Infrared</td>
<td>c11, c12, c13, c14, c15, c16</td>
</tr>
<tr>
<td>4</td>
<td>Difference between channels</td>
<td>c08-c14, c11-c14, c14-c15, c11-c15, c11+c15-2*c14</td>
</tr>
<tr>
<td>5</td>
<td>Cloud characteristic</td>
<td>Cloud bottom height, cloud top height, cloud top pressure, cloud top temperature</td>
</tr>
<tr>
<td>6</td>
<td>Others</td>
<td>Topography</td>
</tr>
</tbody>
</table>

ResBlock consists of the convolutional network layers that are connected to a rectified linear unit (ReLU), while its upper layer information provides an effective transmission function. The ResBlock layers (Figure 2 and pink shading area in Figure 1) were added in each channel to enable the deeper neural network and this addition can minimize the introduction of further errors [11]. Each image’s
horizontal dimension becomes one quarter of their original dimension after going through the convolution layers. The Max Pooling layers used in PERSIANN-CNN are removed from PERSIANN-CWB because these layers lead to severe loss of information during the model trials. Features that are extracted by this encoder process are concatenated as a three-dimension array. In decoder section of PERSIANN-CWB, two Conv2DTranspose layers in PERSIANN-CNN are replaced with two cells, which is composed by ResBlock, bilinear upsampling, and convolution layers, to restore the horizontal dimension – a process that can produce better result. In the end, another ResBlock and a convolution layer with kernel size of 1 × 1 are added to predict the rainfall amount. We then applied the momentum method of the stochastic gradient descent as optimizer. The model was trained with the backpropagation algorithm using the PyTorch framework.

3.3. Evaluation metrics

To measure the similarity of images during the training, we calculated the root mean square error (RMSE) and correlation loss that ranges from 0 to 1. R-Squared ($R^2$) reveals the variability and difference between the predicted result and the observation. A value of $R^2$ closer to 1 is ideal as it explains how many percentages of variability of precipitation can be explained by model. RMSE measures the residual between predicted value and observation by summing up the square of individual difference and estimating an averaged error. A value closer to 0 is ideal. Threat score (TS) is commonly used in meteorology community to evaluate precipitation forecast. It measures the percentage of positive detection area over the sum of observation and false-alarm areas. A value closer to 1 is ideal.

4. Result

4.1. Overall performance

Hourly precipitation data derived by PERSIANN-CWB during August 2019 are compared with CWB QPESUMS and other sources of precipitation estimate. Figure 3a presents the evaluation outcome in terms of the box and whisker plot of $R^2$. It appears that PERSIANN-CWB has larger $R^2$ values (closer to 1) than other datasets. RMSE of PERSIANN-CWB is smaller than other datasets (Figure 3b). Threat score ($P \geq 0.5 \text{mm/hour}$) indicates that PERSIANN-CWB outperforms other datasets as well for it has the smallest inter-quartile range (Figure 3c). These evaluation metrics indicate that the deep learning-based PERSIANN-CWB offers an improvement in rainfall retrieval over its predecessor, PERSIANN-CNN.

![Figure 3](https://example.com/figure3.png)

*Figure 3.* Box and whisker plot of the (a) $R^2$, (b) RMSE, and (c) threat score ($P \geq 0.5 \text{ mm/hour}$) of hourly rainfall estimations on August 2019. PERSIANN-CWB and PERSIANN-CNN are shaded in light pink and yellow color, respectively.

To put PERSIANN-CWB’s rainfall retrieval into perspective, we plot in Figure 4 the horizontal distribution of precipitation in the afternoon of 29 August 2019. Both PERSIANN-CWB (Figure 4b)
and PERSIANN-CNN (Figure 4c) have the better depiction of rainfall distribution than other datasets, and PERSIANN-CWB apparently captures the rainfall magnitude more realistically. IMERG V06 (Figure 4g) only shows rainfall in the southwest coast of Taiwan and it is probably caused by the misguidances of the cloud top information in the IR image, which can be referred to Figure 4d. The rainfall depiction of CWB SCaMPR (Figure 4e) also is skewed by the cloud top shape in the IR image. CMORPH RT (Figure 4f) produced a very different rainfall estimation from others with questionable location and amount.

Figure 4. Comparison of hourly rainfall estimation of (a) CWB QPESUMS, (b) PERSIANN-CWB, (c) PERSIANN-CNN, (d) Himawari 8 infrared image (IR), (e) CWB SCaMPR, (f) NOAA CMORPH RT, and (g) NASA IMERG V06 for 07 UTC 29 August, 2019.

Another example shows the midnight rainfall along the southwest coastline of Taiwan (Figure 5). Midnight rainfall is frequent in summer as a result from the convergence between the larger-scale southwesterly flows and the land breeze. Figure 5 shows that PERSIANN-CWB outperforms others in this type of rainfall events. In this case, both CWB SCaMPR (Figure 5e) and IMERG V06 (Figure 5g) are dictated by the shape of cloud top appeared on IR image (Figure 5d). CMORPH RT produced unrealistic inland rainfall (Figure 5e). Both PERSIANN-CWB (Figure 5b) and PERSIANN-CNN (Figure 5c) depict the correct shape of the rainband while PERSIANN-CWB produced the closest rainfall amount and distribution to the observation (CWB QPESUMS).
4.2. Evaluation of daily rainfall estimation

To evaluate the daily precipitation pattern, we accumulated the hourly estimation of CWB QPESUMS, PERSIANN-CWB, PERSIANN-CNN, CWB SCaMPR, CMORPH RT, and IMERG V06 during August 2019. Figure 6a shows that the daily $R^2$ of each dataset against CWB QPESUMS from 1 to 31 August 2019 indicates the best performance being PERSIANN-CWB (red triangle), which reveals a marked improvement from PERSIANN-CNN (pink open square). The box and whisker plot of $R^2$ (Figure 6d) shows that PERSIANN-CWB consistently outperforms all other datasets with the smallest interquartile range (its median and mean of $R^2$ are higher than other datasets). We should note that the daily variation and amount of $R^2$ are not dependent on weather types.

Figure 5. Same as Figure 4 except for 16UTC August 12, 2019.
Figure 6. (a), (b), (c) are $R^2$, RMSE, and threat score ($P \geq 1$ mm/day) of daily precipitation on August 2019. Legend of symbols is on right top corner of (a). (d), (e), (f) are box and whisker plot of (a), (b), (c), respectively.

Figure 6b shows the RMSE of the six datasets against CWB QPESUMS. The RMSE of PERSIANN-CWB overall are smaller than other datasets. The box and whisker plot of RMSE (Figure 6e) shows PERSIANN-CWB has a smaller median, 75th percentile, and upper extreme than other datasets. Threat score calculated by areas of $P \geq 1$mm/day shows that PERSIANN-CNN and PERSIANN-CWB have a comparably better precipitation estimate than the other datasets (Figure 6c). The box and whisker plot of RMSE (Figure 6f) shows that both PERSIANN-CNN and PERSIANN-CWB have a higher median and smaller interquartile range than other datasets.

Taiwan’s summer rainfall has a predominant diurnal/afternoon signal [16]. To examine the horizontal distribution of the daily-accumulated rainfall from a diurnal event, we plot in Figure 7 the daily rainfall amounts on 11 August 2019 within Taiwan. It is apparent that both PERSIANN-CNN and PERSIANN-CWB portray the rainfall pattern and amount better than other datasets, especially in the central mountains. Moreover, PERSIANN-CWB appears to depict the typical pattern of the diurnal rainfall in Taiwan [16] better than PERSIANN-CNN, highlighting its improved performance based on the CNN framework and the presented deep learning modification.
Figure 7. Comparison of daily rainfall estimation of (a) CWB QPESUMS, (b) PERSIANN-CWB, (c) PERSIANN-CNN, (d) CWB SCaMPR, (e) NOAA CMORPH RT, and (f) NASA IMERG V06 for 11 August, 2019.

To examine different rainfall events, Figure 8 shows a 4 August 2019 case characterized with a cellular structure, which was caused by a synoptic-scale weather system east of Taiwan with both inland and offshore precipitation (Figure 8a). Both PERSIANN-CWB (Figure 8b) and PERSIANN-CNN (Figure 8c) capture the narrow band of precipitation in the west hill of the central mountains, but only PERSIANN-CWB depicts the precipitation cells embedded in the eastern Taiwan precipitation zone. However, PERSIANN-CNN appears to capture the eastern Taiwan rainband intensity better than PERSIANN-CWB. Regardless, the difference in performance between PERSIANN-CWB and other datasets is striking.

Figure 8. Same as Figure 7 except for August 4, 2019.
4.3. Typhoon rainfall

The improvement of PERSIANN-CWB is not without limit. One area of underperformance lies in the estimation of typhoon rainfall. As shown in Figure 9 for Typhoon Bailu on 04 UTC 24 August 2019, PERSIANN-CWB does not estimate the correct amount of precipitation along the east hill of the central mountains (and no other dataset could, either). This precipitation band along the eastern mountain range, which is the windward side relative to the typhoon, is primarily orographic rainfall. The typhoon rainbands were not properly produced by any dataset, either. Warm cloud or orographically induced precipitation is a known challenge for IR rainfall retrievals (Hong et al. 2004) and this appears to be hampering PERSIANN-CWB too. Since the presented model was trained for July and August only, and the typhoon season spans from August through October, more trials and developments are needed to establish a tailored training method for typhoon precipitation.

![Figure 9. Same as Figure 4 except for 04 UTC August 24, 2019. Eye of typhoon Bailu located in southeast Taiwan.](image)

5. Conclusion

Improving precipitation retrieval techniques from meteorological satellites has been an important task pursued by various meteorological agencies worldwide. The complexity of the problem lies in the indirect relationship between cloud top observation and ground precipitation. To provide the broad coverage and timely updates of precipitation estimate, the use of geostationary satellites’ IR observation in retrieving rainfall information is paramount. The emergence of deep learning technology in the past decade has provided a new opportunity to refine the rainfall retrieval algorithms. Given that newer geostationary satellites carry advanced instruments providing three times more IR and WV channels, the deep learning algorithms (especially the convolution neural network) have enabled us to retrieve more cloud characteristics from the multiple IR and WV spectra through time-efficient methods that extract features from the images and then solve the nonlinear problems.

By adding four more coevolution layers to extract all of H8’s IR and WV channels, the difference of channels, and cloud characteristics based upon PERSIANN-CNN, the improvement through PERSIANN-CWB is evident. Incorporated with the presented modification of neural network architecture and feature engineering of H8 satellite data, PERSIANN-CWB was able to produce a rainfall retrieval over the subtropical island of Taiwan with a strong locality of rainfall features.
Evaluation of PERSIANN-CWB based on hourly and daily precipitation in Taiwan shows consistent improvements over PERSIANN-CNN and a superior performance to SCaMPR, CMORPH RT, and IMERGE V06. Thus, PERSIANN-CWB may benefit aviation and maritime activities over the oceans, where weather radar coverage is lacking, in addition to providing reliable precipitation monitoring for weather and climate.

Author Contributions: Conceptualization, C.C.C, J.D.T, Y.C.C; Data Resources/Methodology, C.C.C, J.D.T; Methodology/Software/Writing, K.K; Formal analysis/Validation/ Visualization/Writing, J.D.T; All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Central Weather Bureau’s Maritime Disaster Prevention Information System Project.

Acknowledgments: CWB SCaMPR data were prepared and provided by Mr. Jian-Guo Wang. Mr. Chun-Jen Lin and Miss Shao-Ci (Candy) Jheng’s efforts in build training system and pre-trials are greatly appreciated.

References
