SCIENTIFIC REP**O**RTS

OPEN

Prediction of a typhoon track using a generative adversarial network and satellite images

Mario Rüttgers, **Sangseung Lee**, **Soohwan Jeon & Donghyun You**

Tracks of typhoons are predicted using a generative adversarial network (GAN) with satellite images as inputs. Time series of satellite images of typhoons which occurred in the Korea Peninsula in the past are used to train the neural network. The trained GAN is employed to produce a 6-hour-advance track of a typhoon for which the GAN was not trained. The predicted track image of a typhoon favorably identifies the future location of the typhoon center as well as the deformed cloud structures. Errors between predicted and real typhoon centers are measured quantitatively in kilometers. An averaged error of 95.6 km is achieved for tested 10 typhoons. Predicting sudden changes of the track in westward or northward directions is identified as a challenging task, while the prediction is significantly improved, when velocity fields are employed along with satellite images.

Every year tropical cyclones cause death and damage in many places around the world. Cyclones are formed when water at the sea surface becomes warm, evaporates, and rises in a form of clouds, and while it cools down, the condensation releases strong energy in a form of winds. Rotation of the earth gives a cyclone its spinning motion. At the center usually a hole forms, which is called an eye of the cyclone. At the eye, the pressure is low and energetic clouds and winds get attracted. Warm air can rise through the hole and, favored by high altitude winds that can create a suction effect, increase the energy of the system. Crosswinds, vertical wind shear or dry air, on the other hand, can block this mechanism[1]. Depending on water conditions and surrounding winds tropical cyclones and their associated storm surges can be a great danger when they find their way to a populated land.

The destructive force of a tropical cyclone comes from the wind speed of the rotating air and the rainfall that can cause disastrous flooding[2]. The highest wind speeds are usually found near the typhoon center. Strong rain, on the other hand, can be experienced near dense cloud structures. Therefore, the track of a cyclone is characterized by both the coordinate of the typhoon center and the shape or the distribution of clouds. Sun *et al.*[3] found that tropical cyclones are becoming stronger, larger, and more destructive in the context of global warming. Kossin[4] mentions in his study that translation speeds of tropical-cyclones have decreased globally by 10% and in Asia by 30% over the period 1949–2016. The lower the velocity of the cyclone as a whole system, the more rain falls on one location and the higher the risk for flooding. Thus, slower translation speeds can cause even more destructive floods.

The present work concentrates on tracks of cyclones that form in the north-western Pacific Ocean, also known as typhoons. The geographical focus lies on the Korean peninsula. In the past, Korea has suffered from numerous typhoons of different sizes. Typhoon Sarah was the strongest one. The cyclone hit the island Jeju on September 16th in 1959 during the traditional Thanksgiving festival leaving 655 dead, 259 missing, and more than 750,000 homeless[5]. In the younger history of the country, especially typhoons Rusa (August 2002) and Maemi (September 2003) have caused fear and death among the Korean population. Together they have been responsible for 376 casualties and a damage of 11.5 billion USD[6,7]. Xu *et al.*[8] investigated paths of super typhoons approaching China over the past 50 years. They identified a trend of strong typhoons getting attracted by the moist environment in southeastern China. This trend also affects damages on the Korean peninsula, since the size of a super typhoon can be similar to the spatial distance between the Chinese east coast and the Korean west coast. To save lives and reduce such damages in the future, accurate forecast methods need to be established.

Accuracy is only one criterion when talking about the prediction of natural disasters. Speed and flexibility play a significant role as well. Forecasts should be done quickly and forecast tools should be able to react immediately on sudden changes. Finally they should be inexpensive. Current predictions in South Korea are done by conducting numerical simulations on a Cray XC40 supercomputer with 139,329 CPUs. This method consumes a

Pohang University of Science and Technology, Department of Mechanical Engineering, Pohang, 37673, Korea. Correspondence and requests for materials should be addressed to D.Y. (email: dhyou@postech.ac.kr)

1

huge amount of computational time. The maintenance costs for such expensive hardware are also very high. An efficient forecast method that considers all these criteria is necessary.

In the work of Kovordanyi and Roy[9] nine existing forecast techniques are listed. Subjective assessment is an evaluation of the cyclone's behavior based on large-scale changes of the flow field. In analogue forecasts, features of the cyclone are compared to all previous storms in the same region and the movement is derived from past experiences. In the third technique, the steering current is estimated by analyzing winds at certain locations and altitudes. The statistical technique uses regression analyses, whereas the dynamical technique uses numerical modeling and simulation. Besides computer based approaches, empirical forecasting based on experiences of meteorologists are considered as a good complement to other techniques. The persistence method allows short-term predictions, but relies on the cyclone to keep its recent track. Satellite-based techniques use satellite images to make forecasts of the track and intensity of tropical cyclones based on cloud patterns. Finally, combinations of the previously mentioned approaches are defined as hybrid techniques.

In this study a satellite-based technique is combined with a deep learning method. Lee and Liu[10] probably firstly used satellite images of tropical cyclones for track prediction with the help of neural networks. But images did not function as input data for the network. In a pre-processing step, information like the Dvorak number, the maximum wind speed or the cyclone's position were extracted from images and fed to the network for a time-series prediction. The model has shown improvements of 30% over the bureau numerical tropical cyclone prediction model used in Guam for forecasting tropical cyclone patterns and tracks. Kovordanyi and Roy[9] were the first who actually used satellite images as input data for a neural network. The network favorably detected the shape of a cyclone and predicted the future movement direction. Hong *et al.*[11] utilized multi-layer neural networks to predict a position of the cyclone's eye in a single high-resolution 3D remote sensing image. The network learns coordinates of an eye from labeled images of the past data and predicts them in test images. Kordmahalleh *et al.*[12] have used a sparse recurrent neural network (RNN) to predict trajectories of cyclones coming from the Atlantic Ocean or the Caribbean Sea, also known as hurricanes. They used dynamic time warping (DTW), which compares the trajectory of a target hurricane to all hurricanes of a dataset. For the prediction process, only hurricanes which have similarities in trajectories to that of the target hurricane, were used. One drawback is that they assumed monotonic behaviour, which means that a hurricane never comes back to a location where it was before. In reality this is not always the case. Alemany *et al.*[13] also used an RNN to predict hurricane trajectories, but instead of assuming monotonic behaviour they considered all types of hurricanes. They used a grid-based RNN that takes the wind speed, latitudinal and longitudinal coordinates, the angle of travel, and the grid identification number from past motions of the hurricane as inputs. With the grid identification number, spatial relations on a map were learned. Their forecasts could achieve a better accuracy than results of Kordmahalleh *et al.*[12]. Zhang *et al.*[14] utilized matrix neural networks (MNNs) for predicting trajectories of cyclones that occurred in the South Indian ocean. They stated that MNNs suit better to the task of cyclone trajectory prediction than RNNs, because MNNs can preserve spatial information of cyclone tracks.
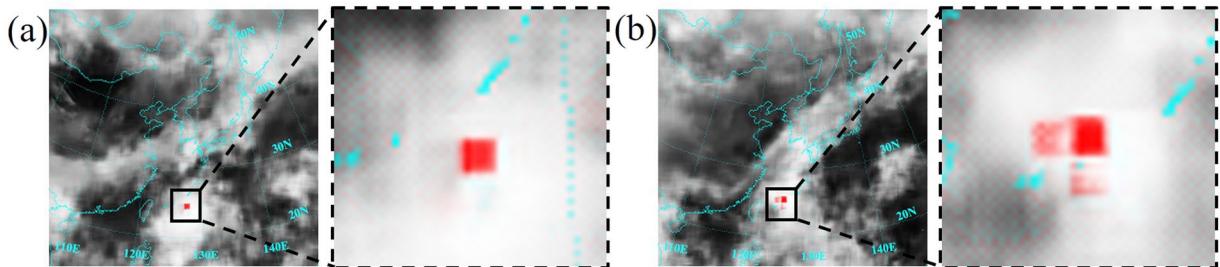
In the previous research, deep learning methods were used to identify the center of a cyclone in a satellite image or to predict a trajectory of a cyclone using discrete meteorological data without cloud images. None of the previous studies using neural-networks reported the capability of predicting both the coordinate of a typhoon center and future shape of cloud structures around the typhoon. But, as mentioned before, a safe warning system should be able to forecast both, cyclone centers and their area of impact in a form of cloud structures. To the best of authors' knowledge, a combination of satellite images and a deep learning method has not yet been used to predict both future typhoon centers and the future cloud appearance. In this study, a generative adversarial network (GAN) is applied for both tasks. As an input it uses chronologically ordered satellite images from the past and as an output it creates images that show the typhoon hours ahead.

A conceivable alternative to the present approach of predicting the coordinate of a typhoon center can be predicting the direction and the speed of a typhoon, from which, afterwards, the future position of the typhoon center is calculated. However, to predict the moving direction and the speed of a typhoon using a GAN, it is still necessary first to forecast the typhoon center, where the direction and the speed of the typhoon are defined. The present GAN employs satellite cloud images along with a marked typhoon center as the input set, but without explicit information about the moving direction or the speeds at the typhoon center.

Predictions can be generated within seconds using a single graphics processing unit (GPU). Thus, the method is significantly quicker and cheaper than the conventional methods, like numerical simulations, where hundreds of thousands CPU are used for one simulation. Tracks of typhoons can be predicted on a real time basis by updating input data as a function of real time. While relearning an updated whole dataset with new inputs is not necessary, training the network with the updated dataset can be conducted on a real time basis by employing a fine-tuning technique or by using multiple GPUs. The GAN is a deep learning technique used to generate samples in forms of images, music or speech. It has been introduced by Goodfellow *et al.*[15], who successfully generated new samples to the MNIST and CIFAR-10 image datasets. Their multi-scale convolution network architecture has originally been proposed by Mathieu *et al.*[16] for image prediction of a video sequence. The inspiration for the present study comes from a problem in fluid dynamics. Lee and You[17] used a GAN to successfully predict unsteady flow over a cylinder. They reported that the adversarial training of the GAN leads the network to generate better prediction of flow fields by extracting flow features in an unsupervised manner, compared to convolutional neural networks.

Results

The GAN has been trained and tested on an NVIDIA Tesla K40c GPU. Ten typhoons in the test dataset are: Faye (Juli 1995), Violet (September 1996), Oliwa (September 1997), Saomai (September 2000), Rammasun (June/July 2002), Maemi (September 2003), Usagi (July/August 2007), Muifa (July/August 2011), Neoguri (June/July 2014), and Malakas (September 2016). The testing for each typhoon is done in sequences. One sequence contains

Figure 1. (**a**) Clearly identifiable predicted typhoon center. (**b**) Three possible predicted typhoon centers.

chronologically ordered input images, the ground truth image, and the generated image. A sequence is named with the date and time (universal time coordinated - UTC) of the corresponding ground truth image, for example: "1993072718" stands for 27th of July in 1993 at 6 pm (UTC).

Prediction of the typhoon center.   The accuracy in the prediction of the typhoon center coordinate is measured by comparing the labeled red square of the ground truth image with the predicted red square of the generated image, rather than coordinates in the Cartesian space. This is because convolutional neural networks have difficulty in learning a mapping between the coordinate in the Cartesian space and the coordinate in the one-hot pixel image space[18]. The red square in the predicted image is detected by applying a color filter on the generated image and a convolution with the size of the red square on the filtered image. The pixel with the highest value in the convolved image gives the location of the predicted typhoon center. In most of the cases images with a clearly identifiable red square are generated, as shown in Fig. 1(a). In some cases, however, the GAN provides several alternatives, as visualized in Fig. 1(b). Each possible predicted typhoon center has its own color strength. The color intensities can be interpreted as probabilities; a strong red square shows a confident prediction, a weak mark stands for a possible predicted typhoon center with a low probability. In cases like in Fig. 1(b), the location with the strongest color intensity is chosen. After identifying the (x, y)-pixel coordinates of the predicted typhoon center, they are transformed back to latitudinal (ϕ) and longitudinal (λ) coordinates via georeferencing.
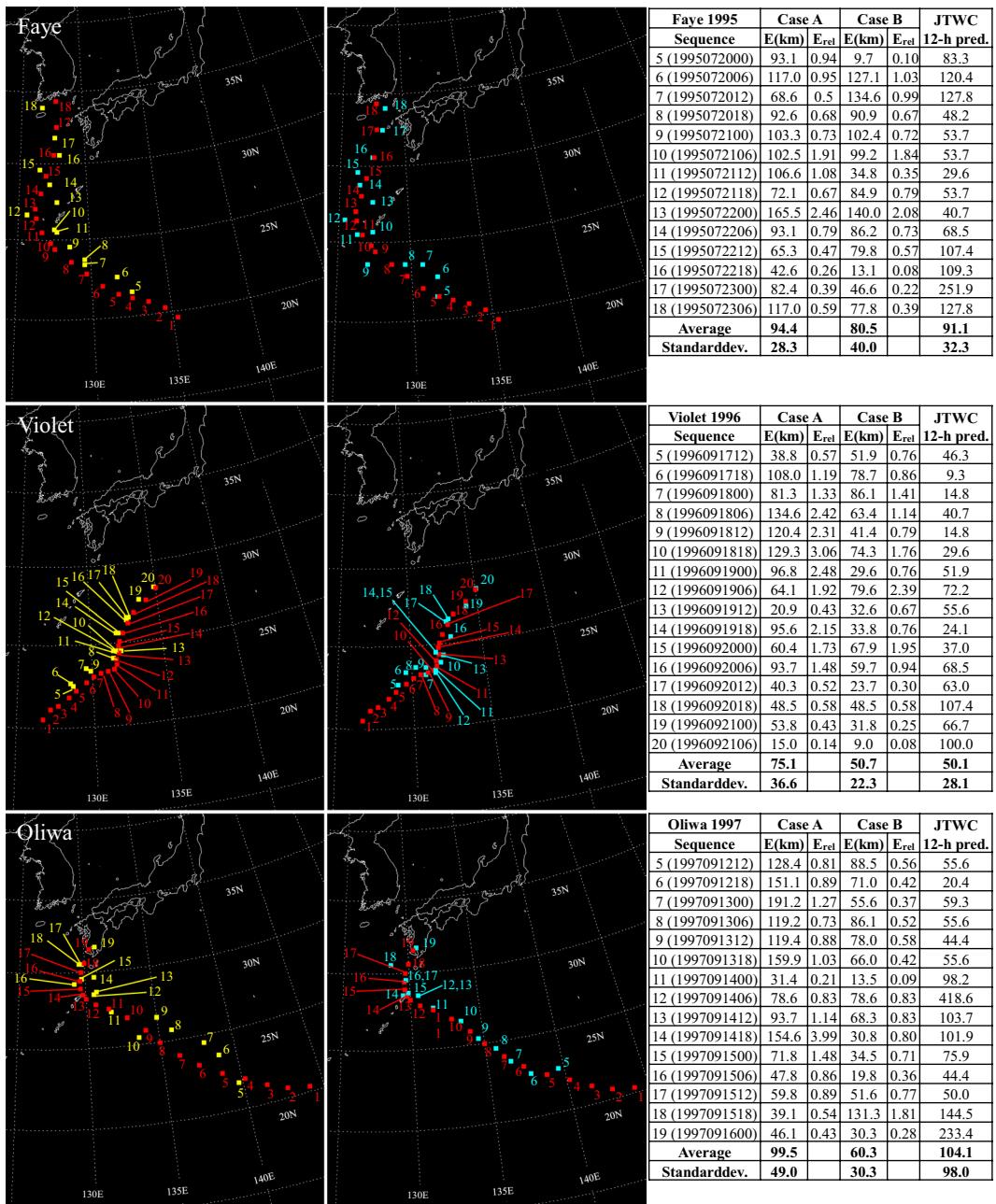
Results for ten test typhoons are presented in Figs 2–5. Images in the left column of the figures are predicted by using only satellite images as input data, and are referred to Case A, where locations of the typhoon center are marked with yellow squares. Images in the right column of the figures are predicted by using satellite images along with velocity fields at 10 m height as input data, and are referred to Case B, where locations of the typhoon center are marked with blue squares. Results in both cases are contrasted to locations of the real typhoon center, represented by red squares. Tables in the figures contain absolute errors, relative errors, and errors of forecasts that have been conducted by the Joint Typhoon Warning Center (JTWC) for typhoons Faye, Violet, Oliwa, Saomai, Rammasun, and Maemi[19–24], and by the Regional Specialized Meteorological Center (RSMC) Tokyo - Typhoon Center for typhoons Usagi, Muifa, Neoguri, and Malakas[25–28].

Focusing firstly on Case A, it is noticeable how errors increase, when typhoons experience sudden northward or westward course changes. It is observed for all ten test typhoons. Figure 2 shows increased errors when typhoon Faye turns northward at sequences 10, 11, and 13, and similarly for Violet at sequences 8–11 and Oliwa at sequences 13–15. In Fig. 3, similar observations are made. When typhoon Saomai changes its path westward at the sequence 15 and northward around the sequence 21 errors increase noticeably. Typhoons Rammasun and Maemi also show difficulty in the prediction when they get deflected northward at sequences 10 and 8–10, respectively. Figures 4 and 5 show results for typhoons Usagi, Muifa, Neoguri, and Malakas. The dragon shaped typhoon Muifa underwent sudden course changes at sequences 16–18 or 24–26. The youngest typhoon among the ten test cases, Malakas, shows the highest error at the sequence 13, when its course has a sudden change to the north.

The path of a typhoon is highly influenced by steering flow that is controlled by storms' ambient environment. Sometimes, when a storm is surrounded by multiple systems whose circulations compete to each other, the steering flow is difficult to predict. As a results, the storms' track is hard to predict. Wu et al.[29] mention that typhoons in the Northwestern Pacific usually experience sudden northward or westward course changes. In the north-turning case, winds are enhanced on the southeast side of tropical cyclones, in west-turning cases, north-easterly winds are enhanced on the west side. One example is illustrated in Fig. 6, which shows satellite images and velocity fields for typhoon Neoguri at sequences 5 (Fig. 6(a)) and 11 (Fig. 6(b)). At the sequence 11, when the typhoon experiences a northward deflection, strong winds are noticed at the southeast side of the center, compared to the velocity field at the sequence 5. It may be difficult for the GAN to learn such phenomena only from satellite images.
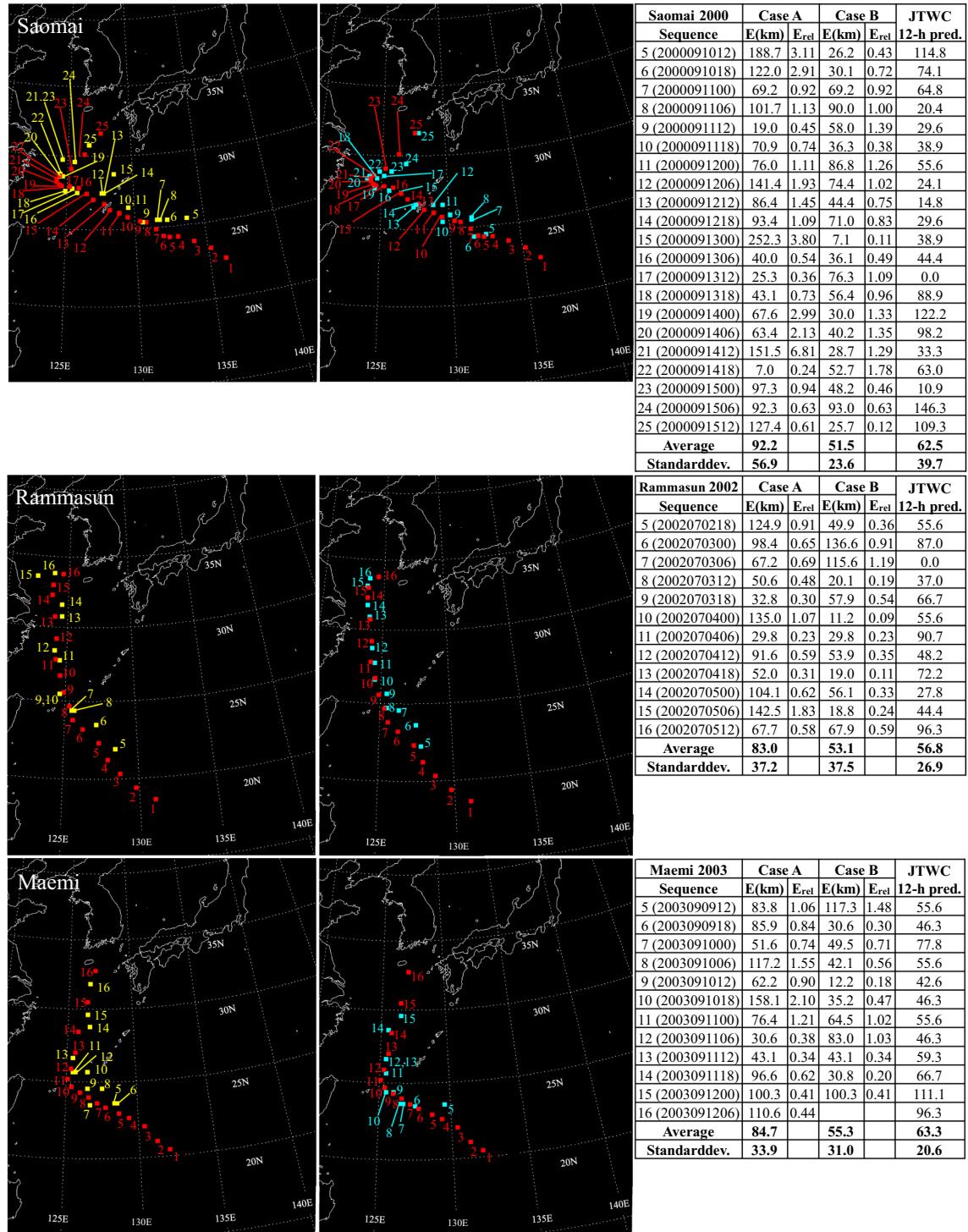
Results in Case B indicate that combining satellite images with images of surface velocity fields has a positive impact on typhoon predictions. Although surface winds cannot represent the complete steering flow, they contain enough information to reduce the average error for all test typhoons, except for typhoon Malakas. Especially, the previously mentioned northward or westward course changes are predicted much more accurately in Case B. Except for results for typhoons Faye and Rammasun, standard deviations of the error are also reduced for all test typhoons. Therefore, adding information of the velocity field to the network seems to be beneficial to improving the predictive capability of the present GAN-based method.

Several findings are gained by analyzing errors in predictions conducted by the JTWC and the RSMC. Firstly, it is observed that for typhoons Faye, Oliwa, Saomai, Rammasun, Maemi, and Muifa, errors in Case B are lower

Faye 1995 Sequence	Case A E(km)	E_{rel}	Case B E(km)	E_{rel}	JTWC 12-h pred.
5 (1995072000)	93.1	0.94	9.7	0.10	83.3
6 (1995072006)	117.0	0.95	127.1	1.03	120.4
7 (1995072012)	68.6	0.5	134.6	0.99	127.8
8 (1995072018)	92.6	0.68	90.9	0.67	48.2
9 (1995072100)	103.3	0.73	102.4	0.72	53.7
10 (1995072106)	102.5	1.91	99.2	1.84	53.7
11 (1995072112)	106.6	1.08	34.8	0.35	29.6
12 (1995072118)	72.1	0.67	84.9	0.79	53.7
13 (1995072200)	165.5	2.46	140.0	2.08	40.7
14 (1995072206)	93.1	0.79	86.2	0.73	68.5
15 (1995072212)	65.3	0.47	79.8	0.57	107.4
16 (1995072218)	42.6	0.26	13.1	0.08	109.3
17 (1995072300)	82.4	0.39	46.6	0.22	251.9
18 (1995072306)	117.0	0.59	77.8	0.39	127.8
Average	94.4		80.5		91.1
Standarddev.	28.3		40.0		32.3

Violet 1996 Sequence	Case A E(km)	E_{rel}	Case B E(km)	E_{rel}	JTWC 12-h pred.
5 (1996091712)	38.8	0.57	51.9	0.76	46.3
6 (1996091718)	108.0	1.19	78.7	0.86	9.3
7 (1996091800)	81.3	1.33	86.1	1.41	14.8
8 (1996091806)	134.6	2.42	63.4	1.14	40.7
9 (1996091812)	120.4	2.31	41.4	0.79	14.8
10 (1996091818)	129.3	3.06	74.3	1.76	29.6
11 (1996091900)	96.8	2.48	29.6	0.76	51.9
12 (1996091906)	64.1	1.92	79.6	2.39	72.2
13 (1996091912)	20.9	0.43	32.6	0.67	55.6
14 (1996091918)	95.6	2.15	33.8	0.76	24.1
15 (1996092000)	60.4	1.73	67.9	1.95	37.0
16 (1996092006)	93.7	1.48	59.7	0.94	68.5
17 (1996092012)	40.3	0.52	23.7	0.30	63.0
18 (1996092018)	48.5	0.58	48.5	0.58	107.4
19 (1996092100)	53.8	0.43	31.8	0.25	66.7
20 (1996092106)	15.0	0.14	9.0	0.08	100.0
Average	75.1		50.7		50.1
Standarddev.	36.6		22.3		28.1

Oliwa 1997 Sequence	Case A E(km)	E_{rel}	Case B E(km)	E_{rel}	JTWC 12-h pred.
5 (1997091212)	128.4	0.81	88.5	0.56	55.6
6 (1997091218)	151.1	0.89	71.0	0.42	20.4
7 (1997091300)	191.2	1.27	55.6	0.37	59.3
8 (1997091306)	119.2	0.73	86.1	0.52	55.6
9 (1997091312)	119.4	0.88	78.0	0.58	44.4
10 (1997091318)	159.9	1.03	66.0	0.42	55.6
11 (1997091400)	31.4	0.21	13.5	0.09	98.2
12 (1997091406)	78.6	0.83	78.6	0.83	418.6
13 (1997091412)	93.7	1.14	68.3	0.83	103.7
14 (1997091418)	154.6	3.99	30.8	0.80	101.9
15 (1997091500)	71.8	1.48	34.5	0.71	75.9
16 (1997091506)	47.8	0.86	19.8	0.36	44.4
17 (1997091512)	59.8	0.89	51.6	0.77	50.0
18 (1997091518)	39.1	0.54	131.3	1.81	144.5
19 (1997091600)	46.1	0.43	30.3	0.28	233.4
Average	99.5		60.3		104.1
Standarddev.	49.0		30.3		98.0

Figure 2. For typhoons Faye, Violet, and Oliwa, predictions using satellite images only (Case A) are shown in the left column with yellow squares, and predictions using satellite images along with velocity fields at 10 m height (Case B) are shown in the right column with blue squares. For both cases, center coordinates of the real typhoon are represented by red squares. Absolute errors, relative errors, and errors in kilometers for 12-hour predictions reported in annual tropical cyclone reports of the Joint Typhoon Warning Center (JTWC)[19–21] as a function of the sequence are summarized in tables.

than errors of 12-hour predictions conducted by the JTWC and 24-hour forecasts of the RSMC. Secondly, six of the ten test typhoons show lower standard deviations for results in Case B compared to predictions reported by the JTWC and the RSMC. This may be partly due to the shorter prediction interval (6-hours) in the present study than those of the JTWC and the RSMC. Thirdly, whereas sudden northward or westward deflections are the challenging regime in Case A, forecasts conducted by the JTWC and RSMC show no increased errors when sudden course changes occur. However, their challenging regime seems to be interactions between storms and land, since for all ten typhoons errors increase before landfall. Difficulties in predictions before landfall are noticeable in Case A for typhoons Faye, Saomai, Rammasun, Maemi, Muifa, and especially Neoguri. In Case B, these difficulties are handled much better by the GAN.

Saomai 2000 Sequence	Case A E(km)	Case A E_{rel}	Case B E(km)	Case B E_{rel}	JTWC 12-h pred.
5 (2000091012)	188.7	3.11	26.2	0.43	114.8
6 (2000091018)	122.0	2.91	30.1	0.72	74.1
7 (2000091100)	69.2	0.92	69.2	0.92	64.8
8 (2000091106)	101.7	1.13	90.0	1.00	20.4
9 (2000091112)	19.0	0.45	58.0	1.39	29.6
10 (2000091118)	70.9	0.74	36.3	0.38	38.9
11 (2000091200)	76.0	1.11	86.8	1.26	55.6
12 (2000091206)	141.4	1.93	74.4	1.02	24.1
13 (2000091212)	86.4	1.45	44.4	0.75	14.8
14 (2000091218)	93.4	1.09	71.0	0.83	29.6
15 (2000091300)	252.3	3.80	7.1	0.11	38.9
16 (2000091306)	40.0	0.54	36.1	0.49	44.4
17 (2000091312)	25.3	0.36	76.3	1.09	0.0
18 (2000091318)	43.1	0.73	56.4	0.96	88.9
19 (2000091400)	67.6	2.99	30.0	1.33	122.2
20 (2000091406)	63.4	2.13	40.2	1.35	98.2
21 (2000091412)	151.5	6.81	28.7	1.29	33.3
22 (2000091418)	7.0	0.24	52.7	1.78	63.0
23 (2000091500)	97.3	0.94	48.2	0.46	10.9
24 (2000091506)	92.3	0.63	93.0	0.63	146.3
25 (2000091512)	127.4	0.61	25.7	0.12	109.3
Average	**92.2**		**51.5**		**62.5**
Standarddev.	**56.9**		**23.6**		**39.7**

Rammasun 2002 Sequence	Case A E(km)	Case A E_{rel}	Case B E(km)	Case B E_{rel}	JTWC 12-h pred.
5 (2002070218)	124.9	0.91	49.9	0.36	55.6
6 (2002070300)	98.4	0.65	136.6	0.91	87.0
7 (2002070306)	67.2	0.69	115.6	1.19	0.0
8 (2002070312)	50.6	0.48	20.1	0.19	37.0
9 (2002070318)	32.8	0.30	57.9	0.54	66.7
10 (2002070400)	135.0	1.07	11.2	0.09	55.6
11 (2002070406)	29.8	0.23	29.8	0.23	90.7
12 (2002070412)	91.6	0.59	53.9	0.35	48.2
13 (2002070418)	52.0	0.31	19.0	0.11	72.2
14 (2002070500)	104.1	0.62	56.1	0.33	27.8
15 (2002070506)	142.5	1.83	18.8	0.24	44.4
16 (2002070512)	67.7	0.58	67.9	0.59	96.3
Average	**83.0**		**53.1**		**56.8**
Standarddev.	**37.2**		**37.5**		**26.9**

Maemi 2003 Sequence	Case A E(km)	Case A E_{rel}	Case B E(km)	Case B E_{rel}	JTWC 12-h pred.
5 (2003090912)	83.8	1.06	117.3	1.48	55.6
6 (2003090918)	85.9	0.84	30.6	0.30	46.3
7 (2003091000)	51.6	0.74	49.5	0.71	77.8
8 (2003091006)	117.2	1.55	42.1	0.56	55.6
9 (2003091012)	62.2	0.90	12.2	0.18	42.6
10 (2003091018)	158.1	2.10	35.2	0.47	46.3
11 (2003091100)	76.4	1.21	64.5	1.02	55.6
12 (2003091106)	30.6	0.38	83.0	1.03	46.3
13 (2003091112)	43.1	0.34	43.1	0.34	59.3
14 (2003091118)	96.6	0.62	30.8	0.20	66.7
15 (2003091200)	100.3	0.41	100.3	0.41	111.1
16 (2003091206)	110.6	0.44			96.3
Average	**84.7**		**55.3**		**63.3**
Standarddev.	**33.9**		**31.0**		**20.6**

Figure 3. For typhoons Saomai, Rammasun, and Maemi, predictions using satellite images only (Case A) are shown in the left column with yellow squares, and predictions using satellite images along with velocity fields at 10 m height (Case B) are shown in the right column with blue squares. For both cases, center coordinates of the real typhoon are represented by red squares. Absolute errors, relative errors, and errors in kilometers for 12-hour predictions reported in annual tropical cyclone reports of the Joint Typhoon Warning Center (JTWC)[22–24] as a function of the sequence are summarized in tables.
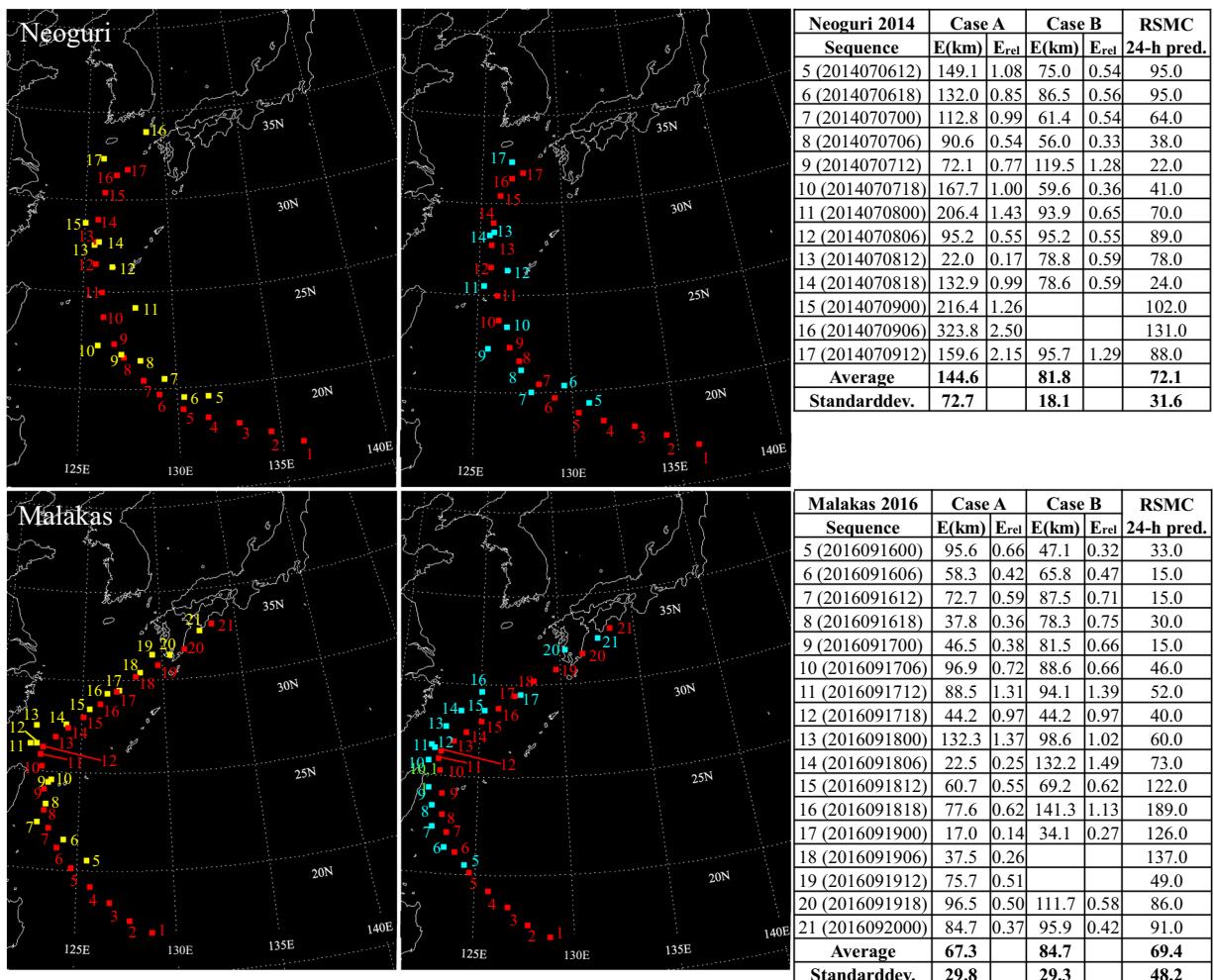
Table 1 compares the frequency distribution of *E* of Cases A and B for all sequences. For Case A, in 74.5% of the cases the error is less than or equal to 120 km, compared to 92.5% in Case B. In Case A, a combined 42.4% have an accuracy of less than or equal to 80 km, which could be increased to 64.8% in Case B. In Case B, only 12 sequences are found above 120 km. The total error in Case A, 95.6 km, could be reduced by 27.7% to 69.1 km in Case B.

Usagi 2007	Case A		Case B		RSMC
Sequence	E(km)	E_{rel}	E(km)	E_{rel}	24-h pred.
5 (2007073018)	163.4	1.32	148.6	1.20	15.0
6 (2007073100)	72.6	1.05	101.6	1.47	11.0
7 (2007073106)	129.2	1.19	78.0	0.72	79.0
8 (2007073112)	108.5	0.95	26.8	0.23	76.0
9 (2007073118)	200.9	1.41	86.1	0.60	67.0
10 (2007080100)	78.5	0.61	49.8	0.39	98.0
11 (2007080106)	199.5	1.16	98.3	0.57	152.0
12 (2007080112)	79.8	0.48	120.1	0.73	150.0
13 (2007080118)	100.5	0.46	175.1	0.80	128.0
14 (2007080200)	103.7	0.54	106.8	0.55	152.0
15 (2007080206)	109.3	0.69	109.3	0.69	45.0
16 (2007080212)	58.1	0.41	92.2	0.65	73.0
17 (2007080218)	65.1	0.45	175.4	1.21	83.0
18 (2007080300)	39.9	0.21	98.8	0.51	69.0
Average	107.8		104.8		85.6
Standard dev.	48.2		40.1		44.8

Muifa 2011	Case A		Case B		RSMC
Sequence	E(km)	E_{rel}	E(km)	E_{rel}	24-h pred.
5 (2011073106)	161.6	3.51	87.0	1.89	124.0
6 (2011073112)	188.7	2.24	43.1	0.51	100.0
7 (2011073118)	98.9	1.11	98.9	1.11	106.0
8 (2011080100)	64.0	1.30	64.0	1.30	137.0
9 (2011080106)	61.6	0.57	92.7	0.85	119.0
10 (2011080112)	118.8	1.33	17.3	0.19	99.0
11 (2011080118)	61.5	0.69	23.7	0.27	56.0
12 (2011080200)	56.5	0.85	56.5	0.85	83.0
13 (2011080206)	104.4	1.34	84.0	1.08	108.0
14 (2011080212)	77.0	0.98	77.0	0.98	117.0
15 (2011080218)	73.6	0.98	90.5	1.20	107.0
16 (2011080300)	137.2	2.04	79.0	1.17	84.0
17 (2011080306)	49.4	0.69	49.4	0.69	46.0
18 (2011080312)	155.6	2.17	62.0	0.86	52.0
19 (2011080318)	86.3	0.94	39.0	0.43	102.0
20 (2011080400)	88.6	0.95	43.6	0.47	42.0
21 (2011080406)	65.7	1.63	102.8	2.54	75.0
22 (2011080412)	99.6	1.81	52.1	0.94	55.0
23 (2011080418)	76.7	1.48	95.5	1.85	64.0
24 (2011080500)	165.6	3.16	41.0	0.78	67.0
25 (2011080506)	65.1	0.97	87.6	1.30	75.0
26 (2011080512)	141.2	2.63	19.5	0.36	75.0
27 (2011080518)	64.8	0.78	157.9	1.91	100.0
28 (2011080600)	168.6	1.28	94.3	0.72	108.0
29 (2011080606)	166.8	1.50	40.8	0.37	83.0
30 (2011080612)	83.5	0.80	73.3	0.70	66.0
31 (2011080618)	83.2	0.66	34.8	0.28	112.0
32 (2011080700)	158.0	1.00	27.5	0.17	115.0
33 (2011080706)	81.7	0.55	46.8	0.32	86.0
34 (2011080712)	205.2	1.31	164.5	1.05	93.0
Average	107.0		68.2		88.5
Standard dev.	44.7		35.5		24.7

Figure 4. For typhoons Usagi, and Muifa, predictions using satellite images only (Case A) are shown in the left column with yellow squares, and predictions using satellite images along with velocity fields at 10 m height (Case B) are shown in the right column with blue squares. For both cases, center coordinates of the real typhoon are represented by red squares. Absolute errors, relative errors, and errors in kilometers for 24-hour predictions reported in annual reports on the activities of the Regional Specialized Meteorological Center (RSMC) Tokyo - Typhoon Center[25,26] as a function of the sequence are summarized in tables.

Prediction of the cloud shape. As an example for the generation of images that show the cloud shape, Fig. 7 illustrates the prediction for the sequence 6 of typhoon Maemi. Generated images do not have the same sharpness than the ground truth image, instead they are blurry. Blurriness is a well known challenge in video prediction tasks, which is tried to be overcome by using a gradient difference loss function in the present study. However, although generated images suffer from a certain degree of blurriness, the main structure of clouds is still visible. Furthermore, in the generated image in Case B the spinning motion of the typhoon seems to be reproduced more realistically than in the generated image in Case A.

This is shown clearer in Fig. 8, where generated cloud images in Cases A and B of four randomly chosen sequences are presented. Generated images in Case A seem to be static and do not represent any dynamics. In the sequence 11 of typhoon Violet, for example, the image in Case A cannot reproduce the spinning motion of the typhoon (see Fig. 8(a)). The image in Case B, on the contrary, does not only illustrate the typhoon more

Neoguri 2014	Case A		Case B		RSMC
Sequence	E(km)	E_{rel}	E(km)	E_{rel}	24-h pred.
5 (2014070612)	149.1	1.08	75.0	0.54	95.0
6 (2014070618)	132.0	0.85	86.5	0.56	95.0
7 (2014070700)	112.8	0.99	61.4	0.54	64.0
8 (2014070706)	90.6	0.54	56.0	0.33	38.0
9 (2014070712)	72.1	0.77	119.5	1.28	22.0
10 (2014070718)	167.7	1.00	59.6	0.36	41.0
11 (2014070800)	206.4	1.43	93.9	0.65	70.0
12 (2014070806)	95.2	0.55	95.2	0.55	89.0
13 (2014070812)	22.0	0.17	78.8	0.59	78.0
14 (2014070818)	132.9	0.99	78.6	0.59	24.0
15 (2014070900)	216.4	1.26			102.0
16 (2014070906)	323.8	2.50			131.0
17 (2014070912)	159.6	2.15	95.7	1.29	88.0
Average	144.6		81.8		72.1
Standarddev.	72.7		18.1		31.6

Malakas 2016	Case A		Case B		RSMC
Sequence	E(km)	E_{rel}	E(km)	E_{rel}	24-h pred.
5 (2016091600)	95.6	0.66	47.1	0.32	33.0
6 (2016091606)	58.3	0.42	65.8	0.47	15.0
7 (2016091612)	72.7	0.59	87.5	0.71	15.0
8 (2016091618)	37.8	0.36	78.3	0.75	30.0
9 (2016091700)	46.5	0.38	81.5	0.66	15.0
10 (2016091706)	96.9	0.72	88.6	0.66	46.0
11 (2016091712)	88.5	1.31	94.1	1.39	52.0
12 (2016091718)	44.2	0.97	44.2	0.97	40.0
13 (2016091800)	132.3	1.37	98.6	1.02	60.0
14 (2016091806)	22.5	0.25	132.2	1.49	73.0
15 (2016091812)	60.7	0.55	69.2	0.62	122.0
16 (2016091818)	77.6	0.62	141.3	1.13	189.0
17 (2016091900)	17.0	0.14	34.1	0.27	126.0
18 (2016091906)	37.5	0.26			137.0
19 (2016091912)	75.7	0.51			49.0
20 (2016091918)	96.5	0.50	111.7	0.58	86.0
21 (2016092000)	84.7	0.37	95.9	0.42	91.0
Average	67.3		84.7		69.4
Standarddev.	29.8		29.3		48.2

Figure 5. For typhoons Neoguri, and Malakas, predictions using satellite images only (Case A) are shown in the left column with yellow squares, and predictions using satellite images along with velocity fields at 10 m height (Case B) are shown in the right column with blue squares. For both cases, center coordinates of the real typhoon are represented by red squares. Absolute errors, relative errors, and errors in kilometers for 24-hour predictions taken from annual reports on the activities of the Regional Specialized Meteorological Center (RSMC) Tokyo - Typhoon Center[27,28] as a function of the sequence are summarized in tables.

realistically, but also cloud patterns in the remaining parts of the image. In the sequence 6 of typhoon Rammasun, the prediction in Case B again resolves the spinning motion of the typhoon much better than the prediction in Case A (see Fig. 8(b)). Details, like the cloud structure east or north of the typhoon center, are generated much more reliably. In the sequence 15 of typhoon Maemi the cyclone has almost reached the Korean peninsula. Parts of the cloud structure start to dissipate, but a significant part follows a strong westerly wind at a high altitude, called jet stream. In Fig. 8(c) it can be noticed how the result in Case B reproduces the suction effect of the jet stream in a smoother way than the predicted image in Case A. Finally, in Fig. 8(d) the image in Case B highlights much better how surrounding clouds move into the eyewall than the image in Case A.

Discussion

The application of a deep learning method for typhoon track prediction in forms of typhoon center coordinates and cloud structures has been explored. Learning only from satellite images, 42.4% of all typhoon center predictions have absolute errors of less than 80 km, 32.1% lie within a range of 80–120 km and the remaining 25.5% have an accuracy above 120 km. The averaged error lies at 95.6 km. In general, errors increase when typhoons undergo sudden northward or westward course changes. Predictions in this challenging regime get much more accurate, when satellite images are combined with the velocity field at 10 m height. In that case, 64.8% of all predictions are below 80 km, 27.7% lie within a range of 80–120 km and only 7.5% remain with an accuracy above 120 km. The averaged error could be reduced by 27.7% to 69.1 km. Furthermore, it has been shown that the GAN is able to generate images that reproduce cloud appearance.

From a user's perspective, it is helpful to make predictions in a quicker and cheaper manner. Current predictions in many countries rely on highly expensive numerical simulations using the state-of-the-art supercomputers, of which acquisition is limited by few advanced countries. Furthermore, it is expected that the present deep-learning-based method can be utilized in combination with other conventional techniques, especially, for

Figure 6. Satellite images and velocity fields at 10 m height of typhoon Neoguri at sequences 5 (**a**) and 11 (**b**). Black arrows indicate directions of positive zonal and meridional velocity components. (Source of satellite images: Korean Meteorological Administration (KMA). http://www.weather.go.kr/weather/images).

Error (km)	Case A	Percentage (%)	Case B	Percentage (%)
0–40	18	10.9	41	25.9
41–80	52	31.5	62	38.9
81–120	53	32.1	44	27.7
≥121	42	25.5	12	7.5
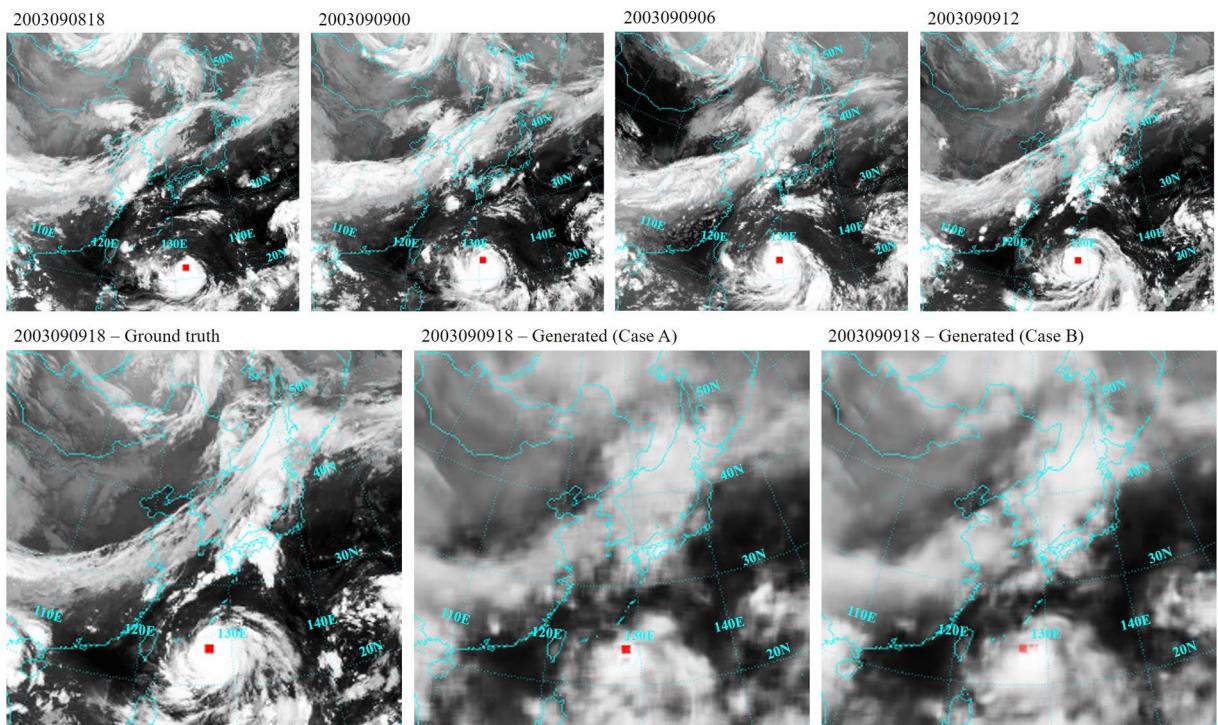
Table 1. Frequency distribution for the prediction of all sequences for Case A (Only satellite images) and Case B (Satellite images combined with data of the velocity field at 10 m height).

non-user-biased analysis and decision making processes. Although, in the present study, predictions are conducted for a 6-hour interval, extension of the prediction time-interval is straightforward once the acquisition of more satellite images is possible. To improve the prediction accuracy, the next step will be adding physical information to the input data, like the sea surface temperature, the surface pressure, and velocity fields at various heights. Just learning from satellite images that show the cloud structure and the typhoon center is a good starting point but not sufficient for learning whole complex phenomena that are responsible for the creation and motion of typhoons.

Methods

Satellite images.   Input satellite images, which were captured by satellites at the altitude of 35,786 km, have been provided by the Korean Meteorological Administration (KMA)[30]. They contain 76 typhoons from 1993 till 2017 that hit or were about to hit the Korean peninsula. Whereas Hong *et al.*[11] used satellite images directly, in this study pre-processing of the images is inevitable. During the 25 years of capture time, different satellites have been operating. Thus, raw images have different perspectives on the Korean peninsula. Three different types of image perspectives and pixel sizes are provided (see Fig. 9). The network takes images with the same pixel size and learns better if all images have the same or a similar perspective on the north-western Pacific Ocean.

Figure 9(a) shows an image of typhoon Rusa in August 2002 as an example for the first perspective of the first period from 1993 till 2010. The images have a pixel size of 512 × 512. Initially, the images have been generated by the Geostationary Meteorological Satellite (GMS) number 4 of Japan[31]. The GMS-4 satellite was replaced by the GMS-5 satellite in 1996 with a design life of five years. But the replacement failed and the GMS-5 had to operate for three more years, before the US National Oceanic & Atmospheric Administration (NOAA) helped out with its GOES-9 Satellite in 2003. Two years later Japan could use the Multi-functional Transport Satellite (MTSAT) type 1 R until 2010. All satellites used the Visible and Infrared Spin-Scan Radiometer (VISSR) technique to obtain

Figure 7. Prediction of the sequence 6 (2003090918) for typhoon Maemi (Source of satellite images: Korean Meteorological Administration (KMA). http://www.weather.go.kr/weather/images).

visible and infrared spectrum mappings of the earth and its cloud cover. The upper black bar, containing the satellite name, the date and the time written in yellow, as well as the black frame are unnecessary information for the learning phase. Furthermore, the pixel size causes memory issues in the test phase. Thus, images have been cropped and resized to 250×238 pixels (see Fig. 9(d)).
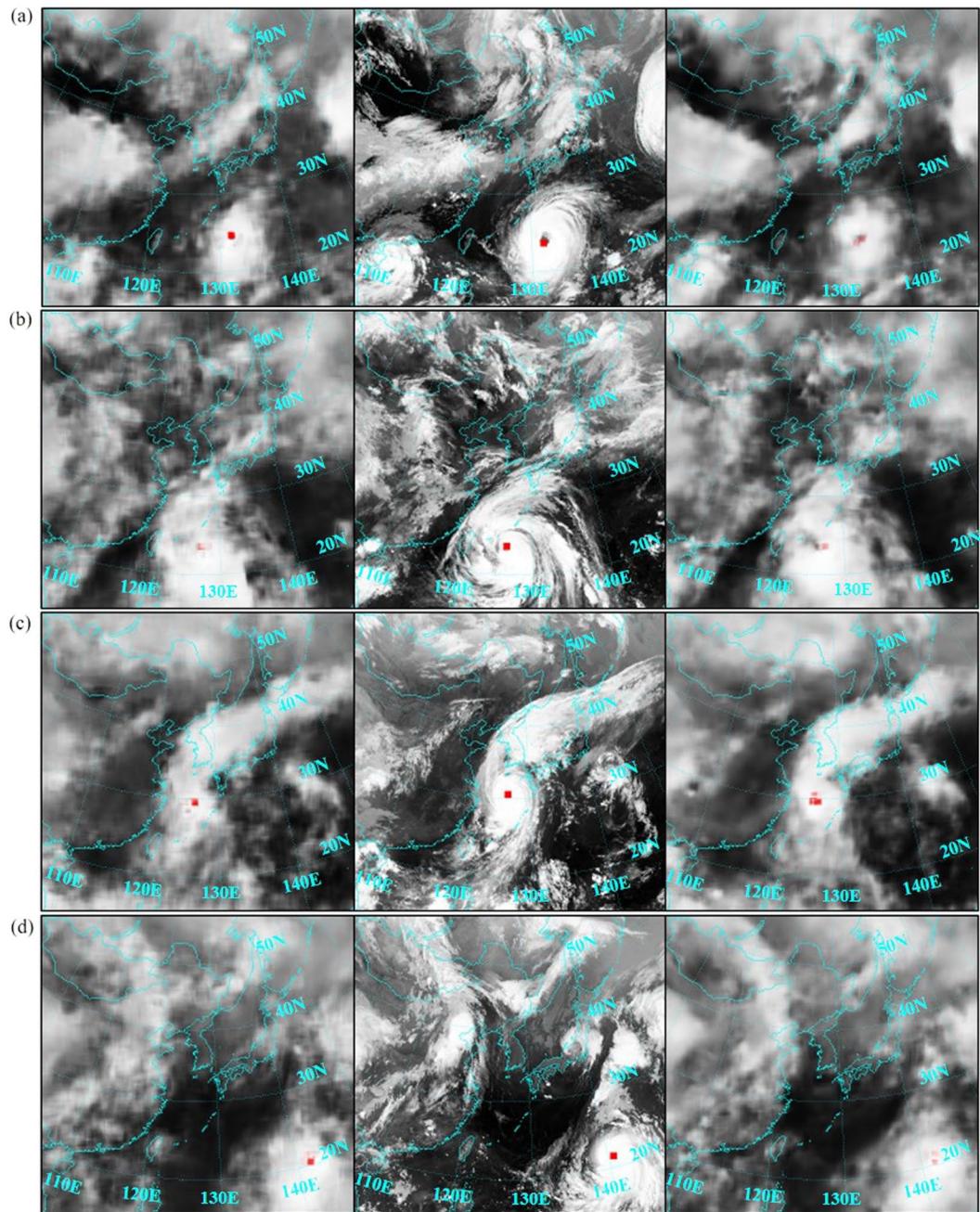
The perspective of images captured between 2011 and 2014 is presented in Fig. 9(b). The image with the pixel size of 512×412 shows typhoon Neoguri that formed in July 2014. It has been captured by Korea's first multi-purpose geostationary meteorological satellite, namely Communication, Ocean and Meteorological Satellite (COMS). As in the previous case, the upper black bar is cropped and the images are resized to 250×238 pixels (see Fig. 9(e)). The view on the earth in Fig. 9(e) is similar to the view in Fig. 9(d).

An example for raw images of the remaining data is shown in Fig. 9(c). The operating satellite is the same like in the previous case. The 512×433 pixel shot shows typhoon Goni from August 2015. Except for a different thickness of the upper black bar, the pre-processing steps match with the previous case and lead to images like in Fig. 9(f).

All images have the 'png' format with three color channels (R(red)G(green)B(blue)). To improve the visibility of the country boarders, their color has been changed from yellow to blue. In total 1,628 images are stored, with a time step size of 6 hours between images. There are two accuracy criteria for the typhoon track prediction, a quantitative and a qualitative criteria. In the quantitative criterion, the difference between the coordinate of the predicted typhoon center and that of the ground truth is taken into consideration. In the qualitative criterion, the predicted shape of clouds and the shape of the ground truth cloud are compared.

Every satellite image has been labeled with a red square at the typhoon center. The latitudinal (ϕ) and the longitudinal (λ) coordinates of the typhoon centers are provided by the Japan Meteorological Agency (JMA)[32]. In order to label each satellite image with its red square, ϕ and λ coordinates have to be transfered to (x, y)-pixel coordinates in the image. This process is known as georeferencing and is illustrated in Fig. 10. Images are split into training and test data. The training data contain 1,389 images of 66 typhoons, the test data are 239 images of 10 typhoons.
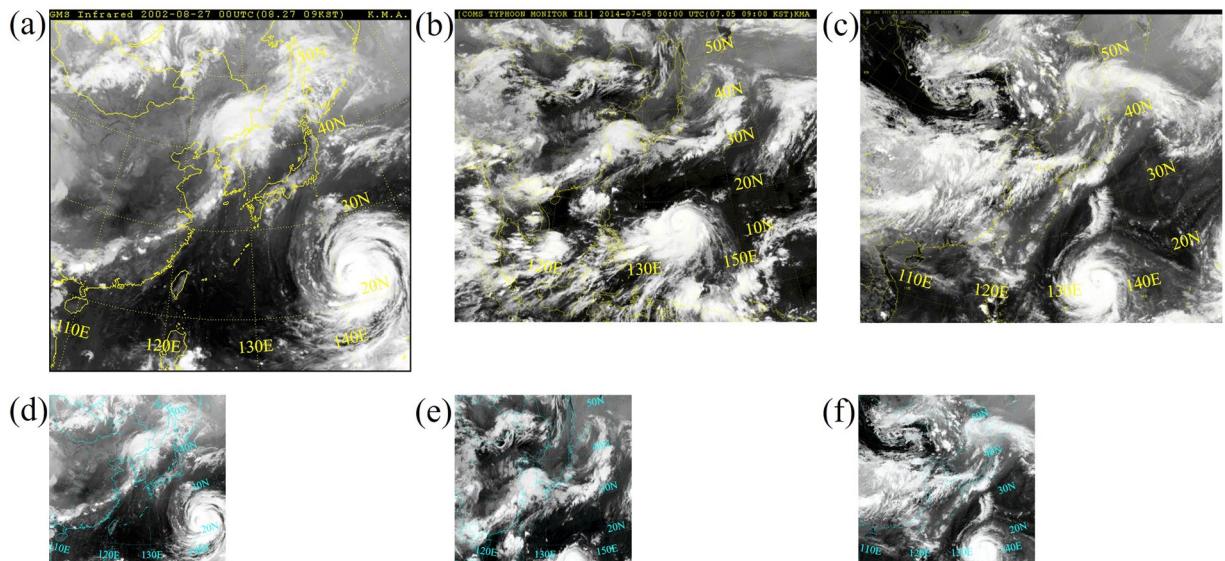
Velocity fields.   In order to help the GAN to better learn the movement of typhoons, the public dataset ERA-interim[33] is employed. ERA-interim is a global atmospheric reanalysis starting from 1979 that has continuously been updated until today. It uses a fixed version of a numerical weather prediction system (IFS - Cy31r2) to produce reanalyzed data. The available data have time steps of 6 hours and their dates fit to the satellite images described in the previous section. Although ERA-interim data cover the whole globe, for this work only the area around the Korean peninsula is selected. Raw data with grid resolution of 0.75 degrees are refined to 0.125 degrees by applying linear interpolation, which leads to a resolution of nearly 13.8 km[34]. Images from reanalysis data are generated with the software Panoply and the map type Lambert Conformal Conic to match the view of the satellite images[35]. In the present study, reanalysis data are used by the GAN to learn information about the surface velocity field at 10 m height. Figure 11 gives an example for typhoon Maemi. The velocity field in Fig. 11(b) is
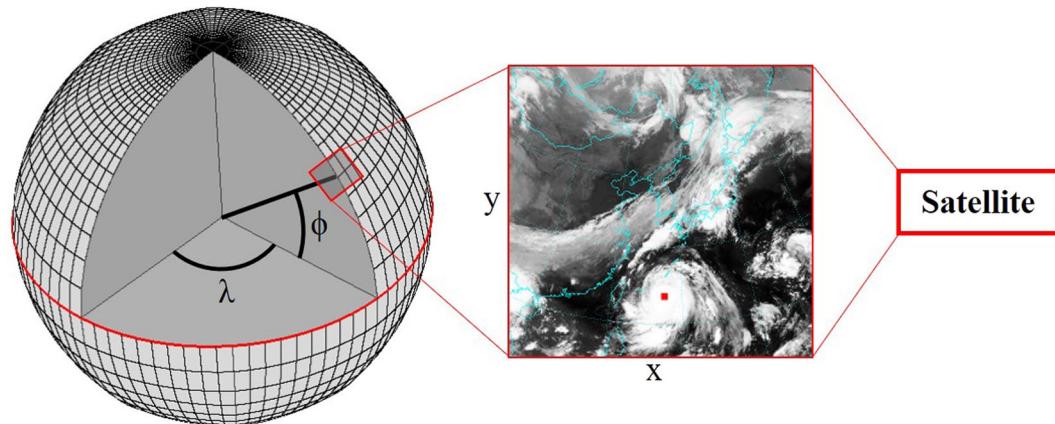
Figure 8. Generated images by using satellite images as input (Case A) in the left-most column, and images by using satellite images along with velocity fields at 10 m height (Case B) in the right-most column, compared to corresponding ground truth images in the middle column: (**a**) Sequence 11 of typhoon Violet, (**b**) sequence 6 of typhoon Rammasun, (**c**) sequence 15 of typhoon Maemi and (**d**) sequence 5 of typhoon Usagi (Source of satellite images: Korean Meteorological Administration (KMA). http://www.weather.go.kr/weather/images).

contrasted with the satellite image in Fig. 11(a). Color changes in the zonal and meridional velocity components located in the white circles indicate a spinning motion of the typhoon.

Deep learning methodology. The deep learning methodology is different for the training and the testing steps. Figure 12(a) shows the concept of adversarial training. For simplicity, a case is shown where only satellite images function as input. As mentioned previously in this work, each full scale satellite image consists of three color channels (RGB). They are represented throughout Fig. 12 by three overlapped layers. Before using the training data as input images to the GAN they get cropped to a total number of 5,000,000 clips with a pixel size of 32×32. One clip contains a set of m consecutively cropped satellite images from the past (input, I) and one ground truth image.
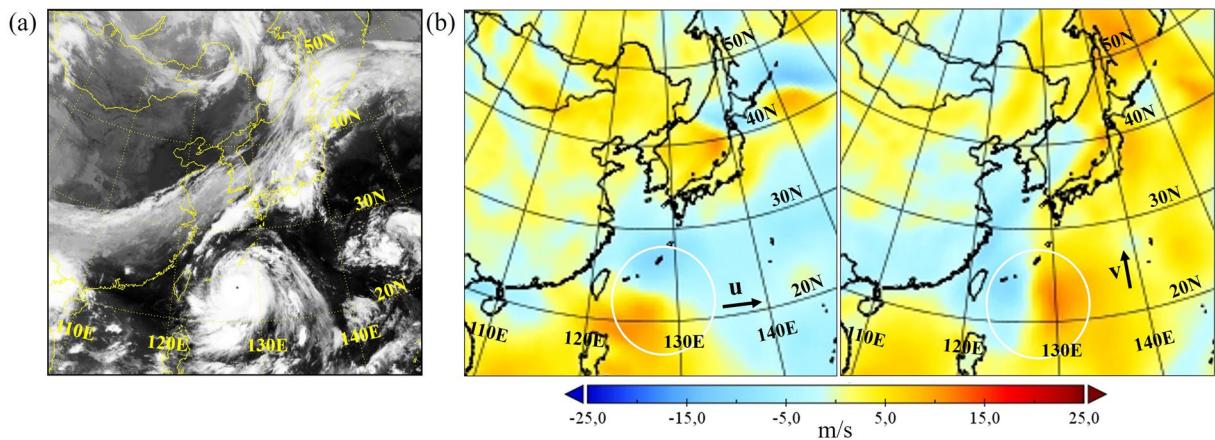
Figure 9. Images of typhoon Rusa in August 2002 as an example for the shots taken from 1993 to 2010 before (**a**) and after (**d**) preprocessing, images of typhoon Neoguri in July 2014 as an example for the shots taken from 2011 to 2014 before (**b**) and after (**e**) pre-processing and images of typhoon Goni in August 2015 as an example for the shots taken from 2015 to 2017 before (**c**) and after (**f**) pre-processing (Source of satellite images: Korean Meteorological Administration (KMA). http://www.weather.go.kr/weather/images).
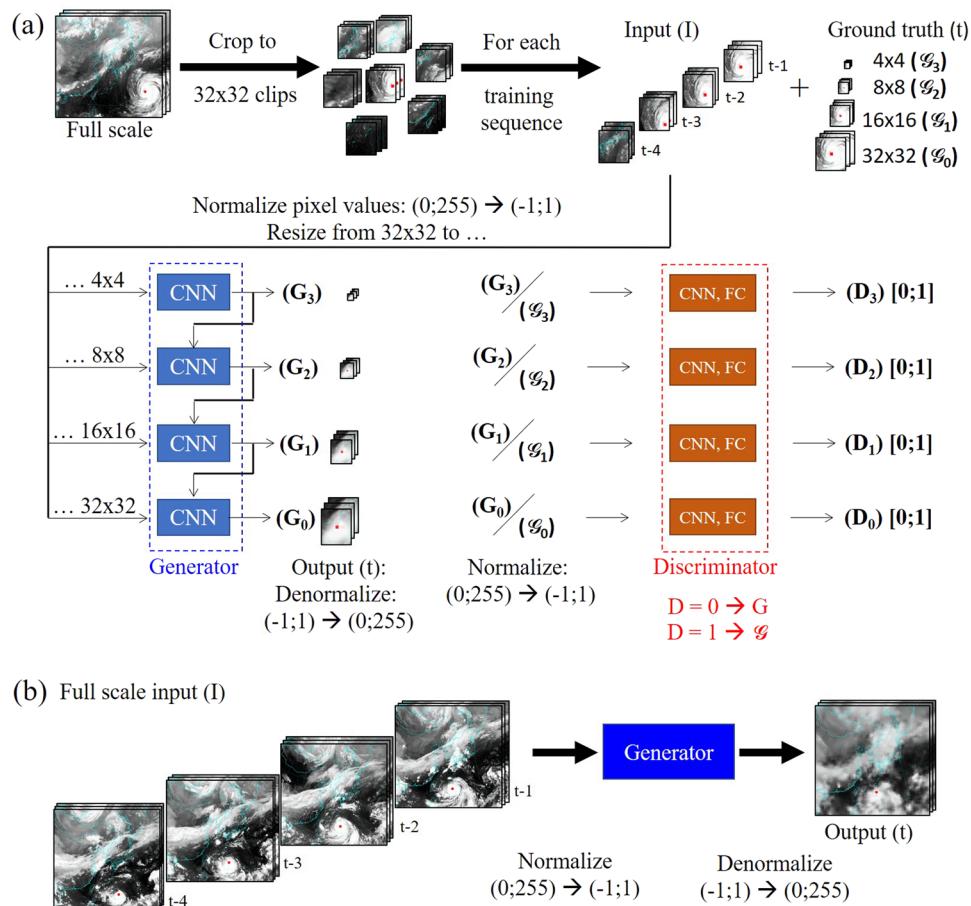
Figure 10. Georeferencing of the typhoon center coordinates (Source of the satellite image: Korean Meteorological Administration (KMA). http://www.weather.go.kr/weather/images).

An important hyperparameter in this study is the number of input images (denoted by m). A high number of m increases memory usage and therefore the learning time. A low number might cause a shortage of input information for the GAN. Three cases with $m = 2$, 4, and 6 are trained and tested. Using currently available satellite images, the computation with $m = 4$ is found to produce the best result. In case of $m = 2$, temporal variation of a typhoon seems not to be sufficiently captured, while in case of $m = 6$, the result is deteriorated due to reduced numbers of training data sets. Considering, for example, typhoon Maemi with 16 sequences, using $m = 4$ allows to start training from sequence 5 and provides a number of 12 training sequences. The choice of $m = 6$, on the other hand, only allows 10 training sequences, starting from sequence 7.

Cropped training clips are trained using variations of deep learning networks in a GAN-based video modeling architecture[16]. This architecture contains a pure convolutional neural network (CNN), so called generator, and a network that combines convolutional layers with fully connected layers (FC), so called discriminator. The generator network takes a clip, normalizes its pixel values from (0; 255) to (−1; 1), predicts a 32×32 part of a satellite image at a future occasion (output, G_0), and denormalizes the predicted pixels from (−1; 1) to (0; 255). GANs are multi-scale networks, therefore the generator network generates images on different scales (G_0, G_1, G_2, G_3). Generated images with pixel sizes of 4×4, 8×8, and 16×16 function as additional inputs for the next bigger scale, as pointed out in Fig. 12(a). This helps to keep long-range dependencies and preserves high resolution information throughout the convolutional learning. The discriminator network, on the other hand, tries to classify between the generated and the ground truth images of each scale. The output layer provides a binary output

Figure 11. (**a**) Satellite image and (**b**) zonal [u] and meridional [v] velocity fields at 10 m height of typhoon Maemi on September 10th in 2003 at 0:00 Coordinated Universal Time (UTC) (Source of the satellite image: Korean Meteorological Administration (KMA). http://www.weather.go.kr/weather/images).

Figure 12. Overall structure of the utilized GAN for (**a**) training and (**b**) testing (Source of satellite images: Korean Meteorological Administration (KMA). http://www.weather.go.kr/weather/images).

(D), where 0 stands for a generated image and 1 means the ground truth image. Testing steps are done with full scale test data, containing images with 250×238 pixels. They are not cropped. As visualized in Fig. 12(b), input images are taken and normalized by the generator network which generates a full scale image at a future occasion and denormalizes it.

The pixel size of 32×32 is selected for efficient usage of GPU memory. Even the network is trained with small size clips (32×32), it can eventually learn the same weights of the network that are learned with larger size clips.

Generator model (fully convolutional architecture)			
Convolution layers (top row: feature map sizes, bottom row: kernel sizes)			
For generating G_0	For generating G_1	For generating G_2	For generating G_3
ch × (m + 1), 128, 256, 512, 256, 128, 3	ch × (m + 1), 128, 256, 512, 256, 128	ch × (m + 1), 128, 256, 128, 3	ch × m, 128, 256, 128, 3
7 × 7, 5 × 5, 5 × 5, 5 × 5, 5 × 5, 7 × 7	5 × 5, 3 × 3, 3 × 3, 3 × 3, 3 × 3, 5 × 5	5 × 5, 3 × 3, 3 × 3, 5 × 5	3 × 3, 3 × 3, 3 × 3, 3 × 3
Discriminator model (convolution layers → max pooling layers → fully connected layers)			
For output D_0	For output D_1	For output D_2	For output D_3
ch, 128, 256, 512, 128	ch, 128, 256, 256	ch, 64, 128, 128	ch, 64
7 × 7, 7 × 7, 5 × 5, 5 × 5	5 × 5, 5 × 5, 5 × 5	3 × 3, 3 × 3, 3 × 3	3 × 3
Max pooling layers (kernel sizes)			
2 × 2	2 × 2	2 × 2	2 × 2
Fully connected layers (neuron numbers)			
1024, 512, 1	1024, 512, 1	1024, 512, 1	512, 256, 1

Table 2. Configuration of the deep learning network. The number of channels (ch) depends on the type of input data. If the GAN takes satellite images only as input data, three channels are needed. If images of zonal and meridional velocity fields are added to the input set, ch increases to 5. The number of input images (m) defines the temporal range of information from which the GAN learns. For the results in Cases A and B, m is set to 4, which means the GAN learns from information of the past 24 hours (4 × 6 hours).

This is because kernel sizes for convolutions are under 7 × 7, so the network learns spatial characteristics under the space size of 7 × 7 pixels. However, during tests, typhoon satellite images on larger size (250 × 238) patches are predicted. This is possible because of the fully convolutional architecture of the generator.

The configuration of the deep learning network is summarized in Table 2 (see[16] for detailed algorithms). An open source code[36] is employed with modifications in input channels so that it is capable of changing the number of prior sequences and of accounting for additional variables such as velocity fields. The number of channels (ch) depends on the type of input data. If the GAN takes satellite images only as input data, three channels are needed. If images of zonal and meridional velocity fields are added to the input set, ch increases to 5. The network is trained from scratch.

The generator model is trained to minimize a combination of loss functions as follows:

$$L_{generator} = 1/4 \sum_{k=0}^{3} \lambda_{l2} L_2^k + \lambda_{gdl} L_{gdl}^k + \lambda_{adv} L_{adv}^{G,k}, \tag{1}$$

where $\lambda_{l2} = 1$, $\lambda_{gdl} = 1$, and $\lambda_{adv} = 0.05$. Let $G_k(I)$ be the predicted image from a convolutional neural network of G_k and $\mathcal{G}_k(I)$ be a $\frac{1}{2^k}$ resized image from the provided ground truth image. The L_2^k loss function evaluates the explicit difference between the predicted and provided images as

$$L_2^k = ||G_k(I) - \mathcal{G}_k(I)||_2^2. \tag{2}$$

Generative techniques for video modeling are known to suffer from the blurriness[16,37,38]. To reduce the blurriness, Mathieu et al.[16] proposed a gradient difference loss function (GDL) and reported that a GAN with the GDL improves the blurriness compared to other methods including a recurrent neural network based method proposed by Ranzato et al.[38]. The GDL, $\mathcal{L}_{gdl}^k$, compares the difference between gradients of the predicted and provided images as

$$L_{gdl}^k = \sum_{i=0}^{n_x-2} \sum_{j=0}^{n_y-2} \left\{ \left| \left| \mathcal{G}_k(I)_{(i+1,j+1)} - \mathcal{G}_k(I)_{(i,j+1)} \right| - \left| G_k(I)_{(i+1,j+1)} - G_k(I)_{(i,j+1)} \right| \right| \right.$$
$$\left. + \left| \left| \mathcal{G}_k(I)_{(i+1,j+1)} - \mathcal{G}_k(I)_{(i+1,j)} \right| - \left| G_k(I)_{(i+1,j+1)} - G_k(I)_{(i+1,j)} \right| \right| \right\}, \tag{3}$$

where n_x and n_y are numbers of pixel in the width and the height of an image, and (i, j) is a pixel coordinate. $\mathcal{L}_{adv}^{G,k}$ loss function is applied to support the generator model to delude the discriminator network by generating images which are indistinguishable from the provided ground truth images as

$$L_{adv}^{G,k} = L_{bce}(D_k(G_k(I)), 1), \tag{4}$$

where L_{bce} is the binary cross entropy loss function defined as

$$L_{bce}(a, b) = -b\log(a) - (1 - b)\log(1 - a), \tag{5}$$

for scalars a and b between 0 and 1.

The discriminator model is trained to minimize a loss function as follows:

$$L_{discriminator} = \frac{1}{4} \sum_{k=0}^{3} [L_{bce}(D_k(\mathcal{G}_k(I)),1) + L_{bce}(D_k(G_k(I)),0)]. \tag{6}$$

This supports the network not to be deluded by the generator model by extracting important features of typhoons in an unsupervised manner.

For each testing, the generator model is fed with m consecutive full scale satellite images from the past and generates a full scale typhoon satellite image at future occasion (see Fig. 12(b)).

Errors in the prediction of typhoon centers.   In the current study two different errors are investigated. The first one describes the distance between the predicted coordinate (ϕ_{pred}, λ_{pred}) and the real coordinate (ϕ_{real}, λ_{real}), named as an absolute error (E). E is calculated in kilometers (km) by applying the haversine formula[39], with the earth radius R taken at the location of the real coordinate:

$$E = 2R\arcsin\sqrt{\sin^2\left(\frac{\phi_{pred} - \phi_{real}}{2}\right) + \cos\phi_{real}\cos\phi_{pred}\,\sin^2\left(\frac{\lambda_{pred} - \lambda_{real}}{2}\right)}. \tag{7}$$

However, the absolute error does not give sufficient information about the prediction quality. Tracks of slowly moving typhoons, for example, are more difficult to predict than tracks of fast ones. It is therefore necessary to introduce a relative error (E_{rel}), that is calculated as follows:

$$E_{rel}(t) = \frac{E}{2R\arcsin\sqrt{\sin^2\left(\frac{\phi_{real}(t) - \phi_{real}(t-6h)}{2}\right) + \cos\phi_{real}(t)\cos\phi_{real}(t-6h)\sin^2\left(\frac{\lambda_{real}(t) - \lambda_{real}(t-6h)}{2}\right)}}. \tag{8}$$

This type of error considers the ratio between E and the distance that a typhoon has traveled over the last 6 hours, where t stands for the time of a certain sequence.

Data Availability
All the data used in the present work are available upon request.

References
1. Kepert, J. Tropical cyclone structure and dynamics. In *Global Perspectives on Tropical Cyclones: from science to mitigation*, 3–53 (World Scientific, 2010).
2. Schrader, C. Hurrikane und taifune: wirbelstürme bewegen sich immer langsamer, http://www.spiegel.de/wissenschaft/natur/hurrikane-und-taifune-wirbelstuerme-bewegen-sich-immer-langsamer-a-1211499.html (2018).
3. Sun, Y. *et al.* Impact of ocean warming on tropical cyclone size and its destructiveness. *Sci. Reports* **7**, 1–10 (2017).
4. Kossin, J. A global slowdown of tropical-cyclone translation speed. *Nat.* **558**, 104–107 (2018).
5. Tilden, C. Annual typhoon report. Tech. Rep., U.S. fleet weather central/Joint typhoon warning center (1959).
6. Kim, A., Tachikawa, Y. & Takara, K. Recent flood disasters and progress of disaster management system in Korea. *Annu. Disas. Prev. Res. Inst., Kyoto Univ.* **50 B** (2007).
7. National emergency management. Maemi - disaster reports, https://web.archive.org/web/20131029185710/ http://eng.nema.go.kr/sub/cms3/3_2.asp (2018).
8. Xu, X. *et al.* Does warmer china land attract more super typhoons? Sci. *Reports* **3**, 1–8 (2013).
9. Kovordanyi, R. & Roy, C. Cyclone track forecasting based on satellite images using artificial neural networks. *ISPRS J. Photogramm. Remote. Sens.* **64**, 513–521 (2009).
10. Lee, R. S. T. & Liu, J. N. K. Tropical cyclone identification and tracking system using integrated neural oscillatory elastic graph matching and hybrid RBF network track mining techniques. *IEEE Transactions on Neural Networks* **11**, 680–689 (2000).
11. Hong, S., Kim, S., Joh, M. & Song, A. Globenet: convolutional neural networks for typhoon eye tracking from remote sensing imagery. *arXiv:1708.03417v1* (2017).
12. Kordmahalleh, M. M., Sefidmazgi, M. G. & Homaifar, A. A sparse recurrent neural network for trajectory prediction of atlantic hurricanes. In *Proc. Genet. Evol. Comput. Conf.* 957–964 (2016).
13. Alemany, S., Beltran, J., Perez, A. & Ganzfried, S. Predicting hurricane trajectories using a recurrent neural network. *arXiv:1802.02548v2* (2018).
14. Zhang, Y., Chandra, R. & Gao, J. Cyclone track prediction with matrix neural networks. In *International Joint Conference on Neural Networks* (2018).
15. Goodfellow, I. *et al.* Generative adversarial networks. *arXiv:1406.2661v1* (2014).
16. Mathieu, M., Couprie, C. & LeCun, Y. Deep multi-scale video prediction beyond mean square error. *arXiv: 1511.05440* (2015).
17. Lee, S. & You, D. Data-driven prediction of unsteady flow fields over a circular cylinder using deep learning. *arXiv:1804.06076v1* (2018).
18. Liu, R. *et al.* An intriguing failing of convolutional neural networks and the coordconv solution. In *Advances in Neural Information Processing Systems*, 9627–9638 (2018).
19. Etro, J. & Bassi, J. Annual tropical cyclone report. Tech. Rep., Joint Typhoon Warning Center (1995).
20. Dillon, C. & Andrews, M. J. Annual tropical cyclone report. Tech. Rep., Joint Typhoon Warning Center (1996).
21. Dillon, C. & Andrews, M. J. Annual tropical cyclone report. Tech. Rep., Joint Typhoon Warning Center (1997).
22. Furze, P. & Engel, G. Annual tropical cyclone report. Tech. Rep., Joint Typhoon Warning Center (2000).
23. Furze, P. & Engel, G. Annual tropical cyclone report. Tech. Rep., Joint Typhoon Warning Center (2002).
24. Furze, P. & Preble, A. Annual tropical cyclone report. Tech. Rep., Joint Typhoon Warning Center (2003).
25. RSMC Tokyo - Typhoon Center. Annual report on the activities of the rsmc tokyo - typhoon center 2007. Tech. Rep., Japan Meteorological Agency (2007).
26. RSMC Tokyo - Typhoon Center. Annual report on the activities of the rsmc tokyo - typhoon center 2011. Tech. Rep., Japan Meteorological Agency (2011).
27. RSMC Tokyo - Typhoon Center. Annual report on the activities of the rsmc tokyo - typhoon center 2014. Tech. Rep., Japan Meteorological Agency (2014).
28. RSMC Tokyo - Typhoon Center. Annual report on the activities of the rsmc tokyo - typhoon center 2016. Tech. Rep., Japan Meteorological Agency (2016).
29. Wu, L., Li, Z., Duan, J. & Zong, H. Sudden tropical cyclone track changes over the western north pacific: a composite study. *Mon. Weather. Rev.* **141**, 2597–2610 (2013).
30. Korean Meteorological Administration (KMA), https://web.kma.go.kr/eng/ (2018).

31. Sharing Earth Observation Resources(eo). Gms (geostationary meteorological satellite) series of japan, https://directory.eoportal.org/web/eoportal/satellite-missions/g/gms (2018).
32. Japan Meteorological Agency (JMA), https://www.jma.go.jp/jma/indexe.html (2018).
33. ERA-interim Public dataset of the ECMWF, http://apps.ecmwf.int/datasets/data/interim-full-daily/levtype=sfc/ (2018).
34. Dee, D. *et al.* The era-interim reanalysis: configuration and performance of the data assimilation system. *Q. J. Royal Meteorol. Soc.* **137**, 553–597 (2011).
35. Schmunk, R. Panoply netcdf, hdf and grib data viewer, https://www.giss.nasa.gov/tools/panoply/ (2018).
36. Github. Adversarial video generation, https://github.com/dyelax/Adversarial_Video_Generation (2019).
37. Srivastava, N., Mansimov, E. & Salakhudinov, R. Unsupervised learning of video representations using LSTMs. In *International Conference on Machine Learning*, 843–852 (2015).
38. Ranzato, M. *et al.* Video (language) modeling: a baseline for generative models of natural videos. *arXiv preprint arXiv:1412.6604* (2014).
39. Mahmoud, H. & Akkari, N. Shortest path calculation: a comparative study for location-based recommender system. *2016 World Symp. on Comput. Appl. Res. (WSCAR)* (2016).

Acknowledgements

Author Contributions

All authors contributed to the study. M.R. processed the images and analyzed the data. S.L. developed the GAN source code for typhoon prediction. D.Y., M.R., S.L. and J.S. made analyses and discussion of results and all participated in completing the manuscript.

Additional Information

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.