Creating synthetic meteorology satellite visible light images during night based on GAN method

CHENG Wencong

Beijing Aviation Meteorological Institute, Beijing 100085, China

Abstract. Meteorology satellite visible light images is critical for meteorology support and forecast. However, there is no such kind of data during night time. To overcome this, we propose a method based on deep learning to create synthetic satellite visible light images during night. Specifically, to produce more realistic products, we train a Generative Adversarial Networks (GAN) model to generate visible light images given the corresponding satellite infrared images and numerical weather prediction (NWP) products. To better model the nonlinear relationship from infrared data and NWP products to visible light images, we propose to use the channel-wise attention mechanics, e.g., SEBlock to quantitative weight the input channels. The experiments based on the ECMWF NWP products and FY-4A meteorology satellite visible light and infrared channels data show that the proposed methods can be effective to create realistic synthetic satellite visible light images during night.

Key words: Deep learning, Generative adversarial networks, meteorological satellite, atmospheric remote sensing, visible channel images

1 Introduction

With the rapid development of satellite remote sensing technology, meteorological satellites play one of the most important roles in atmospheric analysis and forecast. As a major source of atmospheric sensing, the data are used by forecasters, decision makers, and general public, providing continuous monitoring of clouds in all regions of the Earth. Meteorological satellite images characterize the distribution of clouds, so they can be used to track the evolution of the large-scale weather systems. With the help of the meteorological satellites, people can make more accurate weather prediction. The meteorological satellite data include images in both visible light and infrared channels. There are multiple visible light remote sensing channels for the most modern meteorological satellites, which can be merged to RGB true colors images. The true color visible light images are intuitive for users, which are rich in details and easy to be understand. Since the satellite visible light channels are not available during the night time, it is impossible to observe the clouds changing continuously by satellite visible light channels across day and night. Users need to switch the observing channels and change the analysis mode when the day and night shift, which means to use the visible light channels during day time and infrared channels during the nigh. In this paper, we investigate the problem of synthesize the meteorological satellite visible
light images during night time.

Deep learning have recently greatly improved in capability to adapt to data with complex spatial structures, owing particularly to the introduction of convolutional neural networks (CNN). Deep learning methods, which characterize the probability distribution of the training data, usually leads to synthesis of blurred images [2][3][4] by using the Euclidean distance. Typically a Euclidean distance like L2 distance is used as loss function with the assumption that the data is drawn from a Gaussian distribution. This can be problematic in the case of multimodal distributions. Recent advances in the field of image generation have been driven particularly by the invention of generative adversarial networks (GAN)[5]. These have achieved state of the art results producing very realistic images in an unsupervised setting. GAN based methods use adversarial training to learn to map a simple probability distribution, such like a group of independent standard normal variables, to the training data distribution. CGAN[6] is a relatively straightforward variant of the basic GAN framework, which can learn the conditional probability distribution to a given input.

The goal of this work is to create synthetic meteorological satellite visible channel imageries during night in a realistic way. In this paper, we propose to learn the non-linear mapping from meteorological infrared channels data and numerical weather prediction products to satellite visible channel imageries by introducing CGAN model. We use a channel-wise attention mechanics, e.g., Squeeze-and-Excitation Block(SEBlock)[7], at the front of the model to quantitative weight the input channels. The proposed method is evaluated on the ERA5 reanalysis data and FengYun-4A (FY-4A) geostationary meteorology satellite multi-channels date. Experimental results demonstrate that our method can effectively create realistic synthetic satellite visible imageries during night. As shown in Fig. 1, the first column is the cloud imageries sensed by FY-4A meteorological satellite infrared channels, which have data during day and night. The second column is the real imageries sensed by satellite visible channels. As the night approaching, firstly there is no data for the low and right part of product (shown in the first row), and there is totality blank but noise signals when the midnight time coming (shown in the second row). The third column is the result of the synthetic satellite visible channel imageries generated by our work corresponding to the first two columns. The illustration of this result show that the method proposed in this paper can handle two cases above and create realistic visible channel imageries sensed by meteorological satellite. To the best of our knowledge, this is the first application of the GAN framework in the field of synthetic meteorological satellite visible channels imageries generation.
Fig. 1. Infrared channels, real visible channels and synthetic visible channels imageries at the same times

2 Data

This work is conducted on FY-4A meteorological satellite data and ERA5 NWP (numerical weather prediction) reanalysis data. The data region we chosen covers the latitude from North 10° to North 50° and longitude from East 110° to 150°. The FY-4A satellite data can be downloaded directly from the website of the China National Satellite Meteorological Centre (NSMC) starting from April 2018. So we select the data from April 1, 2018 to July 20, 2018, which are about 2500 hourly data compose as the training dataset. And the hourly data from July 21 to 22, 2018 are chosen as the test dataset.

(1) FY-4A meteorology satellite data

The FengYun-4A (FY-4A) meteorological satellite, a second generation geostationary meteorological satellite of China, was launched on December 11, 2016[8]. It was fixed at a position of 99.5°E above the equator. The FY-4A satellite was equipped with four advanced optical instruments aboard, including an Advanced Geosynchronous Radiation Imager (AGRI), a Geostationary Interferometric Infrared Sounder (GIIRS), a Lightning Mapping Imager (LMI) and Solar X-EUV Imaging Telescope (SXEIT). The AGRI has 14 spectral bands from visible to infrared (0.45 -13.8μm) with high spatial (1 km for visible light channels, 2 km for near-infrared channels, and 4 km for remaining infrared channels) and temporal (full-disk images at the 15-min interval) resolutions. In this work, the synthetic products of the visible channels CH01, CH02 and CH03 are used as target products, and the Middle and long wave Infrared channels from CH07 to CH14 are used as parts of the model input data. The wavelengths, descriptions and usages of the FY-4A AGRI channels in this work are shown in Figure 2.

Tab 1. Channels of the FY-4A AGRI
| Channel ID | wave band (μm) | Description | usage |
|------------|----------------|-------------|-------|
| 01         | 0.45~0.49      | Vis channel corresponding to blue |        |
| 02         | 0.55~0.75      | Vis channel corresponding to green | Use as target data |
| 03         | 0.75~0.90      | Vis channel corresponding to red |        |
| 04         | 1.36~1.39      | Shortwave infrared channel, no data during night |        |
| 05         | 1.58~1.64      | Shortwave infrared channel, no data during night | Don’t use in this work |
| 06         | 2.1~2.35       | Shortwave infrared channel, no data during night |        |
| 07         | 3.5~4.0(High)  |                                              |        |
| 08         | 3.5~4.0(Low)   |                                              |        |
| 09         | 5.8~6.7        |                                              |        |
| 10         | 6.9~7.3        | Middle and long wave Infrared channels, existing data during night | Use as parts of model input data |
| 11         | 8.0~9.0        |                                              |        |
| 12         | 10.3~11.3      |                                              |        |
| 13         | 11.5~12.5      |                                              |        |
| 14         | 13.2~13.8      |                                              |        |

(2)ERA5 reanalysis data.

ERA5 is the 5th generation reanalysis dataset from the European Centre for Medium-Range Weather Forecasts(ECMWF), which is one of the most widely used NWP reanalysis data[9]. This dataset contains records from 1950 to near real-time. The most notable improvements of the ERA5 dataset from its predecessor are the finer spatial grid (31 km), the higher temporal resolution(hourly), the more number of vertical levels (137 levels), a new NWP model (IFS Cycle 41r2) and the increase of the amount of data assimilated. The ERA5 data can be downloaded from the European Centre for Medium-Range Weather Forecasts website. In this work, we use the hourly ERA5 reanalysis products with a spatial resolution of 0.25°×0.25° as parts of the model input data. The selected ERA5 meteorological elements and vertical pressure levels in this work are shown in table 3. There are 75 channels considering the elements and levels.

| Tab 1. Elements of the NWP re-analysis products as the model input |
|---------------------------------------------------------------|
| elements | unit | levels |
|----------|------|--------|
| Fraction of cloud cover | (0 - 1) | 100/200/300/ |
| U-component of wind | m/s | 400/500/600/ |
| V-component of wind | m/s | 700/800/850/ |
| Vorticity | s⁻¹ | 900/950/1000/ |
| Temperature | K | |
| Relative humidity | % | |
| Skin temperature | K | Single level |
3 Methods

The goal of this work is to use meteorological satellite infrared channels data and NWP products to generate corresponding visible channels imageries. Suppose we observe a spatial region represented by an $M \times N$ grid which consists of $M$ rows and $N$ columns. The visible channels imageries can be represented by a vector $S \in S^{M \times N}$, where $S$ denotes the domain of the visible channels data. $R_1, R_2, \ldots, R_c$ is a set of infrared channels data, and $N_1, N_2, \ldots, N_p$ is a set of NWP elements, $E$ denotes the conditional expectations. Then the synthetic visible channel imagery $\hat{S}$ to be generated:

$$\hat{S} = E[S | R_1, R_2, \ldots, R_c, N_1, N_2, \ldots, N_p]$$  \hspace{1cm} (1)

We train a Generative Adversarial Network (GAN) model to generate visible imageries given the corresponding satellite infrared imageries and numerical weather prediction products (NWP). GANs work by training two different networks: a generator network $G$ and a discriminator network $D$. $G$ generate the target samples as realistic as possible from the input data. $D$ is trained to estimates the probability that an input data is drawn from the distribution of the real data, that is, $D$ can classify an input sample as real or synthetic. Follow the GAN principle, both networks are trained simultaneously with $D$ trying to correctly discriminate between real and synthetic data, while $G$ is trying to produce realistic samples that will confuse $D$.

For the generator network $G$, we use a U-Net[10] neural network as the backbone structure, combined with a SEBlock as the frontier module. U-Net is an encoder-decoder with skip-connections between mirrored layers in the encoder and decoder stacks, which is widely used in image segmentation. Specifically, U-Net module add skip connections between each layer $i$ and layer $n-I$, where $n$ is the total number of layers. Each skip connection simply concatenates all channels at layer $i$ with those at layer $n-i$. There are 8 FY-4A infrared channels and 75 NWP products, so there are total 83 channels as the inputs data. Since the resolution of NWP products is lower than the satellite products, an upsample module is added to increase the NWP products resolution to the infrared products. To catch the contribution of the input channels, a Squeeze and extraction Block (SEBlock) is introduced after the input layer, with the goal of improving the quality of representations by explicitly modeling the interdependencies between the input channels. SEBlock module first performs a Squeeze operation to squeeze global spatial information into a channel descriptor. This is achieved by using global average pooling to generate channel-wise statistics. We can get 75 real numbers corresponding to the input features by this step. To make use of the information aggregated in the squeeze operation, an Excitation operation is performed, which aims to fully capture channel-wise dependencies by introducing a fully connected layer (FCLayer). The excitation
operator maps the input specific descriptor to a set of channel weights, and the input data are recalibrated by dot product operation. The U-Net structure is an encoder-decoder network with skip connections, which is widely used in image segmentation. Due to the NWP products have lower resolutions than satellite data, an upsample layer is added after the U-Net structure. Discriminator $D$ is a standard classification convolution network. By calculating the probability of the true satellite data of the input, $D$ can discern whether the input data is a real satellite data or a generated product. During the training process, $D$ tries to correctly distinguish between real and generated satellite data, while $G$ tries to generate the synthetic satellite data as realistic as possible so that $D$ cannot distinguish between them. In order to extract the mapping relationship between the NWP products $x$ and the satellite data $y$, the conditional GAN model is used as the basic structure of the discriminator $D$, that is $x$ and $y$ are used together as the input to $D$ (the discriminator of the basic GAN only uses $y$ as the input).

For $D$, we would like to find its parameters:

$$\text{argmax}_i \log (D(x, y)) + \log (1 - D(x, G(x, z)))$$

(2)

For $G$, we would like to optimize:

$$\text{argmax}_i \log (D(x, G(x, z)))$$

(3)

The loss function for $D$ is defined as:

$$L_D = L_{\text{bce}} (D(x, y), 1) + L_{\text{bce}} (D(x, G(x, z)), 0)$$

(4)

where:

$$L_{\text{bce}} (\hat{a}, a) = -\frac{1}{N} \sum_{i=1}^{N} (a_i \log \hat{a}_i + (1 - a_i) \log (1 - \hat{a}_i))$$

(5)

$N$ is the number of samples in the minibatch of the model input, $a \in \{0, 1\}$ represents the label of the input data ($0$ for generated and $1$ for the real satellite data), and $\hat{a} \in [0, 1]$ is the estimated label by the discriminator $D$. When the value is close to 0, $D$ assumes the input is a generated satellite data, and when the value is close to 1, $D$ assumes that the input is a real satellite data.

The generator $G$ and discriminator $D$ networks are described in detail in Fig.2 of the supporting information. To train the GAN model, we alternated between training the generator with one batch of input data and training the discriminator with two batches, in which one containing real samples and the other containing generated samples. The Adam[18] optimizer was used to train both the generator and the discriminator with a learning rate of 0.001. The mini-batch size is set to 8. We performed the training using 2 of NVIDIA GTX 1080ti graphics processing unit. We train the model for a total 300 epochs, and the full training required approximately 25 hours.
In the case of G, as mention in [11], we use a loss that includes an adversarial term and a reconstruction L1 error. It is defined as:

\[ L_G = \lambda_1 L_{bce}(D(x,G(x,z)),1) + \lambda_2 \| y - G(x,z) \| \]  \hspace{1cm} (6)

Where y is the corresponding ground truth satellite data.

3 Experimental Results

In order to observe and quantify the effect of the synthetic satellite data, the real and synthesized satellite visible images were compared using the mean absolute error (MAE) and root mean squared error (RMSE). In addition, the peak signal-to-noise ratio (PSNR) and structural similarity index measure (SSIM), which are commonly used in the domain of image reconstruction, was evaluated. For satellite images with the resolution of m×n, The formulas in this paper are shown as follow:

\[ MAE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} | I(i, j) - K(i, j) | \]  \hspace{1cm} (7)

\[ RMSE = \sqrt{\frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} (I(i, j) - K(i, j))^2} \]  \hspace{1cm} (8)

\[ PSNR = 20 \times \log_{10}\left( \frac{MAX_i}{\sqrt{MSE}} \right) \]  \hspace{1cm} (9)

\[ SSIM(x,y) = \frac{(2\mu_x\mu_y + c_1)(\sigma_{xy} + c_2)}{\mu_x^2 + \mu_y^2 + \sigma_x^2 + \sigma_y^2 + c_3} \]  \hspace{1cm} (10)

Where \( MSE \) in (9) is the mean squared error. \( MAX_i \) is the Maximum grayscale of the images, that is 255 in this work. In (10), \( \mu_x \) and \( \mu_y \) are mean value of x and y, \( \sigma_x \) and \( \sigma_y \) are standard deviation of x and y, \( \sigma_{xy} \) is the covariance between x and y, the positive constants \( c_1, c_2 \) and \( c_3 \) are used to avoid a null denominator.

We implemented the proposed networks with Pytorch 1.0.1 deep learning framework and trained
them using Intel Xeon 4116×2 CPU, 64GB RAM, and NVIDIA GTX 1080ti 11GB×2 GPU. The networks were trained using the Adam[12] optimizer with learning rate of 0.0002, $\beta_1 = 0.5$, $\beta_2 = 0.999$.

The training data were collected from April 1 to July 20, 2018. The total number of training samples was 1057 (only using the hourly data from 08 am to 17 pm in the daytime). The resolution of the NWP products input to the model is 160×160, and the resolution of the infrared channel products input to the model is 1000×1000. The generated satellite visible images resolution is 1000×1000.

Fig. 4 displays selected examples of generated satellite visible light images and corresponding real satellite images for different time, the first row in each column shows the real FY4A true color visible light images combined by 3 visible channels as RGB, the second row shows the synthetic satellite visible light images generated by this work, the third row shows the synthetic satellite visible light images generated by infrared channels only, the 4th row shows the synthetic satellite visible light images generated by NWP products only, and the 5th row shows one of the FY-4A infrared channel (CH12, 10.3~11.3 μm) images.

Column 1-5 demonstrate various cases where the time changes. The first column shows the samples at Beijing time 17:00 (that is 09:00 UTC, the following time is Beijing time as well) on July 21, 2018 before the sun sets. The second column shows the samples at 19:00 on July 21, 2018. As the time is close to night, the lower right part of the real visible light images has became a shadow space. The 3rd column shows the samples at 0:00 on July 22, 2018, while the visible light image has no data at all. Column 4 shows the samples at 05:00 on July 22, 2018. Column 5 shows the samples at 8:00 on July 22, 2018, when the sun has fully risen.

It is evident from Fig.4 that the proposed method can create realistic-looking satellite visible light images. The textures are similar between the real and generated samples in the day time. There is a tropical cyclone in the middle of the right part in each image, which is the No.10 typhoon “AMPIL” in 2018. It can be found that the visible light images obviously provide more detailed information of the typhoon, which is helpful for analyzing the characteristics and influence of the typhoon. Due to there is no satellite visible light data at night, we cannot conduct the continuous observation of the typhoon using the visible light channels, and need to switch to infrared channels in night time. The continuous typhoon observation can be maintained in visible light channels by using the proposed method. Given the the synthetic visible light images, it can be clearly observed that the typhoon "AMPIL" gradually moved to the northwest of china, and the influence range expanded widely during the night of July 21. By synthetic satellite visible light images, we can detailed observe and analyze the complex meteorological events such as typhoon during night time.
We can extract the weights in the SEBlock of the model while generating the images. By comparing the weights corresponding to each input channel, we can roughly understand the contribution of each
input channel to the final target product. The weights at 8:00 on July 21, 2018 after model inferring is shown in Table 3. It can be seen from the table that the infrared channel data have the greatest contribution to generate the synthetic visible light image during the night time. The reason may be that the resolution of the infrared channel data is the same as that of the visible light images, which is higher than NWP products’. So the infrared channels data can provide more detailed information. As the NWP products, the contribution of temperature is higher according to the weights. We think the reason is the brightness temperature observed by infrared channel of the satellite changes rapidly from day to night, and the NWP temperature element can be used as an important feature to adjust the generated target visible light images. We note that the convolution operation in the deep learning model has certain adjustment ability, so the channel weight extracted from the SEBlock module can only roughly reflect the contribution of each input channel to the target product, and it should not be considered as an quantitative metric to evaluate the importance of the input channels.

| 类别                      | 要素                        | 要素累计权重 |
|--------------------------|----------------------------|-------------|
| NWP products—multi levels| Fraction of cloud cover     | 0.11        |
|                          | U-component of wind         | 0.18        |
|                          | V-component of wind         | 0.19        |
|                          | Vorticity                   | 0.33        |
|                          | Temperature                 | 0.61        |
|                          | Relative humidity           | 0.15        |
| Satellite infrared channels data | CH07                      | 0.81        |
|                          | CH08                       | 0.6         |
|                          | CH09                       | 0.25        |
|                          | CH10                       | 0.76        |
|                          | CH11                       | 1           |
|                          | CH12                       | 1           |
|                          | CH13                       | 0.97        |
|                          | CH14                       | 0.49        |

Since there is no real satellite visible light data during night time, in order to quantitatively evaluate the effect of the synthetic visible light images during night, the optical flow method is used to extrapolate the forecast image at 18:00 from the satellite visible light images at 16:00 and 17:00 on July 21, 2018. At 18:00 there are no real visible channels data can be observed in the lower right 1/4 part (i.e. corresponding to latitude 40°N~50°N, longitude 20°~30°E) of the concerned region. So the lower right 1/4 part of the forecast image obtained by the optical flow at that time is used as the test benchmark. The real and synthetic images for various input data are shown in Figure 6.
Table 2 shows MAE (lower is better), RMSE (lower is better), PSNR (higher is better) and SSIM (higher is better) as input of infrared channels data, NWP products and infrared channels data combined with NWP products. MAE and RMSE is evaluated use the albedo value of visible light channels data, which ranging from 0 to 1.65. PSNR and SSIM are evaluated use raw images directly. It can be found that the synthetic images generated by combined data are quantitatively outperform the others, which shows the advantages of the proposed method.

The visual perception and the quantitative evaluation show that the proposed method can better generate the satellite visible channels images.

5 conclusion and future work

In this work, we have presented a method using GAN to create synthetic satellite visible light images during night time. We leverage the idea from taking the combination of satellite infrared channels data and the NWP products as the model input. The experiment results in this paper suggest that comparing to the model input only with satellite infrared channels data and only with NWP products, the combination one learns to synthesize more realistically looking images. For future work, we will investigate how to merge the geographic information and other data to improve the generated synthetic satellite visible light images, both in perceptual similarity and quantitative evaluation.

References
[1] Y. LeCun et al. Deep learning. Nature, 521(7553):436–444, 2015.
[2] D. Pathak, P. Krahenbuhl, J. Donahue, T. Darrell, and A. A. Efros. Context encoders: Feature learning by inpainting. In CVPR, 2016.
[3] R. Zhang, P. Isola, and A. A. Efros. Colorful image colorization. In ECCV, 2016.
[4] M. Mathieu, Camille Couprie, and Yann LeCun. Deep multi-scale video prediction beyond mean square error. arXiv preprint arXiv:1511.05440, 2015.
[5] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In Advances in Neural Information Processing Systems, pages 2672–2680, 2014.
[6] Isola P. , Zhu J. Y. , Zhou T. , et al. Image-to-Image Translation with Conditional Adversarial Networks. In CVPR 2017.
[7] J. Hu, L. Shen, S. Albanie, et al., Squeeze-and-Excitation Networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017.

[8] ECMWF, 2017. ERA5 data documentation. European Centre for Medium-range Weather Forecast (ECMWF). <https://software.ecmwf.int/wiki/display/CKB/ERA5+data+documentation>.

[9] O. Ronneberger, P. Fischer, T. Brox. U-Net: Convolutional Networks for Biomedical Image Segmentation. In: MICCAI 2015:2-3.

[10] D. Pathak, P. Krahenbuhl, J. Donahue, T. Darrell, and A. A. Efros. Context encoders: Feature learning by inpainting. In CVPR, 2016.

[11] D. Kingma and J. Ba. Adam: A method for stochastic optimization. ICLR, 2015.