Attention-Hurricane: Pixel-wise Spatio-temporal Hurricane Path Prediction

Ashkan Farhangi
Department of Electrical and Computer Engineering
University of Central Florida
ash@knights.ucf.edu

***Abstract*—Entering the Big Data era, there is a rise in computer vision algorithms for modern video datasets to tackle challenging problems such as object and action detection. For many, the use of video datasets in their experiments will produce difficulties in the overall computation and preprocessing. Therefore, there is a crucial need to improve the performance of action detection algorithms. Typically, action detection algorithms tend to be complex and use a great number of parameters to ensure optimal accuracy. However, for efficient computation in cloud computing, there is a need for less complex and computation for action detection datasets. Therefore, we propose a deep recurrent neural network Attention-Hurricane that is designed to be efficient and deliver adequate accuracy in hurricane trajectory prediction. Evaluate the prediction accuracy we have showcased the network accuracy compare to previous state-of-art methods. The prelimnary results are promising as the networks ability to use attention mechanism gives hope for better trajectory prediciton.**

I. Introduction

Every year hurricanes produce loss and destruction in many places around the world. As the disruptive force of a hurricane comes from the wind speed of the rotating air and the rainfall that can cause disastrous flooding, the highest wind speeds are usually found near the hurricane center [**?**]. Furthermore, it found that hurricanes are becoming stronger, larger, and more destructive in the context of global warming [**?**]. The present work concentrates on tracks of hurricanes. As accuracy is only one measure about the prediction of natural disasters, speed and versatility play an important role as well. Forecasts should be done quickly and forecast tools should be able to react immediately on sudden changes. It's also reported that the current hurricane predictions of South Korea are done by conducting numerical simulations on a Cray XC40 supercomputer with 139,329 CPUs which holds aa huge amount of computational time. Not to mention that the maintenance costs for such expensive hardware are also very high. Therefore, an efficient forecast method that considers all these criteria is necessary [**?**]. Recently, the computer vision community has made important progress by using various pattern recognition techniques in visual object tracking, a task to locate a target object in a video, maintaining its identity and yielding its trajectory, given its initial position in the first frame [**?**]. Extreme hurricane event tracking is similar to visual object tracking, but it has unique, challenging aspects: hurricane events may be dependent on longer-term and wider-range Spatio-temporal dynamics among multiple experimental variables than the targets in visual object tracking do on RGB pixels. The target events are not often defined as rigid bodies, flexibly changing their shape with no clear boundary, and are difficult to visually separate them from each other. Thus, it is usually difficult to associate an object of interest with the correct one in consecutive frames. hurricane event data are normally sparsely collected (both temporally and spatially), as it occurs rarely but requires special devices installed in the wild. For example, the hurricane data we use in this paper are collected every three hours, sufficient time for a hurricane to move 350 km. That makes it difficult for us to assume the object would appear nearby in consecutive frames. Because of these challenges, a conventional tracking-by-detection method, which relies heavily on the amount of training data and detects the object mainly by its appearance but relatively neglects Spatio-temporal dynamics, is less suitable for this problem [**?**]. In this work, we introduce a simple but sound end-to-end model, suitable for the hurricane tracking problem. Particularly, the proposed model consists of two sub-modules, the (1) focus learning module to learn where and what to converge, and the (2) tracking module to track what we converged on. The focus learning module is designed to extract the latent feature of the target event from the first frame, given the initial image and the location of the target [**?**]. Given the representation of the target event, the tracking module predicts its location by localizing the learned feature of a target object in the subsequent frames. In both modules, we adopt ConvLSTM-based auto-encoder structure for the two subsequent reasons. First, it learns a mapping from time-series hurricane variables to a time-series density map with event possibilities, suffering less from a blurry boundary of hurricane events. It compactly embeds the target's historical appearance change and movement in hidden states, capturing essential Spatio-temporal dynamics of the target.

Contribution

. We implemanted two Attention mechanism to hurricane trajectory predicition.

. Pixel-wise spatio Attention is implemanted in 2D space

to increase the trajectory prediciton accuary.

. Spatio-temporal Attention mechanism utlizes the hurricanes timely behaviour to increase the models prediction accuracy.

. Our Attention-Hurricane network uses data augmentation that increases the overall efficiency.

II. Background and Related Work

[?] Tracking-by-detection approaches first detect candidates of the target mainly by using convolutional neural networks (CNNs). Afterward, the most likely candidate is chosen. When multiple target objects exist, each target needs to be associated with a distinct candidate. Among various CNN-based trackers, multi-domain CNN (MDNet) [?], [?]achieves the state-of-the-art performance on multiple datasets by learning discriminative features for instance classification To track a target without a rigid boundary or motion against the background, pixel-wise object tracking methods have advantages over tracking-by-detection approaches. Furthermore, [] adopted gradient-boosted decision trees to estimate pixel-level annotation of a segmentation mask.

In the works done by [?], 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 hurricane favorably identifies the future position of the hurricane center as well as the deformed cloud formations. In the work of [] and [] various existing forecast techniques are listed. Including analog forecasts, features of the hurricanes are compared to all previous storms in the same region and the movement is derived from past events 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. 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. To best of our knowledge, [] first used satellite pictures of tropical cyclones for track forecast with the help of neural networks. But images did not work 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 derived from images and fed to the network for a time-series prediction. 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. Furthermore, [] was the first who study which used satellite images as input data for a neural network. The network favorably detected the shape of a hurricane and predicted the future movement direction. Also, [] have used a sparse recurrent neural network (RNN) to predict trajectories of cyclones coming from the 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 forecast process, only hurricanes that have connections in trajectories to that of the target hurricane were used. One disadvantage is that they assumed monotonic behavior, which means that a hurricane never comes back to a position where it was before. In reality, this is not always the case

Recurrent Neural Networks (RNN) have gain momentum as their architecture use more parameters and deep algorithms have more tools to predict in various fields such as financial, medical and trajectories such as hurricanes. To best of our knowledge [] used an RNN to predict hurricane trajectories, but instead of assuming monotonic behavior they examined 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 estimate from past movements of the hurricane as inputs. With the grid identification number, spatial relations on a map were learned. Their forecasts could achieve better accuracy than the results of [?], [?] utilized matrix neural networks (MNNs) for predicting trajectories of cyclones that occurred in the South Indian ocean. It also stated that MNNs suit is better for the task of hurricane 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 hurricane 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 compositions around the typhoon [?] 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 hurricane, from which, afterward, the future position of the typhoon the center is determined. 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 hurricane center.

Recently, convolutional LSTM (ConvLSTM) [30] has been widely applied to handle spatio-temporal dynamics in analyzing video data, e.g., precipitation forecasting [31], motion prediction [8], pixel-wise video prediction [38], crowd counting [41], and segmentation-based tracking [35]. Romera et al. [28] used ConvLSTMs for instance segmentation on a single image.

Conventional extreme climate event detection and tracking methods rely on a number of mathematical simulation-based methods, including an ensemble of multiple prediction models or multi-scale prediction systems Also, ConvLSTM [] and incremental neural networks were proposed to predict the future trajectory of hurricanes and cyclones. Region CNNs

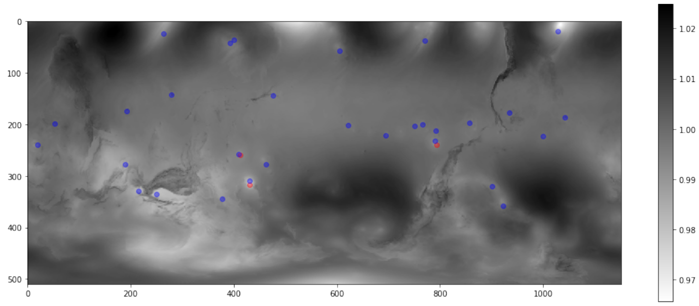


Fig. 1: feature.

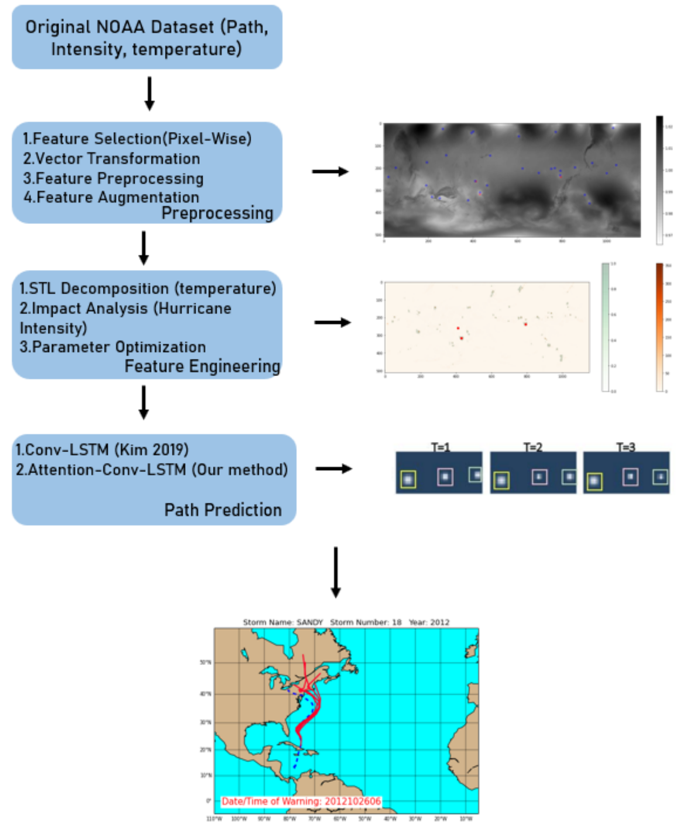


Fig. 2: The networks architecture follows this procedure. In our opinion the feature extraction is imporatant for pixel-wise attention mechansim. The networks input attention mechanism decides the weights of each input vector and our model will use both attention mechanism in the space-time vectors.

were applied to classify different types of extreme hurricane events. Most existing approaches, however, have not addressed unique challenges to manage sparse climate data covering a wide geographic range for an extended period. In this work, we tackle the unique challenges of the climate event tracking problem introduced in Section 1 with the ConvLSTMs-variant model, which is specially designed to seize a wide range of Spatio-temporal dynamics.

In previous studies, various machine learning models have been used to predict hurricane paths

In [2], the authors treated the hurricane path as a time-series problem that can be solved with 5 a simple RNN model. One limitation of this model is not using the hurricane path data with its 6 intensity accordingly. The model simply does not know what is the hurricane severity at each time 7 step. However, the intensity of the hurricane holds great information about the speed, path and other 8 factors that are important in accurate path prediction [1].

Later on, in a collaborative study done by researchers at Google, Microsoft and Lawrence Livermore lab, [3], they further improved the accuracy of their previous models by using a simple Conv-LSTM architecture 1. Although the research in this field is promising, compared to reliable statistical models, these models are not capable of accurately predict a never seen before hurricane. This is due to the fact that most of these models are using the first 80 In the next section, we introduce our solution by tackling the forementioned limitations.

III. Methodology

As mentioned in the limitation section, the simple state-of-art Conv-LSTM architecture does not 18 have enough features to deliver a reliable prediction. Using only the observed atmosphere images 1 19 will not result in a comprehensive model that can be used. 20 In our study, we intend to tackle these limitations by using a more comprehensive dataset from 21 the National Oceanic and Atmospheric Administration [2]. In combination with a better feature vector, 22 we intend to improve the previous state-of-art architecture (Conv-LSTM) by using an input attention 23 mechanism 2 .

$128 \times 256 \times 1$
$128 \times 256 \times 48$
$128 \times 256 \times 24$
$128 \times 256 \times 12$
$128 \times 256 \times 6$

A. *Attention Mechanism*

IV. Experiments and Discussion

A. *Training*

Parameter	Value
input_dim	1000
hidden_layers	101
rnn_layers	2
Res-Net 18	True
LSTM	Dropout = 0.1
GRU	Dropout = 0.2

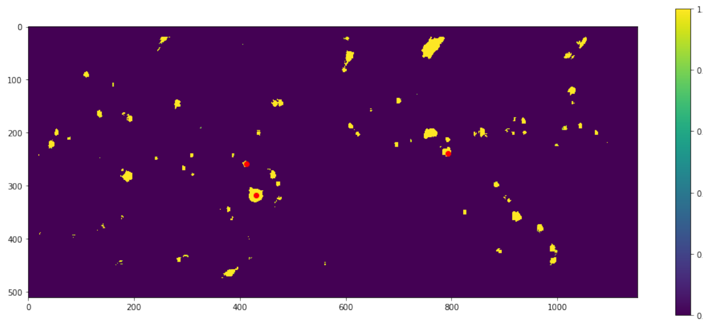


Fig. 3: Pixel-wise Attention mechanism: The use of attention mechanism in input (2D space) will add a new dimension of weights in 2D dimension. The networks training phase will optimize the weights to for adding space attention on parts of images that are will improve the accuracy.

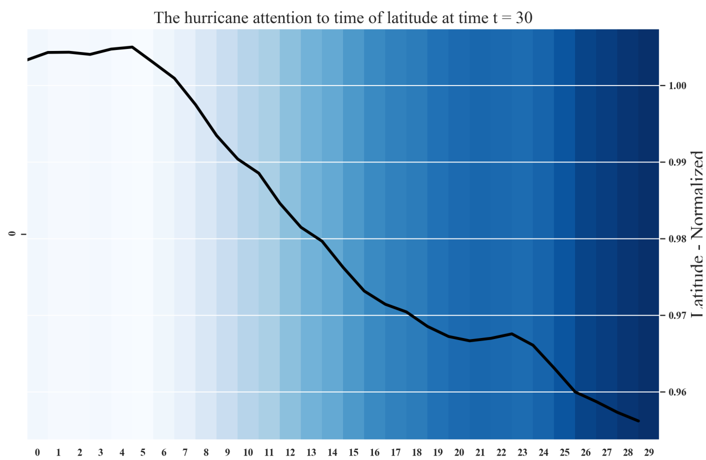


Fig. 4: Spatio-temporal Attention mechanism: In this case, the attention weights have included the temporal dimension. During the training, the networks attention will be set on latter part of the input (when hurricane is latest reported) to predict the next location. Due to visualization limiations, the y-axis will hold only the lattitude value.

Model	IOU: 0.1	IOU: 0.5
Detection-based CNN [?]	14.78%	12.26%
3D Conv Encoder-decoder [?]	22.65%	15.53%
2 layered ConvLSTM	88.96%	87.32%
3 layered ConvLSTM	93.77%	92.18%
4 layered ConvLSTM	92.58%	90.87%
5 layered ConvLSTM	93.49%	91.93%
4 layered Attention-Hurricane (Ours)	91.63%	89.63%
5 layered Attention-Hurricane (Ours)	92.36%	90.45%

Table 1. Comparison of AvgP for different tracking models.

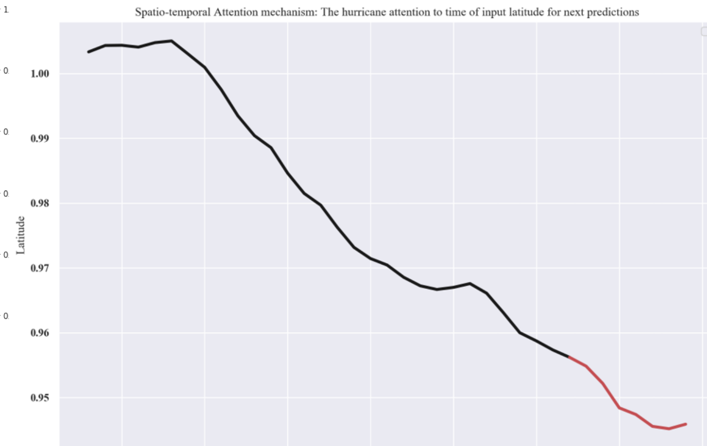


Fig. 5: Lattidude prediction based on previous attention: The models prediction is simlar to latest part of attention where the attention weights were made in lattidude decrease.

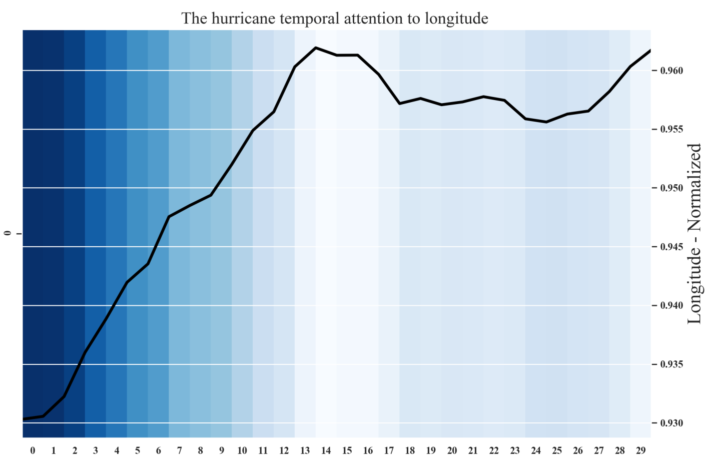


Fig. 6: Lattidude prediction based on previous attention: The models prediction is simlar to latest part of attention where the attention weights were made in lattidude decrease.

V. Conclusion & Future Work

In this work, we showcased the use of GRU in action detection networks. Since our network is faced by a set of constraints we were not able to showcase the accuracy of our performance. However, for future work, we are looking to include the preliminary semantic segmentation in UCF-101 dataset and provide a better accuracy measures to satisfy our readers. As mentioned before, it would be interesting to evaluate CSD [?] in a noisy environment. Specifically, when a part of the object in each data in missing.

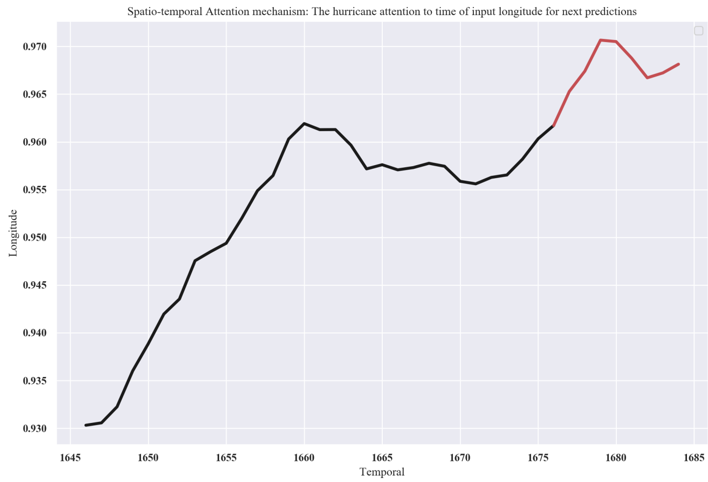
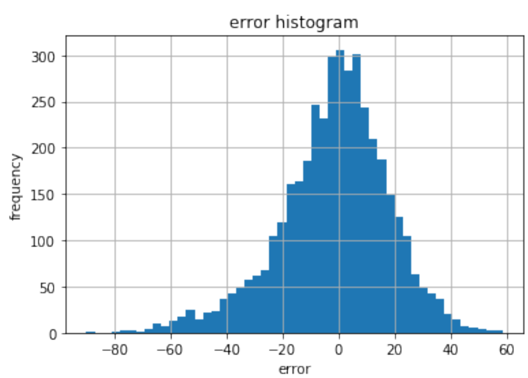
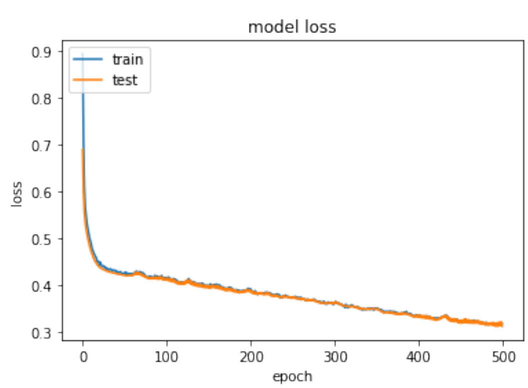


Fig. 7: Lattidude prediction based on previous attention: The models prediction is simlar to latest part of attention where the attention weights were made in lattidude decrease.

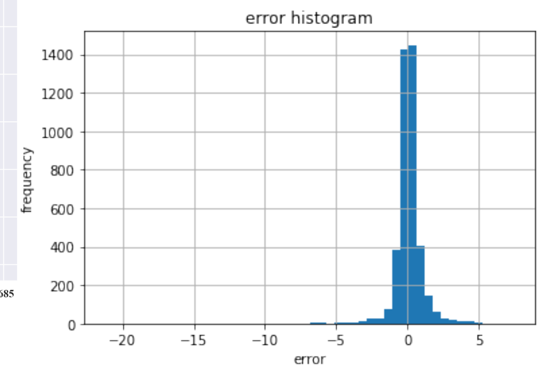


(a) Original Figure

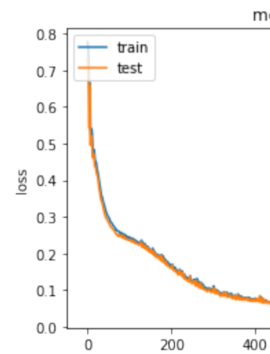


(b) Optical Flow

Fig. 8: Training .

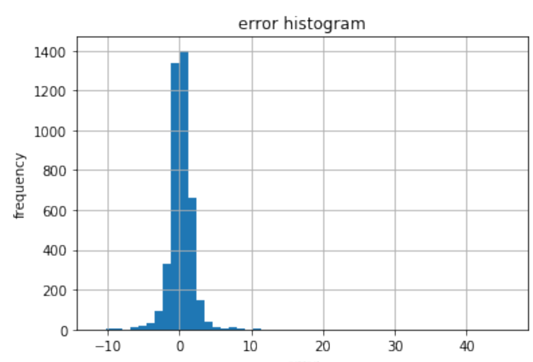


(a) Original Figure.

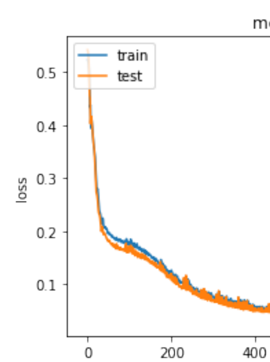


(b) Opti

Fig. 9: Training .



(a) Optical Flowdetection.



(b) Origin

Fig. 10: Training .