MPL-GAN: Toward Realistic Meteorological Predictive Learning Using Conditional GAN

HONG-BIN LIU AND ICKJAI LEE

College of Science and Engineering, James Cook University, Cairns, QLD 4870, Australia

Corresponding author: Ickjai Lee (ickjai.lee@jcu.edu.au)

ABSTRACT Meteorological imagery prediction is an important and challenging problem for weather forecasting. It can also be seen as a video frame prediction problem that estimates future frames based on observed meteorological imageries. Despite it is a widely-investigated problem, it is still far from being solved. Current state-of-the-art deep learning based approaches mainly optimise the mean square error loss resulting in blurry predictions. We address this problem by introducing a Meteorological Predictive Learning GAN model (in short MPL-GAN) that utilises the conditional GAN along with the predictive learning module in order to handle the uncertainty in future frame prediction. Experiments on a real-world dataset demonstrate the superior performance of our proposed model. Our proposed model is able to map the blurry predictions produced by traditional mean square error loss based predictive learning methods back to their original data distributions, hence it is able to improve and sharpen the prediction. In particular, our MPL-GAN achieves an average sharpness of 52.82, which is 14% better than the baseline method. Furthermore, our model correctly detects the meteorological movement patterns that traditional unconditional GANs fail to do.

INDEX TERMS Meteorological prediction, spatio-temporal forecasting, GAN, video prediction.

I. INTRODUCTION

Weather forecasting is one of the main applications of meteorological prediction. It is important for our daily life as well as industrial and agricultural production. Common uses include precipitation nowcasting [1], [2], streamflow prediction [3]–[7], wind speed simulation [8], radiation estimation [9], and temperature forecasting [10], [11]. Numerous techniques have been proposed to predict more accurate weather measurements including Numerical Weather Prediction (NWP), radar map based methods, and satellite imagery based methods. Recently, with the advances of deep learning techniques, researchers adopt *Recurrent Neural Network* (RNN) based methods to improve those traditional approaches in order to address this challenging problem. For example, authors [1] formulated the precipitation nowcasting problem into a spatio-temporal sequence forecasting model, and proposed a LSTM-based model named ConvLSTM for radar echo map prediction. A Seq2Seq-LSTM based model [11] was proposed to improve NWP performance through historical observations. A study [12] proposed an adversarial model to predict cyclone trajectory with satellite imageries. These studies reveal that radar and satellite

imageries play more and more important roles in meteorological prediction, not only because they are more robust, but also they provide end-users with more sequential information and better visualisations of the history from current to predicted atmosphere. However, these approaches share some common drawbacks: they do not generalise well in real-world meteorological datasets especially for long term predictions. To be more specific, the pioneering work, ConvLSTM, produces blurry radar imagery predictions, and keeps going worse as the time step moves forward. These meteorological imageries do not appear to be realistic but blurry resulting in unpleasant visualisations. These drawbacks are mainly due to two reasons. First, these models optimise Euclidean losses such as *Mean Absolute Error* (MAE) and *Mean Square Error* (MSE) across the overall length of sequential meteorological imageries. A few studies introduced various models with MAE and MSE, but produced blurry images [13], [14]. This is mainly due to the assumption that the data is drawn form the Gaussian distribution which only works on a continuous portion of the image while ignoring isolated small regional areas. Second, due to the nature of RNN architecture, small errors are quickly accumulated to become large errors along the generated sequence because of the gap between training and inference [15], [16]. These two causes indicate that it is crucial to include an uncertainty

handling procedure to generate realistic meteorological predictions.

Meanwhile, the video frame prediction can be modelled as a spatio-temporal sequence forecasting problem. For instance, authors [17] extracted a sequence of images from video frames, and proposed an encoder-decoder RNN based model named PredRNN [17] and its improved version PredRNN++ [18]. Nevertheless, these models suffer from the same drawback as ConvLSTM produces blurry predictions. Recently, *Generative Adversarial Network* (GAN) [19] was used to handle uncertainties in video frame prediction [13], [20]–[22]. GAN models match two distributions by one generator and one discriminator playing the minmax game, where the generator learns to generate samples to fool the discriminator and the discriminator learns to distinguish these fake samples. These unconditional GAN based models are able to produce realistic looking videos by learning a high dimensional distribution of complex datasets. However, these models are not suitable for meteorological predictive learning even though they are able to produce realistic looking and temporally coherent video frames. This is because those generated video frames do not model the real-world meteorological changes given by the source of meteorological imagery frames. Note that, meteorological prediction needs to consider the moving entities' (pixel wise) direction, speed, rotation acceleration and other information.

To sum up, on one hand RNN based meteorological predictive models with MAE and MSE produce blurry predictions. On the other hand, GAN based models are able to generate realistic looking video frames but fail to catch the actual atmospheric movement missing local variations and patterns. In this work, we propose a Conditional GAN based model named Meteorological Predictive Learning GAN (in short MPL-GAN) that optimises both the regression loss and GAN loss, and aims at generating realistic meteorological predictions. Optimising regression loss aims at modeling the real-world atmosphere imagery movement which is crucial for weather forecasting. The GAN loss is used to estimate the data distribution to deal with the uncertainty to produce non-blurry predictions.

Our main contributions are summarised as follows:

- to the best of our knowledge, this is the first model that combines regression loss with GAN loss to generate realistic meteorological predictions that provide better visualisations;
- to conduct extensive experiments on a real-world radar imagery dataset. Experimental results demonstrate our model generates non-blurry predictions even in the long term while it catches real-world atmosphere changes;
- to provide an extensive experiment analysis to show that the pure GAN model without the predictive learning module fails to catch the actual atmospheric movement, which demonstrates the effectiveness of our proposed MPL-GAN model detecting meteorological changes.

II. RELATED WORK

This section briefly reviews related studies.

A. METEOROLOGICAL PREDICTIVE LEARNING

Optimal flow based methods [23], [24] have been a long history in the meteorological predictive learning literature. With the recent advances in deep learning, authors [1] explored the possibility of applying RNN, and proposed a model called Convolutional LSTM (ConvLSTM) [1] and its improved version TrajGRU [2] for radar echo imagery prediction. Both approaches tried to optimise the MSE loss. In the meantime, the video frame prediction and the traffic flow prediction can be considered as the same problem. PredRNN and its improved version PredRNN++ are proposed by [17] to tackle this problem, again these methods optimised the MSE loss as well, and they shared the same issue that the prediction gets more blurry over time.

Beyond the meteorological imagery predictive learning, neural network based methods have also been used in numerical weather forecasting. For example, a study [11] proposed a Seq2Seq LSTM to predict temperature, wind speed, and relative humidity. Another study [10] improved such method by introducing a temporal progressive growing schedule sampling strategy. Nevertheless, these approaches suffer from the same long term prediction accuracy degradation.

B. GAN FOR IMAGE AND VIDEO GENERATION

GAN [19] has been the most popular generative model since it was first released in 2014. Since then GAN models have shown their superior abilities especially in image generation, starting from hand-written digit generation [19], [25] to large scale image set generation [26], [27]. Recently, researchers try to push the limit of GAN by generating photo-realistic videos using unconditional GAN [13], [20]–[22], [28]. Those video GANs aim at producing photo-realistic and temporal coherent videos, and they are used to match the high-dimensional data distribution between the two. Note that, there are no other considerations those models take into account. That is, for given initial frames, generated frames do not need to consider the moving entities' direction, speed and other moving information. However, these moving properties play an important role in our study, and our GAN takes them into account unlike traditional unconditional GANs.

III. PROPOSED MPL-GAN

In this section, we will describe our proposed MPL-GAN model that aims to produce realistic looking meteorological imageries. Figure 1 shows the overall architecture of our proposed model that contains a predictive learning model to generate predictions and a Conditional GAN module to map those predictions back to photorealistic distributions. First, we formulate the meteorological predictive learning as follows:

Definition 1: Given a length-k matrix sequence $\mathcal{M}_1^k \equiv [\mathcal{X}_1, \mathcal{X}_2, \ldots, \mathcal{X}_k]$, where each matrix $\mathcal{X}_t \in \mathcal{M}$ represents
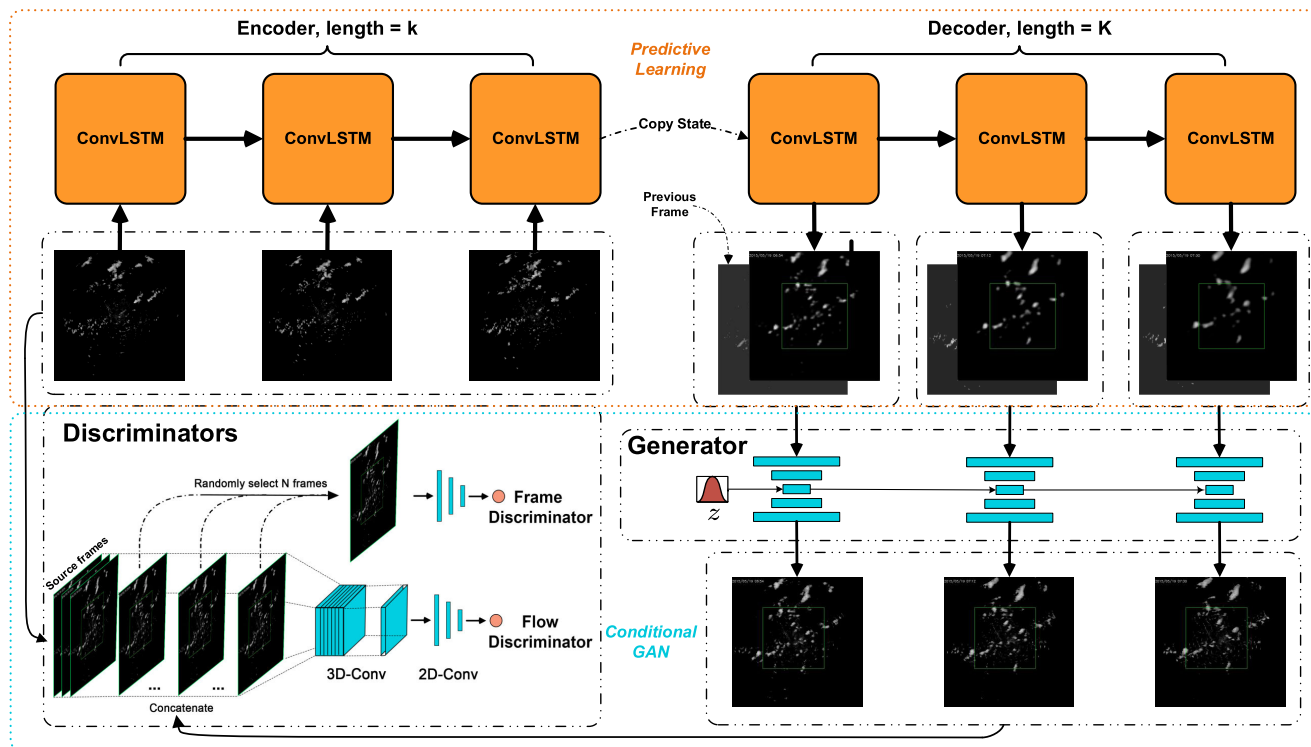
FIGURE 1. Architecture of our proposed model MPL-GAN. The orange box indicates the predictive learning module that models the real-world meteorological movement patterns. We use ConvLSTM for evaluation purpose that can be replaced by other predictive learning approaches if needed. The blue box concludes the Conditional GAN module which consists of the conditional generator and two discriminators $\mathcal{D}_{Fr}$ and $\mathcal{D}_{Fl}$.

the meteorological imagery at time-step t. Metrological predictive learning is to predict a sequence of corresponding meteorological imageries of following K time-steps based on the past frames of $\mathcal{M}$, denoted as $\hat{\mathcal{P}}_1^K \equiv \left[\hat{\mathcal{X}}_1, \hat{\mathcal{X}}_2, \ldots, \hat{\mathcal{X}}_K \right]$.

Note, a meteorological imagery carries important weather information such as rainfall, temperature, and wind speed etc.

A. PREDICTIVE LEARNING

In order to model meteorological changes, we adopt encoder-decoder ConvLSTM [1] as a predictive learning module. As investigated by previous studies [1], [2], predictive learning aims to capture the local spatio-temporal pattern movement such as rotation and scaling. As we mentioned earlier, all existing GAN-based next frame prediction models are not suitable for meteorological predictive learning as these models do not capture the real-world meteorological changes. Furthermore, these GAN models without such predictive learning module are not able to produce long term predictions. For example, a study [13] can only produce a maximum of two frames for future video predictions. In contrast, our MPL-GAN model generates the next prediction conditional on previous ground-truth and current predictive output with conditional GAN. We manage to generate more than 10 frames of non-blurry and realistic meteorological imagery predictions, and yet still model the real-world atmospheric changes with the help of a predictive learning module. This demonstrates that the predictive learning module is crucial

for modelling meteorological changes. Note that, we use ConvLSTM for evaluation purpose in this study, but it can be replaced by any other advanced models such as Traj-GRU [2] and PredRNN++ [18]. On one hand, the predictive learning module is required for modelling meteorological changes, but on the other hand, naive predictive learning models suffer from the blurry image issue and they need to be specially refined for meteorological change analysis. In the next section, we will introduce Conditional GAN to solve the blurry problem caused by the traditional naive predictive learning module.

B. CONDITIONAL GAN

GAN [19] attempts to learn a mapping function $\mathcal{G}$ to map a random noise vector z to an image $\mathcal{X}$, $\mathcal{G} : z \rightarrow \mathcal{X}$. In our settings, we aim to map the blurry prediction produced by ConvLSTM to the original non-blurry distribution. Let $\hat{\mathcal{P}}_1^K$ denote the generated sequence of ConvLSTM, $\mathcal{M}_1^K$ denote the observed ground-truth frames, our goal is to train a conditional Generator $\mathcal{G} : \{z, \hat{\mathcal{P}}_1^K\} \rightarrow \mathcal{M}_1^K$.

1) CONDITIONAL GENERATOR

As the prediction sequence is generated recursively by the ConvLSTM, we train the Conditional Generator $\mathcal{G} : \{z, \hat{\mathcal{X}}_t\} \rightarrow \mathcal{X}_t$ per frame along with the ConvLSTM time steps instead of taking the whole generated sequence to train the GAN generator, where $\hat{\mathcal{X}}_t$ denotes the ConvLSTM output at

time t and $\mathcal{X}_t$ is the ground-truth frame at time t. However, when the ConvLSTM prediction gets more blurry in the later time steps, the Conditional GAN gets harder to map the conditional distributions between the two. To solve this problem, we train the generator along with conditioning on the previous frame $\mathcal{X}_{t-1}$, i.e. $\mathcal{G}: \{z, \mathcal{X}_{t-1}, \hat{\mathcal{X}}_t\} \rightarrow \mathcal{X}_t$. Ideally, we should train the generator conditioning on the current ground-truth frame $\mathcal{X}_t$ and current ConvLSTM prediction frame $\hat{\mathcal{X}}_t$. We use the previous frame instead of the current frame based on the observation that two consecutive meteorological frames are very similar in terms of data distribution. They even look to be very similar since the atmosphere normally changes gradually and slowly. Moreover, during the inference stage, none of the previous and current ground-truth frames would be available. Then, we can replace the previous ground-truth frame $\mathcal{X}_{t-1}$ with the generator output of previous time step during the inference phase, i.e. $\tilde{\mathcal{X}}_t = \mathcal{G}(z, \tilde{\mathcal{X}}_{t-1}, \hat{\mathcal{X}}_t)$. We can also think the other way around, since the ground-truth frames are not available during the inference phase. However, we make an assumption that the output distribution of our generator perfectly matches the actual data distribution, then the data distribution can be carried forward from the last known ground-truth frame to the predicting frame recursively. Therefore, we use the ground-truth frame $\mathcal{X}_{t-1}$ instead of $\tilde{\mathcal{X}}_{t-1}$ during the training phase for training stability, then replace $\mathcal{X}_{t-1}$ with $\tilde{\mathcal{X}}_{t-1}$ during inference.

2) FRAME DISCRIMINATOR

We randomly select N frames among the K time steps to train the frame discriminator $\mathcal{D}_{\text{Fr}}$. The $\mathcal{D}_{\text{bFr}}$ outputs *1* for the true frame $\mathcal{X}_t$ and outputs *0* for the fake frame $\tilde{\mathcal{X}}_t$. Then, we train the frame discriminator by optimising the minmax game defined in the original GAN, we use Hinge Loss [29] for $\mathcal{L}_{\mathcal{D}_{\text{Fr}}}$ defined as follow:

$$\mathcal{L}_{\mathcal{D}_{\text{Fr}}} = \mathbb{E}[\max(0, 1 - \mathcal{D}_{\text{Fr}}(\mathcal{X}_t)) \\ + \mathbb{E}[\max(0, 1 + \mathcal{D}_{\text{Fr}}(\tilde{\mathcal{X}}_t))]. \quad (1)$$

3) FLOW DISCRIMINATOR

Frame discriminator aims to ensure that the generator produces samples matching the actual data distribution, i.e. to ensure the produced samples looked to be realistic. On top of this, similar to the video discriminator proposed by [22], we propose a flow discriminator $\mathcal{D}_{\text{Fl}}$ ensuring the generator produces temporal coherent frames. Similarly, $\mathcal{D}_{\text{Fl}}$ outputs *1* for the real sequence $\mathcal{M}_1^K$ and outputs *0* for the generator sequence $(\mathcal{M}_1^k; \tilde{\mathcal{P}}_k^K)$, here we concatenate the initial source sequence $\mathcal{M}_1^k$ and the conditional generated sequence $\tilde{\mathcal{P}}_k^K$. $\mathcal{L}_{\mathcal{D}_{\text{Fl}}}$ defined as follow:

$$\mathcal{L}_{\mathcal{D}_{\text{Fl}}} = \mathbb{E}[\max(0, 1 - \mathcal{D}_{\text{Fl}}(\mathcal{M}_1^K)) \\ + \mathbb{E}[\max(0, 1 + \mathcal{D}_{\text{Fl}}((\mathcal{M}_1^k, \tilde{\mathcal{P}}_k^K)))]. \quad (2)$$

C. TRAINING

Again, a predictive learning module is essential for modelling the real-world meteorological movement patterns,

and conditional GAN is used to map the blurry predictions generated by predictive learning back to non-blurry imagery distributions. Therefore, we divide the training process into two stages. First, we start training the predictive learning module, when the training of predictive learning module is almost stable then we start the training of the GAN module.

1) PREDICTIVE LEARNING TRAINING

Following the settings of ConvLSTM [1] and TrajGRU [17], we train our Predictive Learning (PL) module by minimising the balance of MSE and MAE losses (B-MSE-MAE) with Stochastic Gradient Descent (SGD) and Back-propagation Through Time (BPTT) [30]. We train the B-MSE-MAE loss until it becomes stable before we start training the Conditional GAN so that the Conditional GAN learns the stabled data distribution. However, we continue to train the PL module along with the GAN module even the loss does not decrease. The intuition behind this is to give the GAN variances of distribution to make the GAN become more robust.

2) CONDITIONAL GAN TRAINING

Training GAN models requires training both the generator and discriminator by optimising the minimax game [19], where the generator learns to fool the discriminator with generated fake samples and the discriminator learns to identify true and fake samples. We follow the same spirit and extend it to training one conditional generator and two separate discriminators. The losses of Frame Discriminator $\mathcal{D}_{\text{Fr}}$ and Flow Discriminator are defined in Equation 1 and Equation 2. Now we define the loss function for conditional generator as follow:

$$\mathcal{L}_{\mathcal{G}} = -\mathbb{E}[\mathcal{D}_{\text{Fr}}(\mathcal{G}(z, \mathcal{X}_{t-1}, \hat{\mathcal{X}}_t)) + \mathcal{D}_{\text{Fl}}(\mathcal{G}^{\phi}(z, \mathcal{X}_{t-1}, \hat{\mathcal{X}}_t))], \quad (3)$$

where ϕ denotes a process of applying the generator recursively with the ConvLSTM time steps to generate a sequence flow of frames. Therefore our overall optimisation goal is to minimise $\mathcal{L}_{\mathcal{D}_{\text{Fr}}}$ and $\mathcal{L}_{\mathcal{D}_{\text{Fl}}}$ that maximises the probability of the discriminators identifying fake frames and fake sequence; and minimises $\mathcal{L}_{\mathcal{G}}$ to maximise the probability of the generator producing samples that the discriminators think they are true.

$$\min_{\mathcal{G}} \max_{\mathcal{D}_{\text{Fr}}} \mathcal{L}_{\mathcal{D}_{\text{Fr}}}(\mathcal{G}, \mathcal{D}_{\text{Fr}}) + \min_{\mathcal{G}} \max_{\mathcal{D}_{\text{Fl}}} \mathcal{L}_{\mathcal{D}_{\text{Fl}}}(\mathcal{G}, \mathcal{D}_{\text{Fl}}). \quad (4)$$

Note, N random frames will be selected for training $\mathcal{D}_{\text{Fr}}$ N times for each training batch b, whereas $\mathcal{D}_{\text{Fl}}$ will be trained once for each b. Moreover, the gradient of $\mathcal{G}$ will be back-propagated multiple times recursively with operation ϕ when training $\mathcal{G}$ with $\mathcal{D}_{\text{Fl}}$. This makes it extremely difficult for training the GAN. Following the principle of [20] and [27], we downsample each frame of the sequence passing to $\mathcal{D}_{\text{Fl}}$ to overcome the difficulties of training.

IV. EXPERIMENT

In this section, we briefly describe the dataset used and provide experimental results.

A. DATASET

We use HKO-7 [1], [2] radar echo imagery dataset to evaluate our proposed MPL-GAN model. The radar echo imagery is recorded every 6 minutes, therefore there are 240 frames per day. Each frame contains 480×480 pixels that cover a $512km \times 512km$ area. We sample data into sequences of the length of 15 frames by a sliding window, 5 for the encoder and 10 for the decoder. In the total number of 993 days of data, we randomly select 80% for the training set, 5% for the validation set, and 15% for the testing set. Unlike the original experimental setting of ConvLSTM and TrajGRU where they try to predict the pixel value and report the precipitation prediction based on that, we focus on imagery frame prediction that is realistic for better visualisation.

B. IMPLEMENTATION AND PARAMETERS

1) PREDICTIVE LEARNING

We use ConvLSTM as the predictive learning module. We implement a three layer ConvLSTM encoder-decoder with the following parameters for each layer: $[3 \times 3 - 64, 3 \times 3 - 192, 3 \times 3 - 192]$.

2) CONDITIONAL GAN

Training GAN is challenging, thus we carefully choose our architecture for the generator and discriminators. For the generator, our architecture is somewhat similar to PG-GAN [27]. In order to match the resolution of generated samples, we upsample the original resolution from 480×480 to 512×512. We randomly select $N = 2$ frames to train the Frame Discriminator $\mathcal{D}_{Fr}$. 3D-Conv of Flow Discriminator consists of three layers set to the following parameters: $[3 \times 3 \times 3 - 64, 3 \times 3 \times 3 - 16, 3 \times 3 \times 3 - 1]$.

We implement our model using PyTorch 1.4, a well known deep learning library developed by Facebook. Our model is trained and evaluated on a server with Nvidia V100 32GB GPU and Intel(R) Xeon(R) Gold 5118 CPU @ 2.30GHz (24 cores). We train our model with Adam optimiser [31] with a learning rate of $1e^{-4}$ for ConvLSTM and $1e^{-3}$ for the generator and discriminators. Batch size is set to 2 due to the resource consumption of conditional GAN model. For all models including baseline methods, we train 100,000 batches and select the best model based on the minimum MSE against the validation set. For MPL-GAN, we train the PL module for 10,000 batches before training the conditional GAN. All experimental results are reported based on the test set.

C. OVERALL EVALUATION

1) BASELINES

In order to evaluate the effectiveness of our proposed model, we compare our model to two baseline methods.

- PL with MSE. In order to show that the PL with MSE produces blurry predictions, we compare our model to the pure ConvLSTM without the conditional GAN module [1].

- PG-GAN. We extend the PG-GAN [27] from image generation to sequence generation with the same architecture of our conditional GAN module, that has a Frame discriminator and a Flow discriminator. This is also very similar to DVD-GAN [22].

2) EVALUATION MATRIX

We use the sharpness measure based on the gradient of two images defined in [13] as follow:

$$\text{Sharp.} = 10 \log_{10} \frac{\max^2_{\tilde{\mathcal{X}}}}{\frac{1}{N} \left(\sum_i \sum_j \left| \left(\nabla_i \mathcal{X} + \nabla_j \mathcal{X} \right) - \left(\nabla_i \tilde{\mathcal{X}} + \nabla_j \tilde{\mathcal{X}} \right) \right| \right)}, \quad (5)$$

where, $\mathcal{X}$ is the ground-truth frame and $\tilde{\mathcal{X}}$ is the output frame; $\nabla_i \mathcal{X} = \left| \mathcal{X}_{i,j} - \mathcal{X}_{i-1,j} \right|$ and $\nabla_j \mathcal{X} = \left| \mathcal{X}_{i,j} - \mathcal{X}_{i,j-1} \right|$; $\max_{\tilde{\mathcal{X}}}$ is the maximum possible value of the image intensities. We report the average sharpness of the test set across K frames in the table as well as the individual frame evaluations as a line chart as shown in Figure 2.
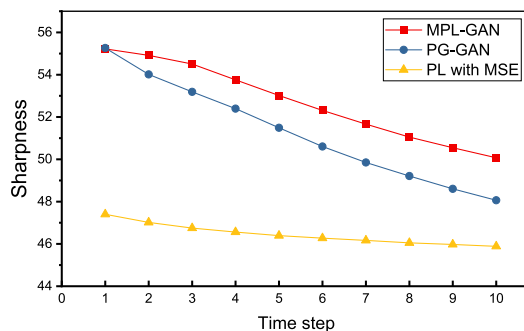

FIGURE 2. Sharpness of 10 prediction time steps based on the test set.

TABLE 1. Experimental results based on the test data, averaged by 10 prediction time steps (Greater number indicates sharper prediction imagery).

Model	Sharpness ↑
PL with MSE [1]	46.48
PG-GAN [27]	51.27
MPL-GAN (ours)	52.82

3) EXPERIMENT ANALYSIS

We conduct both quantitative and qualitative evaluations with two baseline approaches. As shown in Table 1, our proposed MPL-GAN achieves the overall average sharpness of 52.82; PG-GAN achieves a similar result of 51.27; whereas PL with MSE only achieves 46.48. This shows that GAN based models are able to generate sharper meteorological imagery predictions compared to PL with MSE. This is because GAN models handle uncertainties of future frames. Per frame quantitative comparison in FIGURE 2 indicates that GAN based
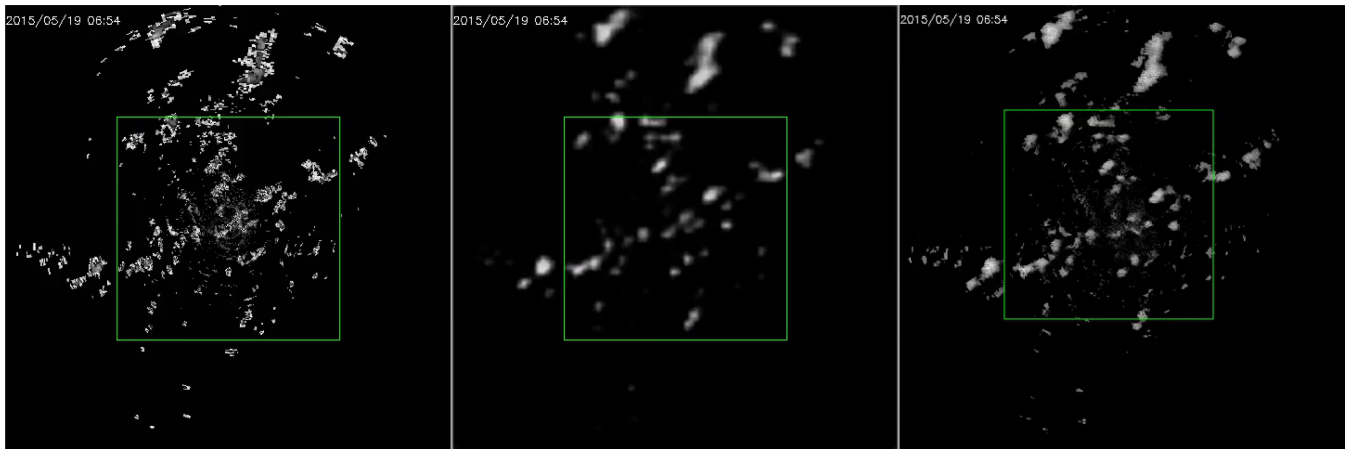
FIGURE 3. Future meteorological prediction samples. *Left*: ground truth. *Middle*: PL with MSE. *Right*: MPL-GAN (ours). *Click to view the animations with Adobe Acrobat Reader.*
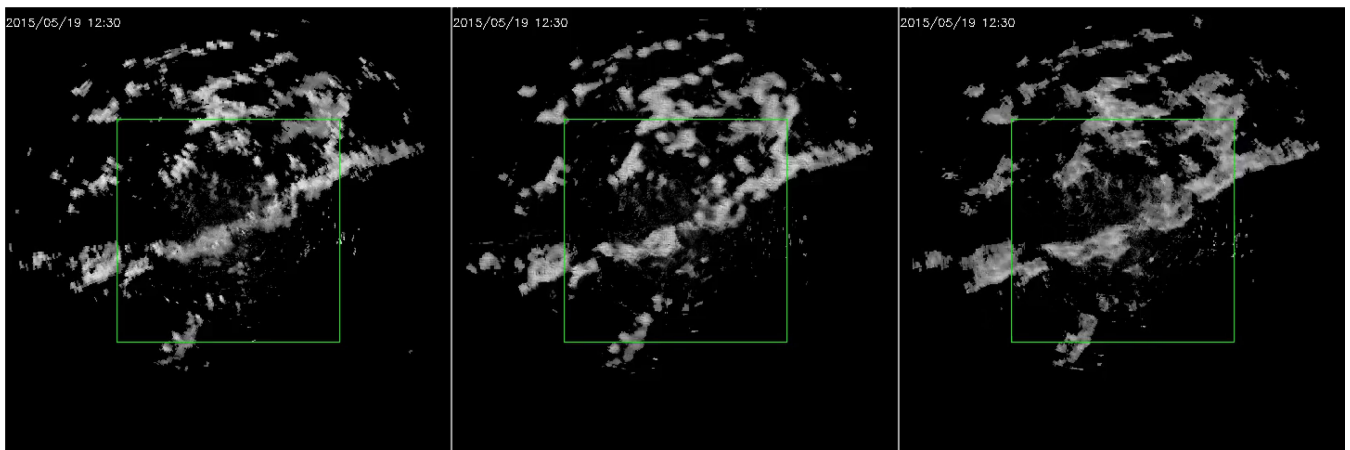

FIGURE 4. Meteorological prediction movement of our MPL-GAN and PG-GAN. *Left*: ground truth. *Middle*: MPL-GAN (ours). *Right*: PG-GAN. *Click to view the animations with Adobe Acrobat Reader.*

models not only beat PL with MSE in the average score but also in the long term predictions. MPL-GAN and PG-GAN achieve a similar score in the first frame, however, PG-GAN drops performance quickly in the long term.

Besides quantitative evaluation, we visualise a sample of prediction sequence in Figure 3 (please view the animation by clicking the figure using Adobe Acrobat Reader). As shown in the animation, both PL with MSE and MPL-GAN catch the real-world meteorological movement patterns. However, PL with MSE generates blurry predictions and continues to get more blurry over time, especially in the long term. Whereas our proposed MPL-GAN continues to generate realistic looking and sharp predictions. Furthermore, if we look at small regions of the prediction frames, PL with MSE tends to omit small areas as the result of MSE loss, whereas our model has more regional details with the incorporation of uncertainty handling in GAN.

Furthermore, the whole experiment aims to find out whether the GAN model is able to solve the blurry prediction

problem caused by PL with MSE. In fact, we can see our proposed model as PL with MSE with the advanced version of PG-GAN with two discriminator heads. As stated previously, GAN based models are able to produce sharper predictions compared to PL with MSE. However, as samples of Figure 4 show that PG-GAN is not able to model the meteorological movements. More specifically, the first generated frame of PG-GAN looks very close to the first generated frame of our model MPL-GAN, but the later frames are just the expansion of the first frame which is obviously not the real-world scenario. On the other hand, with the constraint of the predictive learning module, our proposed model MPL-GAN continues to generate realistic looking and diverse meteorological frames that catch the real-world meteorological movement pattern. We summarise the findings above in Table 2.

In summary, the quality of meteorological imagery prediction is of crucial importance in weather forecasting, and in monitoring climate change. Figure 3 clearly depicts that MPL-GAN produces a quality prediction result identifying

TABLE 2. Performance comparison.

Model	Sharpness	Learns Met. Patt.
PL with MSE	Blurry	✓
PG-GAN	Sharp	✕
MPL-GAN (ours)	Sharp	✓

both global trends and local variations whilst PL with MSE is too blurry and coarse missing local variations and details. PL with MSE is less useful in practice since it misses many localised patterns, and results in inaccurate predictions, however MPL-GAN is practically useful for forecasting weather and monitoring local and global climate change as evidenced in Figure 3 and Figure 4.

V. CONCLUSION

We propose MPL-GAN to solve the blurry prediction problem of predictive learning methods such as ConvLSTM and TrajGRU. We utilised a conditional GAN to handle this problem by mapping the blurry predictions generated by predictive learning methods back to their original non-blurry data distributions. To do that, we recursively apply a conditional generator conditioning on the previous output of itself and the current output of the predictive learning module. Through the novel design of Frame Discriminator and Flow Discriminator, the generator learns to produce temporally coherent and realistic frames. Experiments on a real-world radar echo dataset demonstrate that our proposed MPL-GAN model not only produces sharp and realistic looking meteorological predictions, but also models the real-world meteorological movement patterns with the constraint of predictive learning module. Although our model is able to generate non-blurry predictions, there is room to improve the prediction accuracy. Since the GAN model brings uncertainties improving the sharpness of prediction, but deteriorates the accuracy, our future work will investigate this problem, and evaluate our proposed model with various real-world datasets.

REFERENCES

[1] X. Shi, Z. Chen, H. Wang, D.-Y. Yeung, W.-K. Wong, and W.-C. Woo, "Convolutional LSTM network: A machine learning approach for precipitation nowcasting," in *Proc. 28th Int. Conf. Neural Inf. Process. Syst.*, Cambridge, MA, USA, vol. 1, 2015, pp. 802–810.

[2] X. Shi, Z. Gao, L. Lausen, H. Wang, D.-Y. Yeung, W.-K. Wong, and W.-C. WOO, "Deep learning for precipitation nowcasting: A benchmark and a new model," in *Neural Information Processing Systems*, I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, R. Garnett, Eds. Nice, France: Curran Associates, 2017, pp. 5617–5627.

[3] Z. M. Yaseen, W. H. M. W. Mohtar, A. M. S. Ameen, I. Ebtehaj, S. F. M. Razali, H. Bonakdari, S. Q. Salih, N. Al-Ansari, and S. Shahid, "Implementation of univariate paradigm for streamflow simulation using hybrid data-driven model: Case study in tropical region," *IEEE Access*, vol. 7, pp. 74471–74481, 2019.

[4] S. J. Hadi, S. I. Abba, S. S. Sammen, S. Q. Salih, N. Al-Ansari, and Z. M. Yaseen, "Non-linear input variable selection approach integrated with non-tuned data intelligence model for streamflow pattern simulation," *IEEE Access*, vol. 7, pp. 141533–141548, 2019.

[5] M. Fu, T. Fan, Z. Ding, S. Q. Salih, N. Al-Ansari, and Z. M. Yaseen, "Deep learning data-intelligence model based on adjusted forecasting window scale: Application in daily streamflow simulation," *IEEE Access*, vol. 8, pp. 32632–32651, 2020.

[6] Z. M. Yaseen, S. O. Sulaiman, R. C. Deo, and K.-W. Chau, "An enhanced extreme learning machine model for river flow forecasting: State-of-the-art, practical applications in water resource engineering area and future research direction," *J. Hydrol.*, vol. 569, pp. 387–408, Feb. 2019.

[7] Z. M. Yaseen, M. Fu, C. Wang, W. H. M. W. Mohtar, R. C. Deo, and A. El-shafie, "Application of the hybrid artificial neural network coupled with rolling mechanism and grey model algorithms for streamflow forecasting over multiple time horizons," *Water Resour. Manage.*, vol. 32, no. 5, pp. 1883–1899, Mar. 2018.

[8] N. D. Bokde, A. Feijoo, N. Al-Ansari, and Z. M. Yaseen, "A comparison between reconstruction methods for generation of synthetic time series applied to wind speed simulation," *IEEE Access*, vol. 7, pp. 135386–135398, 2019.

[9] T. Hai, A. Sharafati, A. Mohammed, S. Q. Salih, R. C. Deo, N. Al-Ansari, and Z. M. Yaseen, "Global solar radiation estimation and climatic variability analysis using extreme learning machine based predictive model," *IEEE Access*, vol. 8, pp. 12026–12042, 2020.

[10] H.-B. Liu and I. Lee, "Bridging the gap between training and inference for spatio-temporal forecasting," in *Proc. 24th Eur. Conf. Artif. Intell.*, 2020.

[11] B. Wang, J. Lu, Z. Yan, H. Luo, T. Li, Y. Zheng, and G. Zhang, "Deep uncertainty quantification: A machine learning approach for weather forecasting," in *Proc. 25th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, New York, NY, US, Jul. 2019, pp. 2087–2095.

[12] M. Rättgers, S. Lee, S. Jeon, and D. You, "Prediction of a typhoon track using a generative adversarial network and satellite images," *Sci. Rep.*, vol. 9, no. 1, Dec. 2019, Art. no. 6057.

[13] M. Mathieu, C. Couprie, and Y. LeCun, "Deep multi-scale video prediction beyond mean square error," in *Proc. 4th Int. Conf. Learn. Represent. (ICLR)*, San Juan, Puerto Rico, May 2016, pp. 1–14.

[14] D. Pathak, P. Krahenbuhl, J. Donahue, T. Darrell, and A. A. Efros, "Context encoders: Feature learning by inpainting," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 2536–2544.

[15] S. Bengio, O. Vinyals, N. Jaitly, and N. Shazeer, "Scheduled sampling for sequence prediction with recurrent neural networks," in *Proc. 28th Int. Conf. Neural Inf. Process. Syst.*, Cambridge, MA, USA, vol. 1, 2015, pp. 1171–1179.

[16] A. Venkatraman, M. Hebert, and J. Andrew Bagnell, "Improving multi-step prediction of learned time series models," in *Proc. 29h AAAI Conf. Artif. Intell.*, 2015, pp. 3024–3030.

[17] Y. Wang, M. Long, J. Wang, Z. Gao, and P. S. Yu, "Predrnn: Recurrent neural networks for predictive learning using spatiotemporal lstms," in *Advances in Neural Information Processing Systems*, I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, R. Garnett, Ed. Nice, France: Curran Associates, 2017, pp. 879–888.

[18] Y. Wang, Z. Gao, M. Long, J. Wang, and P. S. Yu, "PredRNN++: Towards a resolution of the deep-in-time dilemma in spatiotemporal predictive learning," in *Proc. 35th Int. Conf. Mach. Learn.*, Stockholm, Sweden, Jul. 2018, pp. 5123–5132.

[19] J. Ian Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Proc. 27th Int. Conf. Neural Inf. Process. Syst.*, Cambridge, MA, USA, vol. 2, 2014, pp. 2672–2680.

[20] T.-C. Wang, M.-Y. Liu, J.-Y. Zhu, G. Liu, A. Tao, J. Kautz, and B. Catanzaro, "Video-to-video synthesis," in *Advances in Neural Information Processing Systems*, S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, R. Garnett, Eds. Nice, France: Curran Associates, 2018, pp. 1144–1156.

[21] M. Saito and S. Saito, "TGANv2: Efficient training of large models for video generation with multiple subsampling layers," 2018, *arXiv:1811.09245*. [Online]. Available: http://arxiv.org/abs/1811.09245

[22] A. Clark, J. Donahue, and K. Simonyan, "Adversarial video generation on complex datasets," 2019, *arXiv:1907.06571*. [Online]. Available: http://arxiv.org/abs/1907.06571

[23] T. Brox, A. Bruhn, N. Papenberg, and J. Weickert, "High accuracy optical flow estimation based on a theory for warping," in *Computer Vision*. Berlin, Germany: Springer, 2004, pp. 25–36.

[24] P. Cheung and H. Y. Yeung, "Application of optical-flow technique to significant convection nowcast for terminal areas in Hong Kong," in *Proc. 3rd WMO Int. Symp. Nowcasting Very Short-Range Forecasting (WSN)*, 2012, pp. 6–10.

[25] A. Radford, L. Metz, and S. Chintala, ''Unsupervised representation learning with deep convolutional generative adversarial networks,'' in *Proc. 4th Int. Conf. Learn. Represent. (ICLR)*, San Juan, Puerto Rico, 2016, pp. 1–16.

[26] A. Brock, J. Donahue, and K. Simonyan, ''Large scale GAN training for high fidelity natural image synthesis,'' in *Proc. 7th Int. Conf. Learn. Represent. (ICLR)*, New Orleans, LA, USA, May 2019, pp. 1–35.

[27] T. Karras, T. Aila, S. Laine, and J. Lehtinen, ''Progressive growing of GANs for improved quality, stability, and variation,'' in *Int. Conf. Learn. Represent.*, 2018, pp. 1–26.

[28] S. Tulyakov, M.-Y. Liu, X. Yang, and J. Kautz, ''MoCoGAN: Decomposing motion and content for video generation,'' in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 1526–1535.

[29] J. Hyun Lim and J. Chul Ye, ''Geometric GAN,'' 2017, *arXiv:1705.02894*. [Online]. Available: http://arxiv.org/abs/1705.02894

[30] P. J. Werbos, ''Backpropagation through time: What it does and how to do it,'' *Proc. IEEE*, vol. 78, no. 10, pp. 1550–1560, Oct. 1990.

[31] D. P. Kingma and J. Ba, ''Adam: A method for stochastic optimization,'' 2014, *arXiv:1412.6980*. [Online]. Available: http://arxiv.org/abs/1412.6980

ICKJAI LEE received the Ph.D. degree in computer science and software engineering from the University of Newcastle, NSW. He is currently a Professor with the Information Technology Academy, James Cook University, QLD. He is actively involved in teaching and research. His research areas include data mining, geospatial databases, space tessellations, the Internet of Things, geo-visualisation, sentiment analysis, deep learning, and trajectory data mining. He has published more than 150 articles in academic forums and has been involved in a number of research projects.

HONG-BIN LIU received the master's degree in computer science from RMIT, Australia. He is currently pursuing the Ph.D. degree with the College of Science and Engineering, James Cook University, under supervision of Prof. I. Lee.

He worked as a full-stack Developer for five years in China and Australia. He is also a Casual Academic with James Cook University, where he involves in teaching several data science subjects. His research interests include artificial intelligence, spatio-temporal reasoning, computer vision, and so on. He has published several top tier conference papers in the field of data mining and artificial intelligence, including ICDM 2019 and ECAI 2020.