Convolutional Auto-encoder Based Sky Image Prediction Model for Minutely Solar PV Power Forecasting

Hua Chai, Zhao Zhen, Kangping Li, Fei Wang
Department of Electrical Engineering
North China Electric Power University
Baoding 071003, China
feiwang@ncepu.edu.cn

Payman Dehghanian
Department of Electrical and Computer Engineering
George Washington University
Washington DC, USA
payman@gwu.edu

Miadreza Shafie-khah
School of Technology and Innovations
University of Vaasa
Vaasa 65200, Finland
mshafiek@univaasa.fi

João P. S. Catalão
Faculty of Engineering of the
University of Porto and INESC TEC
Porto4200-465, Portugal
catalao@fe.up.pt

2020 IEEE Industry Applications Society Annual Meeting | 978-1-7281-7192-0/20/$31.00 ©2020 IEEE | DOI: 10.1109/IAS44978.2020.9334923

***Abstract*—The precise minute time scale forecasting of an individual Photovoltaic power station output relies on accurate sky image prediction. To avoid the two deficiencies of traditional digital image processing technology (DIPT) in predicting sky images: relatively limited input spatiotemporal information and linear extrapolation of images, convolutional auto-encoder (CAE) based sky image prediction models are proposed according to the spatiotemporal feature extraction ability of 2D and 3D convolutional layers. To verify the effectiveness of the proposed models, two typical DIPT methods, including particle image velocimetry (PIV) and Fourier phase correlation theory (FPCT) are introduced to build the benchmark models. The results show that the proposed models outperform the benchmark models under different scenarios.**

Keywords—Solar PV power forecasting, minute time scale, sky image, convolutional auto-encoder, spatiotemporal feature

I. INTRODUCTION

With the increase of solar PV integrated capacity [1], the difficulty of stable operation of power system is also increasing. Influenced by weather conditions and meteorological factors, solar PV power has significant uncertainty [2]. In order to reduce the difficulty of the operation, the key lies in weakening the uncertainty of solar PV power, one kind of effective methods is solar PV power forecasting [3].

Output of a single photovoltaic power plant is mainly influenced by the surface irradiance [4], while the surface irradiance is dominantly impacted by the clouds with complex distribution over the plant [5]. When the clouds change drastically on the minute time scale [6], the surface irradiance will show significant nonlinear fluctuation. The larger the capacity of a single station, the more serious the adverse effect of this fluctuation on power grids. Accurate minute time scale solar PV power forecasting is propitious to the stability of power system operation, the consumption of solar PV power and electricity market operation. Such methods use ground-based sky images commonly and are usually carried out in two steps: the first step is cloud distribution prediction, which means sky image prediction; the second step is to establish a mapping relationship from cloud distribution to the surface irradiance, then to the output. This paper only focuses on the first step.

At present, the traditional sky image prediction research adopts digital image processing technology (DIPT). The research content is to use two adjacent images with high image resolution to calculate the cloud displacement vector, and then use the cloud displacement vector to linearly extrapolate the current image to get a prediction image. The above three images get the same time resolution. The research can be divided into two categories: the first is based on image gray information and the second is based on image Fourier frequency domain. The first type of methods include scale invariant feature transform [7], optical flow [8] and particle image velocimetry (PIV) [9], among which PIV is widely used; however, the stability and accuracy of these methods are poor to some extent [10]. The second type of methods can describe the difference between images by using mathematical expressions in less computing resources and shorter computing time [11], and Fourier phase correlation theory (FPCT) is widely used. However, the DIPT based research needs an idealized hypothesis: the cloud distribution in two adjacent images are same on the minute time scale, which leads to two problems affecting the prediction performance: one is the relatively limited length of the input image sequence, which means the input spatiotemporal information may be not rich enough; the other is the idealized linear assumption, which means the prediction error could be introduced into the models from the beginning.

In addition to DIPT, sky image prediction can also draw lessons from the video prediction models based on deep learning. Video prediction models are usually complex and deep convolutional auto-encoders (CAEs) [12-14] in which convolutions include 2D convolution and 3D convolution. Convolutional neural network is a deep learning model widely used in the field of computer science [15]. An auto-encoder is an unsupervised deep learning model. For video prediction models, their input and output are image sequences and can be flexibly set lengths according to the need, which means that models can easily input more than two images. Thus, more abundant spatiotemporal information can be input and then prediction performance of models can be improved. Moreover, such models can fit the complex nonlinear relationship between input and output automatically [16], which means spatiotemporal features can be effectively learned without idealization, thus avoiding the introduction of correlative errors.

In summary, the first kind of models, CAE based sky image prediction models, are first established, which have 2D

and 3D forms. As controls, then the second kind of models, DIPT based sky image prediction models, are established using PIV and FPCT. Third, 5 different prediction scales are considered on the practical dataset, and the performance of CAEs and DIPT are compared under each scenario. The results under 5 scenarios indicate the performance of CAEs is superior to that of DIPT and the best model under each scenario is given.

In sum, the main contributions in this paper include:

(1) 2D-CAE based and 3D-CAE based sky image prediction models are proposed to avoid the two deficiencies from DIPT based sky image prediction models.

(2) The comparison between CAE based methods and two widely used DIPT methods (PIV and FPCT) is simulated to verify the superiority of the proposed models under 5 different prediction scales.

The rest of this paper is organized as follows. Section II introduces the above two kinds of models. Section III shows data processing and assessment metrics. Section IV presents the simulation design, results and discussion. Finally, Section V highlights the concluding remarks.

II. Methodology

A. Convolutional Auto-encoder (CAE) Based Sky Image Prediction Models

The proposed CAEs consist of convolutional layers and transposed convolutional layers. The two types of layers are the same in the calculation theory, the main difference between them is generally that a convolutional layer generates a down-sampled feature map while a transposed convolutional layer generates an up-sampled one [17]. Their equations are described as follows:

$$\mathbf{Y}_j = f(\sum_i \mathbf{X}_i * \mathbf{W}_{ij} + \mathbf{B}_j) \tag{1}$$

where $\mathbf{W}_{ij}$ denotes the i-th weight kernel in the j-th layer, $\mathbf{X}_i$ represents the input corresponding to $\mathbf{W}_{ij}$, $\mathbf{B}_j$ is the bias of the j-th layer, $\mathbf{Y}_j$ is the output feature map of the j-th layer and $f(\cdot)$ is an activation function.

Convolution operations including 2D convolution and 3D convolution is depicted as Fig. 1. A 2D convolution kernel has two directions to move, so 2D convolution applied on multiple frames stacked together generates a feature map. A 3D convolution kernel has three directions to move, so 3D convolution applied on multiple frames generates an output volume [18]. Under the same conditions, 3D convolution is finer than 2D convolution in terms of spatiotemporal feature extraction, but which means longer computation time.

For the structure of an auto-encoder, there are two kinds of descriptions: one contains an input layer, hidden layers and an output layer, the other contains an encoder, a bottleneck layer and a decoder. For the CAEs proposed in this paper, hidden layers consists of convolutional layers or transposed convolutional layers, and an output layer is a convolutional layer. The output of each layer is successively processed by batch normalization (BN) [19] and an activation function named as LeakyReLU [20]. A CAE with three hidden layers is shown in Fig. 2, whose structure conforms to the remaining CAEs in this paper. In Fig. 2, the encoder contains an input layer and a convolutional layer, the bottleneck layer is a convolutional layer, and the decoder contains a transposed convolutional layer and an output layer.

In CAE based sky image prediction models, an input image sequence information is compressed when passing through an encoder and a bottleneck layer successively. Then the condensed information is used by a decoder to predict a future frame. If t_0 is the time point to implement image prediction and n is the length of an input image sequence, the type of models can be described as follows:

$$SI_{t_0+t_{out},\,CAE} = f_{CAE}([SI_{t_0-(n-1)\cdot t_{in}}, \cdots, SI_{t_0-t_{in}}, SI_{t_0}]) \tag{2}$$

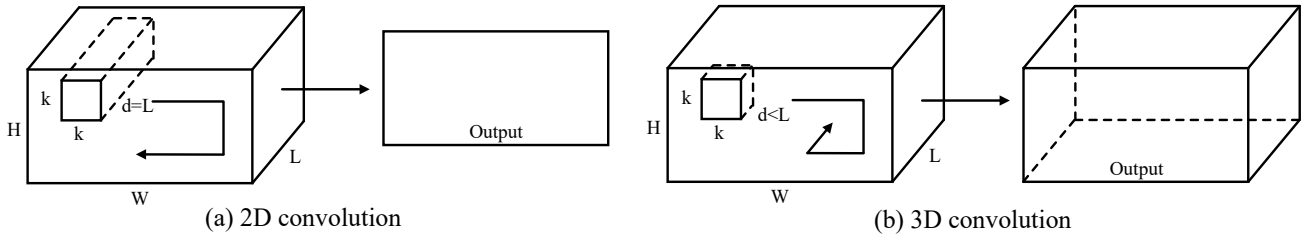
where t_{out} is the time resolution of the predicted sky image; $SI_{t_0+t_{out},\,CAE}$ represents the predicted sky image generated by a CAE; $f_{CAE}(\cdot)$ represents a CAE based sky image prediction model; SI_{t_0} is the sky image at t_0; t_{in} is the time resolution of input images.

B. Digital Image Processing Technology (DIPT) Based Sky Image Prediction Models

PIV first divides both input images into small block regions, and then computes the cloud displacement vector by matching these blocks [9]. For FPCT, it uses Fast Fourier Transform to realize the interconversion between image space domain and frequency domain [11]. Firstly, Fourier Transform of two input images is calculated and then used to calculate the cross-power spectrum. Finally, the cloud displacement vector is obtained by the inverse Fourier Transform of the cross power spectrum. PIV and FPCT are used to construct two determined models which generate a predictive image by linear extrapolation of the image at t_0. The type of models can be described as follows:

$$SI_{t_0+t_{out},\,DIPT} = f_{DIPT}([SI_{t_0-t_{in}}, SI_{t_0}]) \tag{3}$$

where $SI_{t_0+t_{out},\,DIPT}$ is the predicted sky image generated by a DIPT based method; $f_{DIPT}(\cdot)$ is a DIPT based sky image prediction model; besides, $n=2$, $t_{out}=t_{in}$.



(a) 2D convolution (b) 3D convolution

H: Height, d: Height of a kernel, L: Height of input frames, W: Width, k: Side length of a kernel

Fig. 1. Convolution operation.

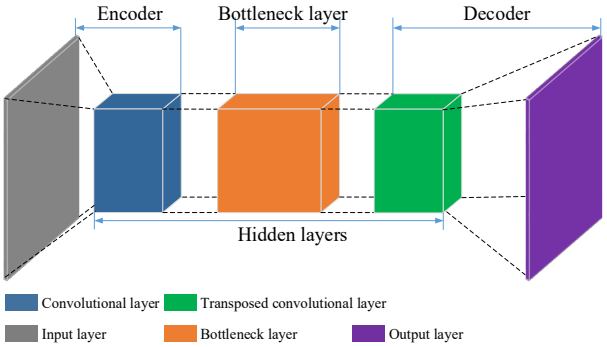


Fig. 2. The architecture of a convolutional auto-encoder.

III. Data Processing and Assessment Metrics

A. Data Processing

The sky images used in this paper are from the PV power station in Alamosa, Colorado, which are available on the National Oceanic and Atmospheric Administration (NOAA) website [21]. Their UTC time span is from May 22 2015 to May 31 2015. For each image, the time resolution is 0.5 min and the size is 288×352.

After data cleansing, the processing of an image is shown as Fig. 3, with the image resolution marked below each image. A sky image has a corresponding cloud analyzed image, as shown in Fig. 3(d) and (a), respectively. The white region in the cloud analyzed image only corresponds to the sky, and it can be used to remove the non-sky noise information in the sky image, thus obtaining the sky information. Specifically, first, the rectangular picture having the white area is obtained as Fig. 3(b) and the clipping coordinates are retained. Second, Fig. 3(c) is obtained by binarizing the pixel values of the Fig. 3(b). Thirdly, Fig. 3(e) is obtained by using the clipping coordinate to cut the Fig. 3(d). Fourthly, Fig. 3(f) is obtained by pixel-wise calculation between Fig. 3(c) and Fig. 3(e). Fifthly, Fig. 3(f) is transformed into the grayscale one as Fig. 3(g). Sixthly, Fig. 3(h) is obtained by downsampling Fig. 3(g). It should be noted that limited to the hardware condition, and in order to reduce the training difficulty of CAE based sky image prediction models, the sky images used are grayscale images with resolution of 32×32. Finally, 16456 image sequences are obtained in the dataset. Each image sequence contains 20 consecutive images, the first 10 are used to construct input and the last 10 are used to construct output. The dataset is divided according to the time order, the training set accounts for 80%, the validation set accounts for 10%, and the test set accounts for 10%. For CAE based sky image prediction models, the training set is used for model training, the validation set is used for model selection, and the test set is used for model test. DIPT based sky image prediction models only uses the test set.

B. Assessment metrics

Structural similarity (SSIM) and mean squared error (MSE) are introduced to evaluate the performances of sky image prediction models. SSIM is used to measure the similarity of structural information between two images, whose value range is [0, 1] [22]; while MSE is used to measure the similarity of gray values between two images. The larger the SSIM of two images is, the more similar their structures are; the smaller the MSE of two images is, the more similar values of their pixels are.

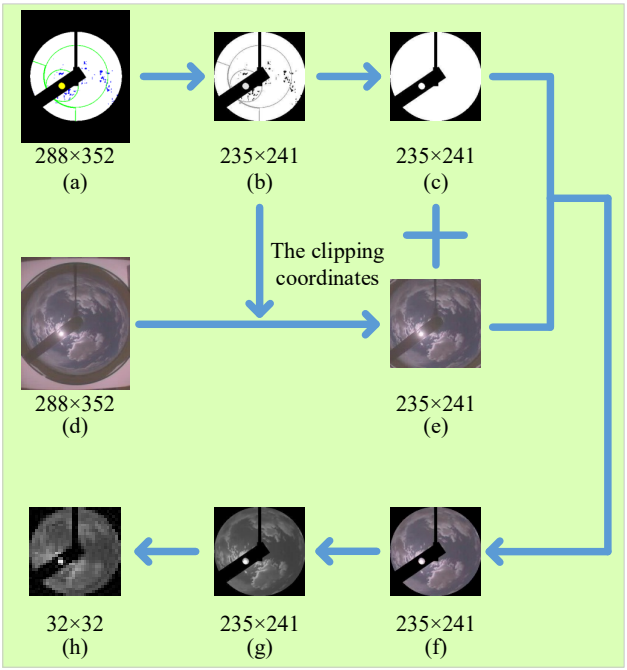


Fig. 3. Sky image processing.

If x and y are two images with same resolution of $M \times N$, their SSIM is the product of luminance comparison $l(x,y)$, the contrast comparison $c(x,y)$ and the structure comparison $s(x,y)$, the three comparisons are expressed as Equations (4)-(6).

$$l(x,y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \tag{4}$$

$$c(x,y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \tag{5}$$

$$s(x,y) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3} \tag{6}$$

where μ_x and μ_y respectively represent the mean value of all the pixels in x and y, σ_x, and σ_y respectively represent the standard deviation, σ_{xy} is covariance between x and y. C_1, C_2 and C_3 are three constants to avoid denominator very close to zero, and they can be respectively described as follows: $C_1 = (K_1 L)^2$, $C_2 = (K_2 L)^2$ and $C_3 = C_2/2$. In general, $K_1 = 0.01$, $K_2 = 0.03$ and $L = 255$. Finally, SSIM is described below:

$$SSIM(x,y) = l(x,y) \cdot c(x,y) \cdot s(x,y) \tag{7}$$

For x and y, MSE is described as follows:

$$MSE(x,y) = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}[x(i,j) - y(i,j)]^2 \tag{8}$$

where $x(i,j)$ and $y(i,j)$ represent the pixel values of coordinate (i,j) in x and y respectively.

IV. Simulation

A. Simulation Design

Two types of models in this paper predict an image. As can be seen from Table I, five prediction scales are set in this paper, each corresponding to one time resolution of the predicted images. The input and output schematics of the two types of models are shown in Fig. 4. The predicted images in the five situations are respectively filled with 5 colors. As shown in Fig. 4(a), 10 consecutive images are fed into the first type of model in each situation. In each situation of Fig. 4(b), the sky image at t_0 and the sky image at $t_0 - t_{in}$ is input into the second type of model, and the latter is the same filling color as the sky image at $t_0 + t_{out}$.

In this paper, the four different structures are set for 2D CAEs and 3D CAEs, respectively, as shown in Fig. 5. The number of hidden layers (NHL) of each structure is 3, 5, 7 and 9 respectively. Since Adam optimizer performing a random gradient descent is used during the training process [23], each evaluation of a CAE is slightly different. To mitigate this randomness, each CAE is evaluated 10 times in each situation, the median of 10 evaluation values is taken as the performance evaluation value. Furthermore, each PIV based or FPCT based sky image prediction model is evaluated only once in each situation due to its determination.

B. Results and Discussion

The comparison results of SSIM and MSE from the two types of models are shown in Fig. 6 and Fig.7, respectively. As the time resolution of the output image decreases, the spatiotemporal correlation between input and output also decreases, which makes the prediction more difficult and finally leads to the obvious deterioration of the performances. Besides, although the optimal values in each scenario are always obtained by a CAE, the optimal model in different scenarios is not only one structure. On the one hand, this indicates the superiority of the predictive performance of the first type models; on the other hand, this also shows that CAE based models need to compare the performance of different structures to find better models, which is exactly the inherent disadvantage of such models.

TABLE I. Input and Output Settings of the Two Types of Models

Scenario Number	CAE based models (n=10)		DIPT based models (n=2)	
	t_{in}/min	t_{out}/min	t_{in}/min	t_{out}/min
1	0.5	0.5	0.5	0.5
2	0.5	1	1	1
3	0.5	2	2	2
4	0.5	3	3	3
5	0.5	4	4	4

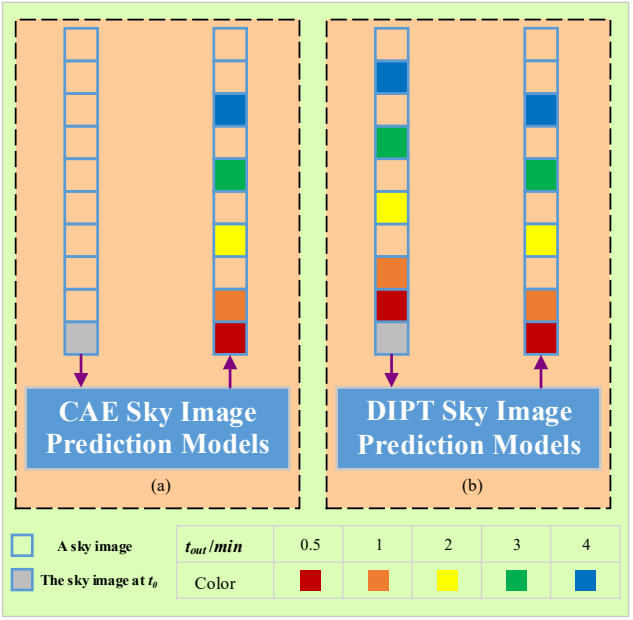


Fig. 4. Input and output of the two types of models.

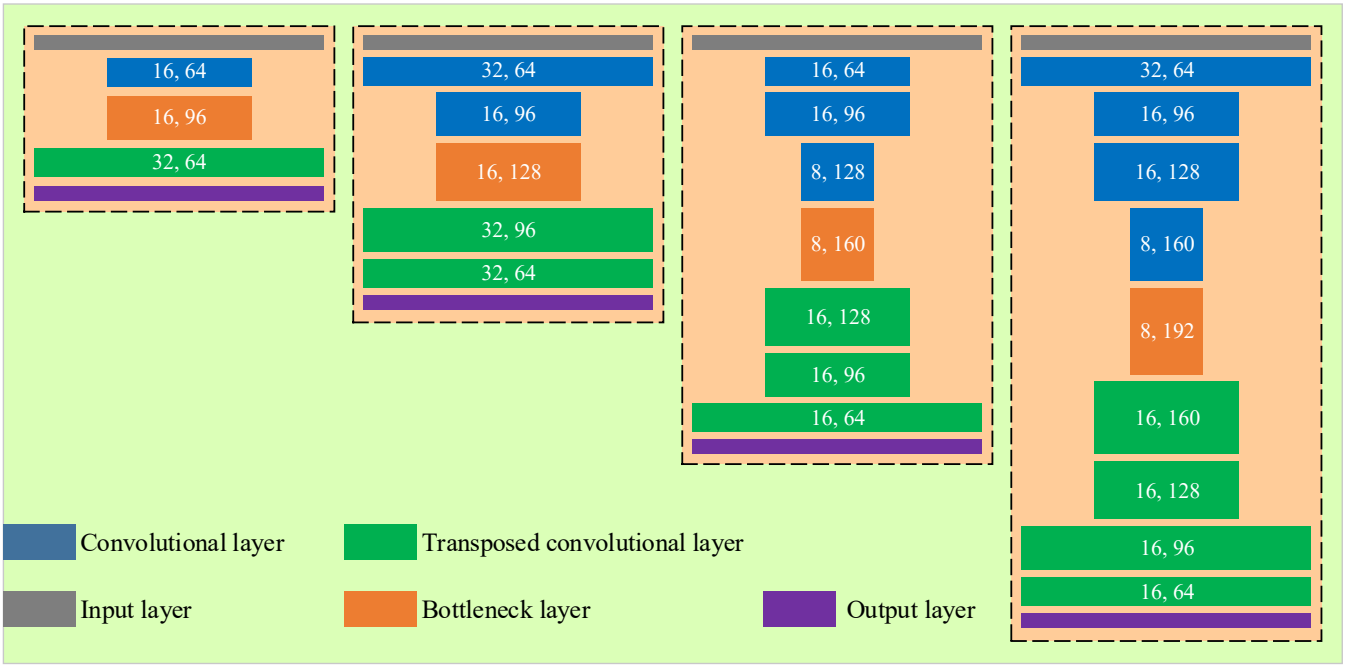


Fig. 5. Four architectures of convolutional auto-encoders.

The mean results of SSIM and MSE from the two types of models are shown in Table 2 and Table 3, respectively, where bold numbers represent the optimal values in each situation. When $t_{out}=1$ min in Table II and $t_{out}=3$ min in Table III, performance evaluations of 2D CAEs is smaller than but very close to the optimal values generated by 3D CAEs. What's more, as shown in Fig. 6 and Fig.7, in different structures under different situations, the best performance of 2D CAEs are basically better than that of 3D CAEs. The above two results show that in each situation, 2D convolution is more superior in sky image prediction compared with the 3D convolution, although 3D convolution has finer extraction of spatiotemporal features than 2D convolution.

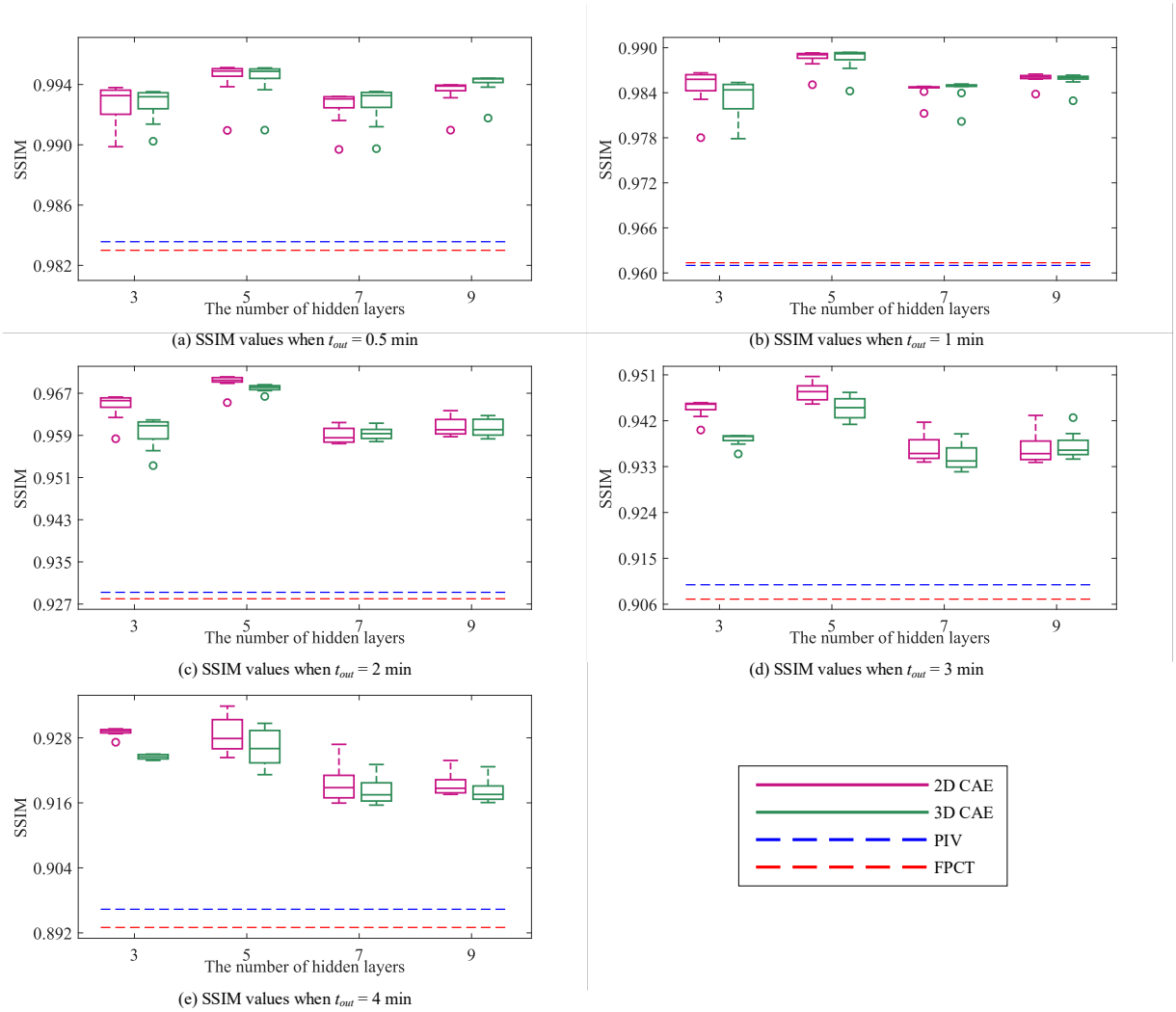
As depicted in Table II and Table III, the best models are 2D CAEs containing 5 and 3 hidden layers respectively when $t_{out}=0.5$ min and $t_{out}=4$ min, but the optimal model cannot be judged directly from the values of SSIM and MSE in the remaining three situations. However, the relatively good models in the remaining three situations can be obtained through details in Fig. 6 and Fig.7. In the remaining three situations, for 2D CAEs containing 5 hidden layers, their SSIM values are or very close to the best but their MSE values are relatively far from the best. While for 2D CAEs containing 3 hidden layers, their SSIM values very close to the best and their MSE values are the best. Therefore, the 2D CAEs with 3 hidden layers are chosen as the best models in these three situations. Finally, the best model in each situation is obtained, with a view to providing reference for further research.

TABLE II. MEAN SSIM VALUES OF TWO TYPES OF MODELS

t_{out}/min	Best 2D CAEs		Best 3D CAEs		PIV	FPCT
	Values	*NHL*	*Values*	*NHL*	*Values*	*Values*
0.5	**0.9953**	5	**0.9953**	5	0.9840	0.9834
1	0.9889	5	**0.9890**	5	0.9608	0.9612
2	**0.9692**	5	0.9678	5	0.9289	0.9277
3	**0.9475**	5	0.9444	5	0.9096	0.9068
4	**0.9291**	3	0.9259	5	0.8962	0.8929

TABLE III. MEAN MSE VALUES OF TWO TYPES OF MODELS

t_{out}/min	Best 2D CAEs		Best 3D CAEs		PIV	FPCT
	Values	*NHL*	*Values*	*NHL*	*Values*	*Values*
0.5	**5.12**	5	5.22	5	13.06	15.72
1	**14.14**	3	15.24	3	33.07	37.45
2	**11.84**	3	12.12	3	67.62	75.47
3	15.90	3	**15.88**	3	96.77	106.18
4	**15.07**	3	15.17	3	123.84	132.96



(a) SSIM values when $t_{out}=0.5$ min

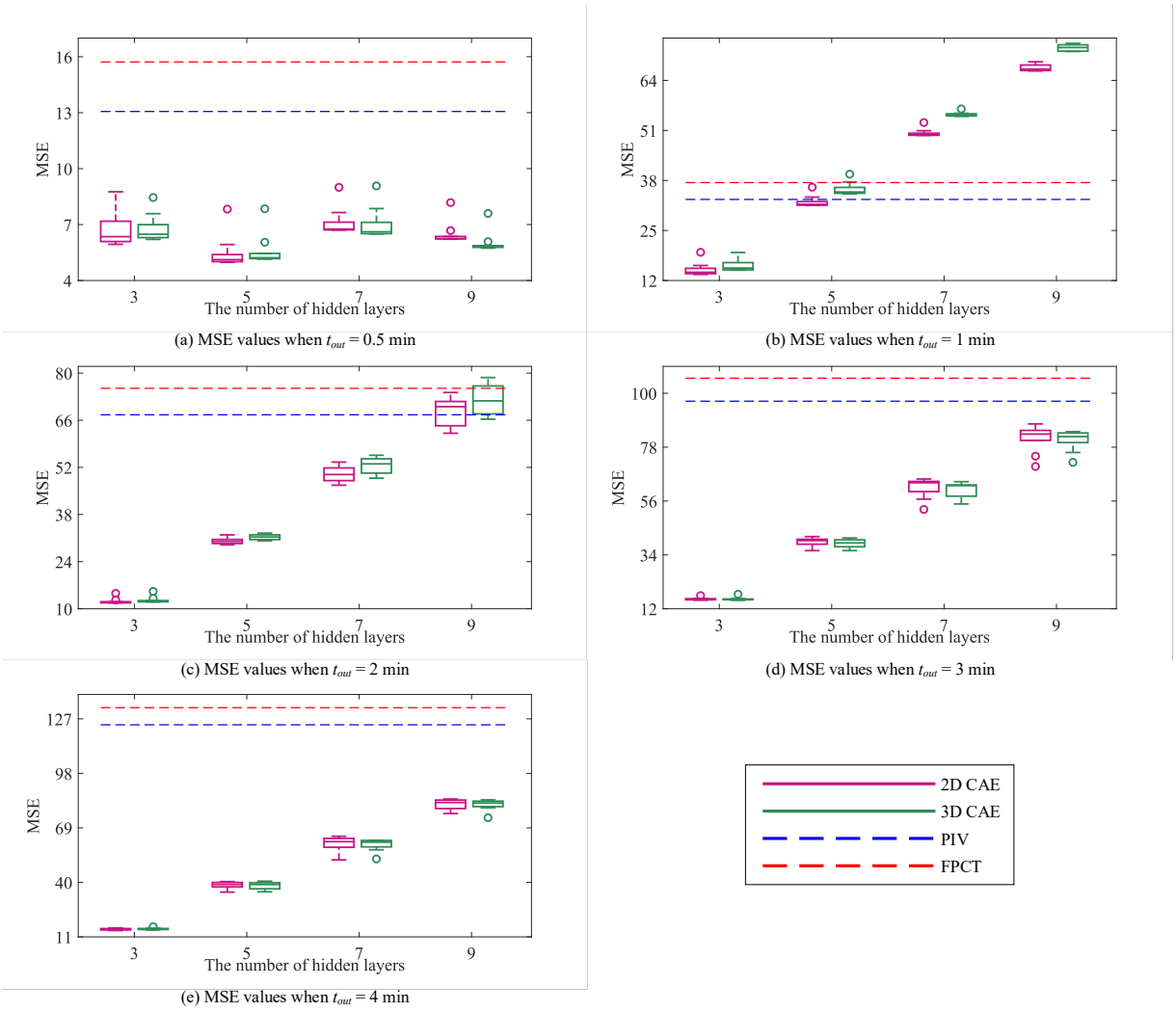
(b) SSIM values when $t_{out}=1$ min

(c) SSIM values when $t_{out}=2$ min

(d) SSIM values when $t_{out}=3$ min

(e) SSIM values when $t_{out}=4$ min

Fig. 6. SSIM comparison of the two types of models.



(a) MSE values when t_{out} = 0.5 min

(b) MSE values when t_{out} = 1 min

(c) MSE values when t_{out} = 2 min

(d) MSE values when t_{out} = 3 min

(e) MSE values when t_{out} = 4 min

Fig. 7. MSE comparison of the two types of models.

V. Conclusion

It is important for the minute time scale output forecasting of a single photovoltaic power station to achieve accurate sky image prediction. To avoid the shortages of DIPT based sky image prediction methods, 2D-CAE based and 3D-CAE based sky image prediction models are proposed. In this paper, four structures are designed for CAE based models, and the control models are built by using two typical DIPT methods: PIV and FPCT. Under five prediction scales, two conclusions can be drawn according to the comparison. The first is that all the best performances are obtained by CAE based models, which means the CAEs are superior to DIPT in sky image prediction under given cases. The second is that 2D CAEs are basically better than 3D CAEs when it comes to the performances of sky image prediction. In addition, the best structure under each scenario are pointed out in order to aid in further study.

In fact, the impacts caused by deep penetration of renewable energy on generation flexibility and operating costs of power grid [25] are significant and reflect in many aspects not only include power forecasting [26-28] for supply-demand balancing and energy trading under various scenarios [29-31], but also refer to load forecasting/load pattern [32-35], demand response applications [36-38], aggregator aggregated capacity forecasting and multi-aggregator scheduling with plenty distributed PV systems [39-41]. The above mentioned research topics will be conducted in the future works.

Acknowledgment

This work was supported by the National Key R&D Program of China (Technology and application of wind power/photovoltaic power prediction for promoting renewable energy consumption, 2018YFB0904200) and eponymous Complement S&T Program of State Grid Corporation of China (SGLNDKOOKJJS1800266). Also, João P. S. Catalão acknowledges the support by FEDER funds through COMPETE 2020 and by Portuguese funds through FCT, under POCI-01-0145-FEDER-029803 (02/SAICT/2017).

References

[1] Installed capacity trends, IRENA. [Online]. Available: https://www.irena.org/solar. (accessed on 1 October 2019).

[2] Y. Wu, S. Chang, and P. Mandal, "Grid-Connected Wind Power Plants: A Survey on the Integration Requirements in Modern Grid Codes," *IEEE Trans. Ind. Appl.*, vol. 55, no. 6, pp. 5584-5593, Nov.-Dec. 2019.

[3] Y. Wu, J. Lin, and H. Lin, "Standards and Guidelines for Grid-Connected Photovoltaic Generation Systems: A Review and Comparison," *IEEE Trans. Ind. Appl.*, vol. 53, no. 4, pp. 3205-3216, Jul.-Aug. 2019.

[4] F. Wang, Z. Xuan, Z. Zhen, K. Li, T. Wang, and M. Shi, "A day-ahead PV power forecasting method based on LSTM-RNN model and time correlation modification under partial daily pattern prediction framework," *Energy Convers. Manag.*, vol. 212, Art. no. 112766, May. 2020.

[5] Z. Zhen, J. Liu, Z. Zhang, F. Wang, H. Chai, Y. Yu, X. Lu, T. Wang, and Y. Lin, "Deep learning based surface irradiance mapping model for solar PV power forecasting using sky image," *IEEE Trans. Ind. Appl.*, vol. 56, no. 4, pp. 3385-3396, Jul.-Aug. 2020.

[6] F. Wang, Z. Xuan, Z. Zhen, Y. Li, K. Li, L. Zhao, M. Shafie-khah, and J. P.S. Catalão, "A minutely solar irradiance forecasting method based on real-time sky image-irradiance mapping model," *Energy Convers. Manag.* vol. 220, Art. no. 113075, Sep. 2020.

[7] F. Wang, K. Li, L. Zhou, H. Ren, J. Contreras, M. Shafie-khah, and J. P. S. Catalão, "Daily pattern prediction based classification modeling approach for day-ahead electricity price forecasting," *Int. J. Electr. Power Energy Syst*., vol. 105, pp. 529-540, Feb. 2019.

[8] M. Lourenco, J. P. Barreto, and F. Vasconcelos, "sRD-SIFT: keypoint detection and matching in images with radial distortion," *IEEE Trans. Robot*., vol. 28, no. 3, pp. 752-760, Jun. 2012.

[9] S. Dev, F. M. Savoy, Y. H. Lee, and S. Winkler, "Short-term prediction of localized cloud motion using ground-based sky imagers," in 2016 *IEEE Region 10 Conference (TENCON)*, 2016, pp. 2563-2566.

[10] A. Zaher, S. Thil, J. Nou, A. Traoré, and S. Grieu, "Comparative study of algorithms for cloud motion estimation using sky-imaging data," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 5934-5939, Jul. 2017.

[11] Z. Zhen, Z. Xuan, F. Wang, R. Sun, N. Duić, and T. Jin, "Image phase shift invariance based multi-transform-fusion method for cloud motion displacement calculation using sky images," *Energy Convers. Manag*., vol. 197, pp. 111853, Oct. 2019.

[12] F. Wang, Z. Zhen, C. Liu, Z. Q. Mi, S. M. Hodge, M. Shafie-khah, and J. P. S. Catalão, "Image phase shift invariance based cloud motion displacement vector calculation method for ultra-short-term solar PV power forecasting," *Energy Convers. Manag*., vol. 157, pp. 123-135, Feb. 2018.

[13] S. Aigner, and M. Körner, "FutureGAN: anticipating the future frames of video sequences using spatio-temporal 3d convolutions in progressively growing GANs," *arXiv preprint arXiv:1810.01325v2*, 2018.

[14] M. Mathieu, C. Couprie, and Y. LeCun, "Deep multi-scale video prediction beyond mean square error," *arXiv preprint arXiv:1511.05440v6*, 2015.

[15] Z. W. Liu, R. A. Yeh, X. O. Tang, Y. M. Liu, and A. Agarwala, "Video Frame synthesis using deep voxel flow," in 2017 *IEEE International Conference on Computer Vision (ICCV)*, 2017, pp. 4473-4481.

[16] J. Gu, Z. Wang, J. Kuen, L. Ma, A. Shahroudy, B. Shuai, T. Liu, X. Wang, G. Wang, J. Cai, and T. Chen, "Recent advances in convolutional neural networks," *Pattern Recognit*., vol. 77, pp. 354-377, May 2018.

[17] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436-444, May 2015.

[18] V. Dumoulin, and F. Visin, "A guide to convolution arithmetic for deep learning," *arXiv preprint arXiv:1603.07285v1*, 2016.

[19] D. Tran, L. Bourdev, R. Fergus, L. Torresani, and M. Paluri, "Learning spatiotemporal features with 3D convolutional networks," in 2015 *IEEE International Conference on Computer Vision (ICCV)*, 2015, pp. 4489-4497.

[20] S. Ioffe, and C. Szegedy, "Batch normalization: accelerating deep network training by reducing internal covariate shift," *arXiv preprint arXiv:1502.03167v3*, 2015.

[21] "Rectifier nonlinearities improve neural network acoustic models," A. L. Maas, A. Y. Hannun, and A. Y. Ng.. [Online]. Available: https://ai.stanford.edu/~amaas/papers/relu_hybrid_icml2013_final.pdf. (accessed on 1 October 2019).

[22] National oceanic and atmospheric administration. [Online]. Available: https://www.noaa.gov/. (accessed on 1 October 2019).

[23] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. Image Process*., vol. 13, pp. 600-612, May 2004.

[24] D. P. Kingma, and J. Ba, "Adam: a method for stochastic optimization," *arXiv preprint arXiv:1412.6980v9*, 2017.

[25] Y. Wu, W. Tan, S. Huang, Y. Chiang, C. Chiu, and C. Su, "Impact of Generation Flexibility on the Operating Costs of the Taiwan Power System Under a High Penetration of Renewable Power," *IEEE Trans. Ind. Appl.*, vol. 56, no. 3, pp. 2348-2359, May-Jun. 2020.

[26] Y. Wu, P. Sun, T. Wu, J. Hong, and M.Y. Hassan, "Probabilistic Wind-Power Forecasting Using Weather Ensemble Models," *IEEE Trans. Ind. Appl.*, vol. 54, no. 6, pp. 5609-5620, Nov.-Dec. 2018.

[27] M. Yang, S. Fan, and W. Lee, "Probabilistic Short-Term Wind Power Forecast Using Componential Sparse Bayesian Learning," *IEEE Trans. Ind. Appl.*, vol. 49, no. 6, pp. 2783-2792, Nov.-Dec. 2013.

[28] J. Ma, M. Yang, X. Han, and Z. Li, "Ultra-Short-Term Wind Generation Forecast Based on Multivariate Empirical Dynamic Modeling," *IEEE Trans. Ind. Appl.*, vol. 54, no. 2, pp. 1029-1038, Mar.-Apr. 2018.

[29] Z. Zhang, H. Tang, P. Wang, Q. Huang, and W. Lee, "Two-Stage Bidding Strategy for Peer-to-Peer Energy Trading of Nanogrid," *IEEE Trans. Ind. Appl.*, vol. 56, no. 2, pp. 1000-1009, Mar.-Apr. 2020.

[30] H. Xu, Y. Lin, X. Zhang, and F. Wang, "Power system parameter attack for financial profits in electricity markets," *IEEE Trans. Smart Grid*, vol. 11, no. 4, pp. 3438-3446, Jul. 2020, DOI: 10.1109/TSG.2020.2977088.

[31] Z. Ding, W. Lee, and J. Wang, "Stochastic Resource Planning Strategy to Improve the Efficiency of Microgrid Operation," *IEEE Trans. Ind. Appl.*, vol. 51, no. 3, pp. 1978-1986, May-Jun. 2015.

[32] X. Wang, W. Lee, H. Huang, R.L. Szabados, Y. Wang, and P.V. Olinda, "Factors that Impact the Accuracy of Clustering-Based Load Forecasting," *IEEE Trans. Ind. Appl.*, vol. 52, no. 5, pp. 3625-3630, Sept.-Oct. 2016.

[33] K. Li, X. Cao, X. Ge, F. Wang, X. Lu, M. Shi, R. Yin, Z. Mi, and S. Chang, "Meta-heuristic optimization based two-stage residential load pattern clustering approach considering intra-cluster compactness and inter-cluster separation," *IEEE Trans. Ind. Appl.*, vol. 56, no. 4, pp. 3375-3384, Jul.-Aug. 2020.

[34] F. Wang, K. Li, N. Duić, Z. Mi, B.M. Hodge, M. Shafie-khah, and J. P. S Catalão, "Association rule mining based quantitative analysis approach of household characteristics impacts on residential electricity consumption patterns," *Energy Convers. Manag.*, vol. 171, pp. 839-854, Sep. 2018.

[35] S. Yan, K. Li, F. Wang, X. Ge, X. Lu, Z. Mi, H. Chen, and S. Chang, "Time-frequency features combination-based household characteristics identification approach using smart meter data," *IEEE Trans. Ind. Appl.*, vol. 56, no. 3, , pp. 2251-2262, May-Jun. 2020.

[36] F. Wang, K. Li, C. Liu, Z. Mi, M. Shafie-Khah, and J. P. S. Catalao, "Synchronous pattern matching principle-based residential demand response baseline estimation: Mechanism analysis and approach description," *IEEE Trans. Smart Grid*, vol. 9, no. 6, pp. 6972–6985, Nov. 2018.

[37] K. Li, L. Liu, F. Wang, T. Wang, N. Dui, M. Shafie-khah and J. P. S. Catalão, "Impact factors analysis on the probability characterized effects of time of use demand response tariffs using association rule mining method," *Energy Convers. Manag.*, vol. 197, Art. no. 111891, Oct. 2019.

[38] X. Lu, K. Li, H. Xu, F. Wang, Z. Zhou, and Y. Zhang, "Fundamentals and business model for resource aggregator of demand response in electricity markets," *Energy*, vol. 204, Art. no. 117885, May 2020.

[39] F. Wang, B. Xiang, K. Li, X. Ge, H. Lu, J. Lai, and P. Dehghanian, "Smart households' aggregated capacity forecasting for load aggregators under incentive-based demand response programs," *IEEE Trans. Ind. Appl.*, vol. 56, no. 2, pp. 1086-1097, Mar.-Apr. 2020.

[40] V. Gupta, R. Kumar, and B.K. Panigrahi, "User Willingness based Decentralized EV Charging Management in Multi-Aggregator Scheduling," *IEEE Trans. Ind. Appl.*, Early Access, DOI: 10.1109/TIA.2020.2993988.

[41] K. Li, F. Wang, Z. Mi, M. Fotuhi-Firuzabad, N. Duić, and T. Wang, "Capacity and output power estimation approach of individual behind-the-meter distributed photovoltaic system for demand response baseline estimation," *Appl. Energy*, vol. 253, Art. no. 113595, Nov. 2019.