Severe Thunderstorm Detection by Visual Learning Using Satellite Images

Yu Zhang, Stephen Wistar, Jia Li, Michael Steinberg and James Z. Wang

***Abstract*—Computers are widely utilized in today's weather forecasting as a powerful tool to leverage an enormous amount of data. Yet, despite the availability of such data, current techniques often fall short of producing reliable detailed storm forecasts. Each year severe thunderstorms cause significant damage and loss of life, some of which could be avoided if better forecasts were available. We propose a computer algorithm that analyzes satellite images from historical archives to locate visual signatures of severe thunderstorms for short-term predictions. While computers are involved in weather forecasts to solve numerical models based on sensory data, they are less competent in forecasting based on visual patterns from both current and past satellite images. In our system, we extract and summarize important visual storm evidence from satellite image sequences in the way that meteorologists interpret the images. In particular, the algorithm extracts and fits local cloud motion from image sequences to model the storm-related cloud patches. Image data from the year 2008 have been adopted to train the model, and historical severe thunderstorm reports in continental US from 2000 through 2013 have been used as the ground-truth and priors in the modeling process. Experiments demonstrate the usefulness and potential of the algorithm for producing more accurate severe thunderstorm forecasts.**

I. INTRODUCTION

A. Background

The numerical weather prediction (NWP) approach has been applied to weather forecasting since the 1950s [1], when the first computer ENIAC was invented [2]. John von Neumann, a pioneering computer scientist and meteorologist, first applied the primitive equation models [3] on ENIAC. Since then meteorologists and modelers have produced increasingly reliable weather forecasts, and they have continued to refine the numerical models, thereby improving the NWP in recent decades. Today, the model of the European Centre for Medium-range Weather Forecasts and the United States' Global Forecast System are two of the most accurate global NWP models. At the 500 hPa geopotential height level of the atmosphere, both models are reported to have annual mean correlation coefficients of nearly 0.9 for 5-day forecasts [4]. Other models, such as the North American Mesoscale model and the High-Resolution Rapid Refresh model use a smaller grid to provide more detailed local forecasts with varying degrees of success.

Y. Zhang and J. Z. Wang are with the College of Information Sciences and Technology, The Pennsylvania State University, University Park, PA 16802, USA. Emails: ryanflower@gmail.com, jwang@ist.psu.edu.

S. Wistar and M. Steinberg are with AccuWeather Inc., State College, PA 16803, USA.

J. Li is with the Department of Statistics, Eberly College of Science, The Pennsylvania State University, University Park, PA 16802, USA.

Manuscript received –; revised –.

Even though NWP can effectively produce accurate weather forecasts of the general weather pattern, it is not always reliable in the prediction of extreme weather events, such as severe thunderstorms, hail storms, hurricanes, and tornadoes, which cause a significant amount of damage and loss of life every year worldwide. Unreliable predictions, either missed or false alarm ones, could cause massive amounts of loss to society. One of the most well-known failed predictions is the Great Storm of 1987 in UK [5]. The Met Office was criticized for not predicting the storm in a timely manner and being partially responsible for the 22 deaths and 2 billion pounds loss. More recently, there were major differences in various global NWP models regarding the landfall of Hurricane Sandy in October 2012 in the northeastern US, and the track of Hurricane Joaquin off the East Coast of the US in late September and early October 2015. On the other hand, the impact of the Blizzard of 2015 was overestimated by the US National Weather Service [6]. Several cities in the US East Coast suffered from unnecessary travel bans and public transportation closures. Much of the loss of life and property due to improper precautions to severe weather events could be avoided with more accurate (for both severity and location) weather forecasts.

The numerical methods used in NWP can efficiently process large amounts of meteorological measurements. However, a major drawback of such methods is that they are sensitive to noise and initial inputs due to the complexity of the models [7], [8]. As a result, although nowadays we have powerful computers to run complex numerical models, it is difficult to get accurate predictions computationally, especially in the forecasts of severe weather. To some extent, they do not interpret the data from a global point of view at a high cognitive level. For instance, meteorologists can make good judgments of the future weather conditions by looking at the general cloud patterns and developing trends from a sequence of satellite images by using geographical knowledge and their experience of past weather events [9], [10]; numerical methods do not capture such high-level clues. Additionally, historical records of severe weather provide valuable references for making weather forecasts, but traditional NWP [1] models do not make the best use of them [11].

To address this weakness of numerical models, we develop a computational weather forecast method that takes advantage of both the high-level visual clues of satellite image sequences and the historical records. In particular, we try to find synoptic scale features of mid-latitude storm systems that could produce severe weather (to be specified later in Section I-C) by modern computer vision and machine learning approaches. It is worth mentioning that the purpose of this work is not to replace numerical models in forecasting. Instead, by tackling the problem using different data sources (*i.e.*, satellite images) and modeling methodologies, we aspire to develop a method that can potentially be used side-by-side with and complement numerical models.

We analyze the satellite image sequence because it provides important clues as meteorologists view evolving global weather systems. Unlike conventional meteorological measures such as temperature, humidity, air pressure, and wind speed, which directly reflect the physical conditions at the sensors' locations, the visual information in satellite images is captured remotely from the orbit, which means a larger geographic coverage with good resolution. The brightness in an infrared satellite image indicates the temperature of the cloud tops [12] and therefore reveals information about the top-level structure of the cloud systems. Visual information in continuous image sequences further indicates the development of the cloud systems, which helps us trace the evidence that corresponds to certain weather patterns, particularly those related to severe thunderstorms. Moreover, certain cloud patterns can be evident in the early stage of severe thunderstorm development. As a result, the satellite imagery is very helpful to meteorologists in their work. It is sensible to incorporate image analysis in automated weather forecasting.

Human eyes can effectively capture the visual patterns related to different types of weather events. The interpretation requires a high level understanding of meteorological knowledge, which has been difficult for computers. However, because the broad spatial and temporal coverage of the satellite images produces too much information to be fully processed by human beings, there is an emerging need for a computerized way to analyze the satellite images automatically. Therefore researchers are seeking computer vision algorithms to process and interpret the visual features automatically from the satellite images, aiming to help meteorologists better analyze and predict weather events. Traditional satellite image analysis techniques include cloud detection with complex backgrounds [13], cloud type classification [14]–[16], detection of cloud overshooting top [17], [18], and tracking of cyclone movement [19]. These approaches, however, only capture local visual features without utilizing many high-level *global visual patterns*, and overlook their *development over time*, which provides more clues for weather forecasts. With the development of advanced computer vision algorithms, some recent techniques put more focus on the analysis of cloud motion and deformation [20]–[23]. In these approaches, cloud motions are estimated to match and compare cloud patches between adjacent images in a sequence. In this paper, the proposed algorithm also uses cloud motion estimation in image sequences. It is different from existing work in that it extracts and models certain patterns of cloud motion, in addition to capturing the cloud displacement.

Historical satellite data archives as well as meteorological observations are readily available for recent years. By analyzing large amounts of past satellite images and inspecting the historical record of severe weather, our system takes advantage of big data to produce severe weather alerts given a certain query image sequence. The result provides an alternative prediction that can validate and correct the conventional forecasts. It can be embedded into an integrative or information fusion system, as suggested in [24], to work with data and decisions from other sources and produce improved storm forecasts. We expect the demonstrated paradigm system be incorporated as an additional tool into an operational forecasting environment in the future with continued exploration and refinement.

B. The Dataset

For our research, we acquired all of the GOES-M [25] satellite imagery for the year of 2008[1], which was an active year containing abundant severe thunderstorm cases for us to analyze. The GOES-M satellite moves on a geostationary orbit and was continually facing the North America area during that year. In addition to the imagery data, each record contains navigation information, which helps us map each image pixel to its real geo-location. The data is multi-spectral and contains five channels covering different ranges of wavelengths, among which we adopted channel 3 (6.5-7.0μm) and channel 4 (10.2-11.2μm) in our analysis because these two infrared channels are available both day and night. Both channels are of 4 km resolution and the observation frequency is four records per hour, *i.e.*, about 15 minutes between the adjacent images. To make sure two adjacent images in a sequence have a noticeable difference, we sampled the original sequence and used a subsequence with a 2-frames-per-hour frame rate. We select the 30-minute sampling interval because the algorithm is based on the optical flow analysis between to adjacent image frames. We need both evident enough visual difference and relatively good temporal resolution. The 30-minute interval is an empirical selection that considers both factors. As the spatial resolution is expected to much improve in future satellites, it is expected that the temporal resolution for the algorithm can be further reduced with the help of more visual details. In the rest of the paper, every two adjacent images in a sequence we refer to are 30 minutes apart without specifying.

Using the imagery and navigation data blocks in each raw data entry, we reconstruct satellite images by an equirectangular projection [26]. The pixel mapping from the raw imagery data to the map coordinate is computed by the navigation data with transformation defined by [25]. The map resolution is set to be 4 km (per pixel), which best utilizes the information of the raw data. We reconstruct the aligned images within the range of 60°W to 124°W in longitude and 20°N to 52°N in latitude, which covers the Continental US (CONUS) area. Hereafter all the image related computations in our approach are performed on the reconstructed images, on which the geo-location of each pixel is directly accessible.

We studied the relationship between satellite images and severe thunderstorms, which may include hailstorms and tornadoes[2]. To relate the satellite data with severe thunderstorms, we retrieved the storm report data from the NOAA National

[1] The satellite imagery is publicly viewable and the raw data archive can be requested on the website of US National Oceanic and Atmospheric Administration (NOAA).

[2] Severe thunderstorms mostly contain hailstorms and tornados.

Weather Service[3], where all the time and locations of storms inside the United States since the year of 2000 are archived. The records act as ground-truth data[4] in the training and provide geographic and temporal priors for the system to make decisions.

C. Synoptic-scale Visual Storm Signatures

Being evidently visible on the satellite imagery, synoptic-scale (spanning several hundred kilometers) weather systems are studied in this paper. Particularly, we are interested in mid-latitude (30°-60°N) synoptic-scale features associated with thunderstorms due to their impact on the CONUS and other densely populated areas around the world.

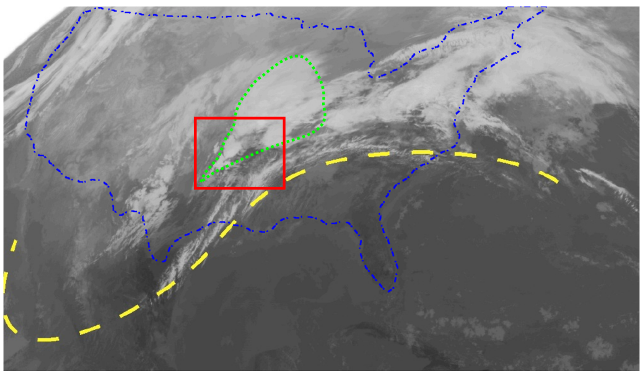


Fig. 1. A sample GOES-M satellite channel 4 image taken at 16:02 (GMT) on February 5, 2008. The boxed area underwent an outbreak of severe thunderstorms later on that day. The jet stream (marked in dashed curve) can be implied from the distribution of clouds. The synoptic-scale system that triggered the severe thunderstorms is demarcated by a comma-shaped cloud patch (surrounded by dashed boundary) in the trough region.

In the northern hemisphere, elongated cold air masses with lower pressures in the higher levels of the atmosphere well above the ground, namely "troughs", will have a dip in the jet stream extending southward at various longitudes around the earth while in other locations a ridge of higher pressure aloft with warm air will send the jet stream bulging northward. These planetary-scale north-to-south oscillations in the jet stream, known as Rossby waves, are major factors in the weather changes that go on in the mid-latitudes. As the ridges and troughs travel around the planet, bringing alternating warm (ridges) and cold (troughs) air masses, the surface and near-surface boundary zones between these differing air masses are prime locations for severe weather given sufficient contributing factors. The east side of a trough, which has a southwest to northeast oriented jet stream, is typically next to an area of ascending warm air, where the development of convective storms is encouraged, given sufficient moisture in the atmosphere [27]. Though jet streams themselves are invisible, they can be revealed by thick cumulus clouds gathering in the unsettled trough regions and areas of clouds spreading northeastward along the east side of a trough. As a result, an elongated cloud-covered area along a southwest to northeast direction is a useful large-scale clue for us to locate severe thunderstorms. Fig. 1 shows a GOES-M channel-4 image taken on February 5, 2008. Clearly in the mid-latitude region the southern boundary of the cloud band goes along a smooth curve, which indicates the approximate southern boundary of the jet stream (sketched in dashed line). Usually the cloud-covered areas within the path of the jet stream to the northwest side of the easily-perceivable boundary are susceptible to severe thunderstorms if other necessary factors are present. In this case, severe thunderstorms developed two hours after the time of the image in the boxed area of the figure. Based on our observation, a large proportion of severe weather producing synoptic-scale systems in the CONUS have similar cloud appearances as in this example.

To more precisely determine locations where severe thunderstorms could potentially form, meteorologists usually look for clouds in "*comma shapes*" [28], [29] near the leading edge of the trough. Such a cloud has a round shape (the head of the comma) and a tail towards the opposite direction of its translation. A comma-shaped cloud patch indicates the presence of positive vorticity advection (PVA) in the atmosphere. This is an area of counterclockwise turning of the air flow in the northern hemisphere, which leads to upward motion and the resultant formation of clouds and often precipitation. When sufficient conditions are present, this area of PVA as indicated by a comma-shaped cloud area can cause the development of severe thunderstorms, most typically in the southern and eastern portion of the comma, including the "tail". A nearby boundary between warm and cold air masses can provide a focal point for the development of thunderstorms. Fig. 1 shows an example of a comma-shaped cloud. A cluster of thunderstorms has formed within the square in the tail of the comma shape in the figure. Because of its distinctive pattern, the comma shape has been one of the most utilized signatures when meteorologists review satellite images for thunderstorms.

It is necessary to mention that the comma shape is still a signature in the synoptic scale and therefore used for generally locating a synoptic-scale weather system rather than individual thunderstorms[5]. Additionally, such a visual signature is not always associated with storms. Other clues need to be incorporated to identify storms. In other words, a synoptic system capable of producing severe thunderstorms would be more likely to have a comma shape and the shape is used as a first approximation of where severe weather is most likely to occur.

The aforementioned visual cloud patterns can often be detected from a single satellite image. However, a single image is insufficient for providing some of the crucial information about the dynamics of a storm system. The cloud shapes are not always clear due to the irregular nature of clouds. In addition, the evolution of storm systems, including cloud areas emerging, disappearing, and deforming, cannot be revealed unless different satellite images in a sequence are compared.

[3]http://www.spc.noaa.gov/climo/online/

[4]Only 2008 records are used as ground-truth for both training and testing so that the change of severe hail criteria from 3/4 of an inch diameter to one inch diameter by the National Weather Service after 2010 does not affect the result.

[5]Individual thunderstorms sometimes develop comma shapes to their clouds, but not until they have become more mature. Such small-scale scale signatures are also more difficult to detect in the satellite image and therefore not the focus of this paper.

In practice, an important step in producing weather forecasts is to review satellite images, during which process meteorologists typically review a series of images, rather than a single one, to capture critical patterns. They use the differences between the key visual clues (*e.g.*, key points, edges, etc.) to track the development of storm systems and better locate the jet streams and comma shapes, among other useful features or patterns. As mentioned in [30], human heuristics are helpful in improving weather forecasts. Following such a guideline, our algorithm attempts to simulate meteorologists' cognitive processes in analyzing temporally adjacent satellite image frames. In particular, two types of motions are regarded crucial to the storm visual patterns: the large-scale cloud translation with the jet stream and the local cloud rotations producing the comma shapes. We consider both types of motions for extracting synoptic storm signatures.

D. Objectives

The algorithm proposed in this paper aims at assisting the prediction of severe thunderstorms (including hailstorms and tornadoes) from a different perspective from NWP. As an initial attempt in this direction, we focus on short-term severe storm prediction. Synoptic scale features abstracted from known visual storm patterns as aforementioned are extracted from satellite images and used as clues to locate regions where thunderstorms would be more likely to happen in the near future. Machine learning is involved in the classification of candidate storm regions by inspecting historical thunderstorm data. The regions, as extracted from synoptic scale features, are not necessarily associated one-to-one with individual severe thunderstorms or groups of such storms, *i.e.*, a region may contain multiple outbreaks of severe thunderstorms and each particular thunderstorm needs to be more precisely located by other techniques. The algorithm in fact simulates meteorologists' cognition processes when reviewing satellite images as a part of the weather prediction process. It could be a helpful tool in automating and refining the current weather prediction. It can also be used to provide assistance to meteorologists so that they can be more efficient and less strained in their daily work.

E. Outline

The rest of the paper is organized as follows. Section II introduces the approach to extract storm signatures and construct storm features. Section III reports the approach and results of the machine learning module that accepts or rejects the extracted storm features. In particular, quantitative benchmarks of the classifier and qualitative case studies that applies the whole workflow are given to demonstrate the effectiveness of the proposed algorithm. Finally, we present the future work and conclude the paper in Section IV.

II. Storm Feature Extraction

Our algorithm extracts synoptic-scale visual cloud features from satellite images. The cloud motion observed from image sequences is analyzed in depth. In order to simulate human cognition and let computers perceive comma shapes introduced earlier, the visual signatures are further abstracted into basic cloud motion components: translation and rotation. Several properties and measurements related to these two components are extracted from cloud motions and analyzed in the following steps.

Fig. 2 illustrates the workflow of the whole system. We employ the *optical flow* between every two adjacent satellite images to capture the cloud motion and discover vortex areas that could potentially be associated with severe thunderstorms. Using the historical records of thunderstorms, vortex features are constructed and a thunderstorm classifier is trained. Given an arbitrary query image sequence, vortices are extracted in the same way and then categorized by the classifier.

In particular, for the feature extraction operations, which will be applied both to the training data and the query data, two steps are carried out. The system first estimates a dense optical flow field describing the cloud motion between adjacent image frames. Second, local vortices are identified with the optical flow and vortex descriptors are constructed. The vortex descriptors combine information from both the visual features and the historical records of severe thunderstorms. The following subsections describe details of these steps.

A. Robust Optical Flow Estimation

Optical flow is a basic technique for motion estimation in image sequences. Given two images $I_t(x, y)$ and $I_{t+1}(x, y)$, the optical flow $\vec{F}_t(x, y) = \{U_t(x, y), V_t(x, y)\}$ defines a mapping for each pixel $g : (x, y) \mapsto (x + U_t(x, y), y + V_t(x, y))$, so that $I_t(g(x, y)) \approx I_{t+1}(x, y)$. The vector field $\vec{F}_t(x, y)$ can be therefore regarded as the pixel-wise motions from image $I_t(x, y)$ to $I_{t+1}(x, y)$. Several approaches for optical flow estimation have been proposed based on different optimization criteria [31]–[33]. In this work we adopt the pyramid Lucas-Kanade algorithm [34] to estimate a dense optical flow field between every two neighboring image frames.

The GOES-M satellite is in a geostationary orbit and thus remains over the same point on the earth at all times. As a result, the only objects moving along a satellite image sequence are the clouds and the optical flow between two images indicates the cloud motion. However, the non-rigid and dynamic nature of clouds makes the optical flow estimation noisy and inaccurate. Fig. 3(a) shows the optical flow estimation result of a satellite image in respect to its previous frame in the sequence. Though the result correctly reflects the general cloud motion, flows in local areas are usually noisy and do not precisely describe the local motions. As to be introduced later, the optical flow properties adopted in this work involve the gradient operation of the flow field. Thus it is important to get reliable and smooth optical flow estimation. To achieve this goal, we both pre-process the images and post-process the estimation results. Before calculating the optical flow between two images, we first enhance both of them using the histogram equalization technique[6] [35]. As a result, more fine details on the images are enhanced for better optical flow estimation.

[6]Each channel is enhanced separately. On a given channel, the same equalization function (estimated from the first frame of the sequence) is applied to both images so that they are enhanced by the same mapping.

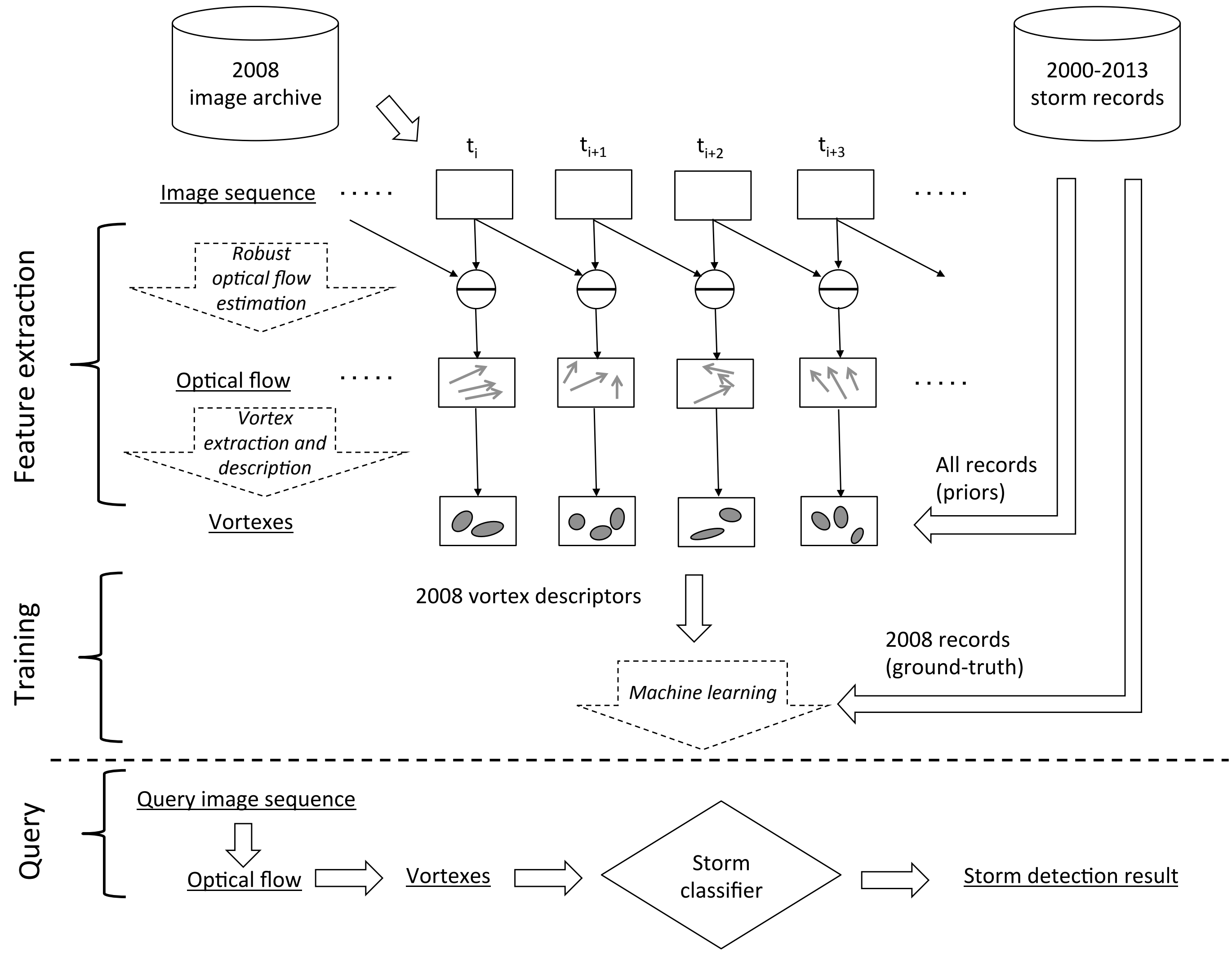


Fig. 2. Workflow of the severe thunderstorm detection system. Optical flows between adjacent image frames are estimated and vortices are extracted base on the flows. Vortex descriptors are then constructed using both the visual information and historical records of severe thunderstorms. Machine learning is performed on the vortex descriptors for the detection of locations of severe thunderstorm.

After the initial optical flow estimation, we smooth the optical flow by applying an iterative update operation based on the Navier-Stokes equation [36] for modeling fluid motions. Given a flow field $\vec{F}(x, y)$, the equation describes the evolving of the flow over time t:

$$\frac{\partial \vec{F}}{\partial t} = (-\vec{F} \cdot \nabla)\vec{F} + \nu \nabla^2 \vec{F} + \vec{f}\ , \tag{1}$$

where $\nabla = (\frac{\partial}{\partial x}, \frac{\partial}{\partial y})$ is the gradient operator, ν is the viscosity of the fluid, and $\vec{f}(x, y)$ is the external force applied at the location (x, y).

The three terms in Eq. (1) correspond to the advection, diffusion, and outer force of the flow field respectively. In the method introduced in [37], an initial flow field is updated to a stable status by iteratively applying these three transformations one by one. Compared with a regular low-pass filtering to the flow field, this approach takes into account the physical model of fluid dynamics, and therefore better approximates the real movement of a flow field. We adopt a similar strategy to smooth the optical flow in our case. The initial optical flow field $\vec{F}_t(x, y)$ is iteratively updated. Within each iteration the time is incremented by a small interval Δt, and there are three steps to get $\vec{F}_{t+\Delta t}(x, y)$ from $\vec{F}_t(x, y)$:

1) add force: $\vec{F}^{(1)}_{t+\Delta t} = \vec{F}_t + \vec{f}\Delta t$;
2) advect: $\vec{F}^{(2)}_{t+\Delta t} = adv(\vec{F}^{(1)}_{t+\Delta t}, -\Delta t)$;
3) diffuse: $\vec{F}_{t+\Delta t} = FFT^{-1}(FFT(\vec{F}^{(2)}_{t+\Delta t})e^{-\nu k^2 \Delta t})$.

The first step is simply a linear increment of the flow vectors based on the external force. In the second step, the advection operator $adv(\cdot)$ uses the flow $\vec{F}^1_{t+\Delta t}(x, y)$ at each pixel (x, y) to predict its location $(x_{t-\Delta t}, y_{t-\Delta t})$ Δt time ago, and updates $\vec{F}^{(2)}_{t+\Delta t}(x, y)$ by $\vec{F}^{(1)}_{t+\Delta t}(x_{t-\Delta t}, y_{t-\Delta t})$. In the last step, the diffusion operation is a low-pass filter applied in the frequency domain, where k is the distance from a point to the origin, and the bandwidth of the filter is determined by the pre-defined fluid viscosity ν and Δt (to be specified later). We do not enforce the mass conservation like the approach in [37] because the two-dimensional flow field corresponding to the cloud movement is not necessarily divergence free. In fact, we use the divergence as a feature for detection of locations of severe thunderstorms in the following procedures.

After several iterations of the update, the flow field converges to a stable status and becomes smoother. The iteration number is typically not large, and we find that the final result is not very sensitive to the parameters ν and Δt. In our system we set the iteration number to 5, the fluid viscosity ν to 0.001, and time interval Δt to 1. The noisy optical flow estimation from the previous stage is treated as the external force field $\vec{f}$, and initially $\vec{F}_t = 0$. Fig. 3(b) shows the smoothed flow field (*i.e.*, $\vec{F}_{t+5\Delta t}$) from the noisy estimation (Fig. 3(a)). Clearly the major motion information is kept and the flow field is smooth for further analysis.

B. Flow Field Vortex Extraction

As introduced in Section I, the rotating and diverging of local cloud patches are two key signatures for detection of locations of severe thunderstorms. In the flow field, these two kinds of evidence are embodied by the *vorticity* and *divergence*. The vorticity of a vector field $\vec{F}(x,y) = \{U(x,y), V(x,y)\}$ is defined as:

$$\begin{aligned}\vec{\omega}(x,y) &= \nabla \times \vec{F}(x,y) \\ &= \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right) \times (U(x,y), V(x,y), 0) \\ &= \left(\frac{\partial V(x,y)}{\partial x} - \frac{\partial U(x,y)}{\partial y}\right) \vec{z}\,;\end{aligned} \tag{2}$$

and the divergence is defined as:

$$\begin{aligned}\mathrm{div}\vec{F}(x,y) &= \nabla \cdot \vec{F}(x,y) \\ &= \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}\right) \cdot (U(x,y), V(x,y)) \\ &= \frac{\partial U(x,y)}{\partial x} + \frac{\partial V(x,y)}{\partial y}\,.\end{aligned} \tag{3}$$

It has been proven that for a rigid body, the magnitude of vorticity at any point is twice the angular velocity of its self rotation [38] (the direction of the vorticity vector indicates the rotation's direction). In our case, even though the cloud is non-rigid, we can regard the vorticity at a certain point as a description of the local rotation.

To reveal the rotation and divergence more clearly, we apply the Helmholtz-Hodge Decomposition [39] to decompose the flow field to a solenoidal (divergence-free) component and a irrotational (vorticity-free) component. Fig. 3(c) and Fig. 3(d) visualize these two components respectively. In both figures, areas with densely overlapped vectors plotted are the places with high vorticity or divergence.

The divergence-free component of the flow field is useful for detecting vortices. On this component, we inspect the local deformation tensor

$$\nabla\vec{F} = \begin{bmatrix} \partial U/\partial x & \partial V/\partial x \\ \partial U/\partial y & \partial V/\partial y \end{bmatrix}\,,$$

which can be decomposed to the symmetric (strain) part $S = \frac{1}{2}(\nabla\vec{F} + \nabla\vec{F}^T)$ and the asymmetric (rotation) part $\Omega = \frac{1}{2}(\nabla\vec{F} - \nabla\vec{F}^T)$. The Q-criterion introduced by [40] takes the difference of their norms:

$$Q = \frac{1}{2}(\|\Omega\|^2 - \|S\|^2) = \frac{1}{4}\|\vec{\omega}\|^2 - \frac{1}{2}\|S\|^2\,, \tag{4}$$

where $\|\cdot\|$ is the Frobenius matrix norm. The Q-criterion measures the dominance of the vorticity. When $Q > 0$, *i.e.*, the vorticity component dominates the local flow, the corresponding location is regarded to be in a vortex region. Fig. 3(e) and Fig. 3(f) visualize the vorticity and Q-criterion of the flow field in Fig. 3(c). In Fig. 3(e), pixels are tinted by the corresponding magnitude of vorticity, and different colors mean different rotating directions. In Fig. 3(f), the vortex regions are highlighted in red. Clearly only pixels with a dominant vorticity component are selected as vortices by the Q-criterion. It is apparent that these vortices are more prone to be located inside the area where severe thunderstorms were reported (highlighted by yellow boundaries) than the removed high-vorticity pixels. It is also observed that the vortices are typically in narrow bending comma shapes as described in Section I-C. Therefore, they are properly to be regarded as the potential storm elements.

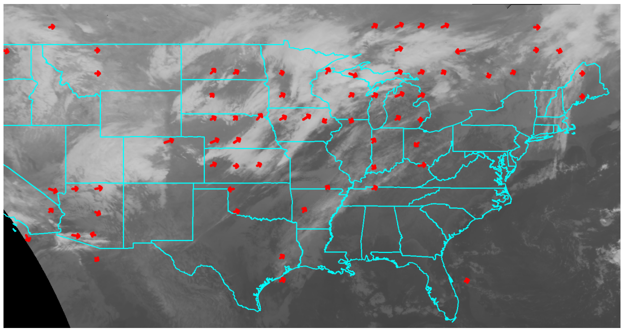
C. Vortex Descriptor

Not all the vortex regions in Fig. 3(f) are related to the location of reported severe thunderstorms[7]. As a result we built a descriptor for each extracted vortex and apply it to a machine learning module (to be introduced later). We introduce the vortex descriptor in this subsection.

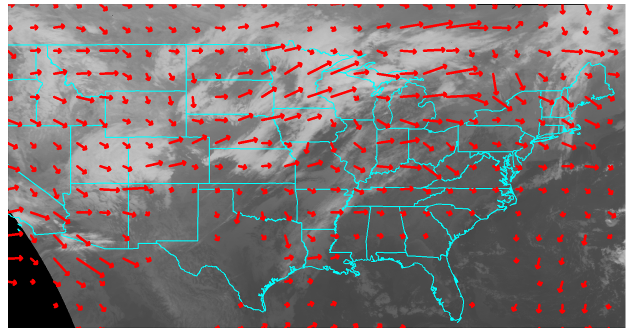
Denoting a certain vortex region in a satellite image as Φ and the area of Φ as $\Pi(\Phi)$, the following visual clues, both static ones from a single image and dynamic ones from the optical flow with respect to the previous image frame, are considered in our approach.

1) Mean channel 3 brightness:
$$w_1(\Phi) = \frac{\sum_{(x,y)\in\Phi} I^{(3)}(x,y)}{\Pi(\Phi)}\,,$$
where $I^{(3)}(x,y)$ is the brightness of pixel (x,y) in the channel 3 image.
2) Mean channel 4 brightness:
$$w_2(\Phi) = \frac{\sum_{(x,y)\in\Phi} I^{(4)}(x,y)}{\Pi(\Phi)}\,,$$
where $I^{(4)}(x,y)$ is the brightness of pixel (x,y) in the channel 3 image.
3) Mean optical flow intensity:
$$w_3(\Phi) = \frac{\sum_{(x,y)\in\Phi} \|\vec{F}(x,y)\|}{\Pi(\Phi)}\,,$$
where $\vec{F}(x,y) = (U(x,y), V(x,y))$ is the optical flow at pixel (x,y) computed between the previous and current image frames.
4) Mean optical flow direction:
$$w_4(\Phi) = \frac{\sum_{(x,y)\in\Phi} \|\theta(x,y)\|}{\Pi(\Phi)}\,,$$
where $\theta(x,y) = \tan^{-1}\left(\frac{V(x,y)}{U(x,y)}\right)$ is optical flow $\vec{F}(x,y)$'s direction.

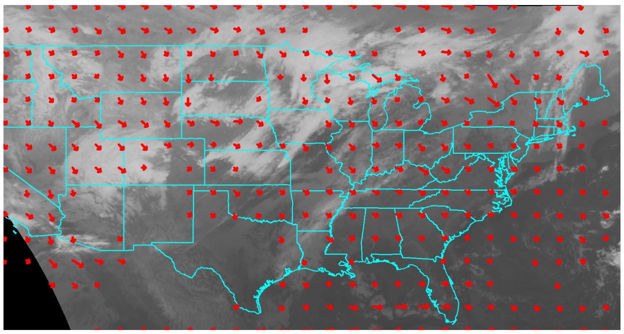
[7]The types of vortices outside of the CONUS are unknown because of lack of records.



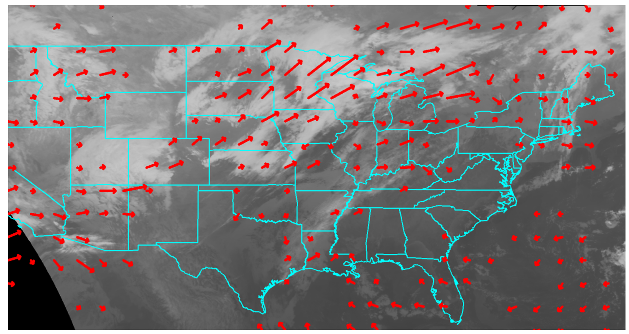
(a) Noisy optical flow estimation



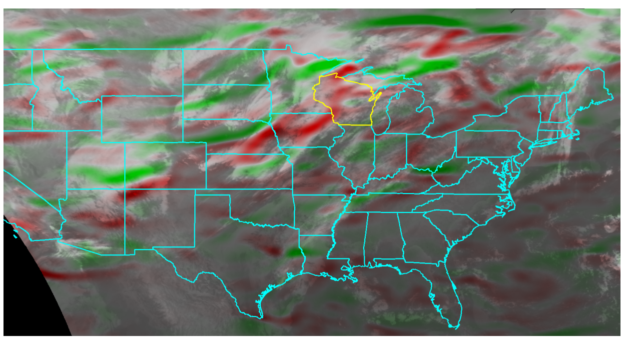
(b) Smoothed optical flow



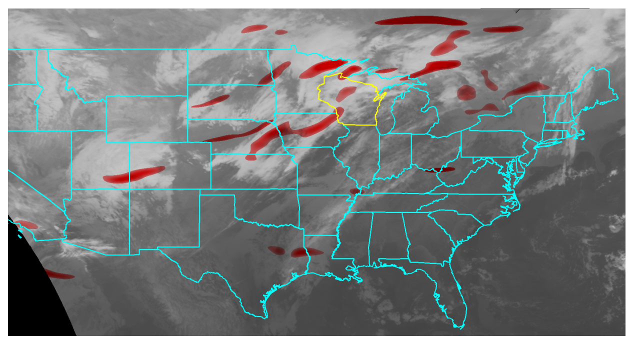
(c) Divergence-free component of the smoothed flow



(d) Vorticity-free component of the smoothed flow



(e) Magnitude of vorticity



(f) Vortex cores and their expansions

Fig. 3. Sample results of the optical flow analysis. GOES-M satellite images covering the continental US area are analyzed (to avoid showing too many missing pixels on the southwest corner, a small stripe of US west coast is not shown on the figures, though these pixels are included in our analysis). Between two adjacent frames (with 30 minutes' interval) in a satellite image sequence, the optical flow from the former frame to the latter one is estimated and processed. The results are plotted on the second frame. Only large enough flow vectors are drawn on figure (a-d). (a) The optical flow estimated by the Lucas-Kanade algorithm. (b) The smoothed optical flow by applying an iterative update based on the Navier-Stokes equation. (c) The divergence-free component of the smoothed flow field. (d) The vorticity-free component of the smoothed flow field. (e) Visualization of the vorticity. Vorticity vectors of the 2D flow field are perpendicular to the image plane. Pixels with vorticity vectors toward the viewer (counter-clockwise rotations) are tinted in green color; and the pixels with vorticity away from the viewer (clockwise rotations) are tinted in red color. The saturation of color means the magnitude of the corresponding vorticity vector. (f) Vortex regions detected by the Q-criterion.

5) Mean vorticity on the solenoidal optical flow component:

$$w_5(\Phi) = \frac{\sum_{(x,y)\in\Phi} \omega(x,y)}{\Pi(\Phi)} ,$$

where $\omega(x,y)$ is the vorticity value of the solenoidal component of optical flow $\vec{F}(x,y)$ (sign of the value indicates the vorticity's direction).

6) Mean divergence on the irrotational optical flow component:

$$w_6(\Phi) = \frac{\sum_{(x,y)\in\Phi} \mathrm{div}\vec{F}(x,y)}{\Pi(\Phi)} ,$$

where $\vec{F}(x,y)$ is the divergence of the irrotational component of optical flow $\vec{F}(x,y)$.

7) Maximal Q-value of the vortex:

$$w_7(\Phi) = \max_{(x,y)\in\Phi} Q(x,y) ,$$

where $Q(x,y)$ is defined in Eq. (4).

We also considered the spatial and temporal distributions of the severe thunderstorms that occurred in the CONUS from 2000 to 2013 using all the historical records of severe thunderstorms through these years[8]. This is shown in Fig. 4. We divided the CONUS area into $4° \times 4°$ grids. Inside each grid, we count the occurrence of severe thunderstorms around each given date (*e.g.*, July 4th) in the history. Storms that occurred from 5 days before to 5 days after the queried date in each year are counted (*e.g.*, for July 4, storms from June 29 to July 9 in every year, a total of 154 days, are included[9]). Denoting the total severe thunderstorm count in grid (i, j) (the grid counted i-th from the left and j-th from the top) for date d is $N_d(i, j)$, the average storm density in the grid is $\rho_d(i, j) = \frac{N_d(i,j)}{154}$.

To describe the severe thunderstorm prior for vortex Φ (extracted from date d), we took the average storm densities of the pixels it covers:

$$w_8(\Phi) = \frac{\sum_{(x,y)\in\Phi} \phi_d(T(x, y))}{\Pi(\Phi)} ,$$

where $T : (x, y) \mapsto (i, j)$ is a mapping from the pixel (x, y) to the grid index (i, j) it belongs to.

With all the features introduced above, for vortex Φ we constructed a vortex descriptor $X(\Phi) = (w_1, w_2, w_3, w_4, w_5, w_6, w_7, w_8)$. Descriptors of training data are fed into a machine learning algorithm, and eventually we obtain a classifier $C(\cdot)$. For an arbitrary vortex Φ, $Y(\Phi) = C(X(\Phi))$ is the predicted status calculated by the classifier. We introduce the machine learning approach in the next section.

III. Storm System Classification and Detection

The vortex regions with large Q values correspond to regions with strong cloud rotations and our approach extracts them as vortices. However, they are not necessarily related to the location of severe thunderstorms because a number of factors go into creating a favorable atmospheric set up for severe weather. The cloud rotation we are tracking is one important ingredient, but there is rotation in clouds over many parts of the world with no resulting severe weather. The development of severe thunderstorms requires sufficient warmth and moisture at the low levels of the atmosphere and a layer of drier and cooler air aloft that contributes to deep instability through multiple layers of the atmosphere. Certain geographical areas have a favorable climatology for severe weather in which these necessary conditions recur frequently (*i.e.*, the central plains of the US) while other geographical areas have factors that inhibit severe weather (*i.e.*, over any body of cold water).

In order to create reliable storm detection, we embedded these vortices and their descriptors into a machine learning framework. Using the descriptors and ground-truth labels of the extracted candidate vortex regions, a classifier was trained to distinguish vortices related to severe thunderstorms from other vortices.

A. Training Vortices and Ground-truth Labels

We extracted vortices from 2008 GOES-M satellite images and used reports of severe thunderstorms in the same year to assign a ground-truth label for each vortex in the CONUS. A classifier is trained to associate the vortex features with the storm labels. Image sequences in the first ten days of each month are used for training and those from the 18th to the 22nd days of each month are used for testing (so that the weather systems in the training and testing data are relatively far apart and independent). In order to get enough difference yet not to lose too much detail, we choose the sampling interval between every two images in a sequence to be 30 minutes. On a computer with a 2 GHz single-core CPU, it took about 1 minute to extract vortexes between two image frames of size 1024×2048.

We used the historical severe storm database to assign ground-truth labels to detected vortices. We denote the historical storm database as $\mathscr{D}$ and each storm entry as (u_i, v_i, t_i) where (u_i, v_i) is the location (latitude and longitude values in degrees) and t_i. Assuming a vortex is detected at location (u, v)[10] and time t (associated with the second frame of the corresponding image pair), if at least one record $(u_i, v_i, t_i) \in \mathscr{D}$ within Δd degrees of latitude and longitude exists between the moments $t - \tau_1$ and $t + \tau_2$ (*i.e.*, $\|u_i - u\| < \Delta d$ and $\|v_i - v\| < \Delta d$ and $t - \tau_1 \leq t_i \leq t + \tau_2$), the vortex is assigned with a positive (storm-related) label; otherwise it is assigned with a negative (no-storm) label.

The spatial and temporal threshold Δd, τ_1, and τ_2 for labeling positive vortices affect the stability and resolution of the classifier. Classifiers with higher resolution are more difficult to train and less stable without enough training data. In our experiment, we selected these thresholds empirically based on our needs and the training data. We chose $\tau_1 = 0.5hr$ and $\tau_2 = 2hr$ because we want to find out vortex visual features that are directly related to on going and near-future thunderstorms. We used storms starting right before a vortex observation to label it (by setting $\tau_1 = 0.5hr$) because we do not want the classifier to miss the strong visual features related to ongoing storms, though we only hope to predict future storms. For the spatial resolution, we chose $\Delta d = 2.5°$ for searching nearby thunderstorms because that is relatively the maximal distance that a storm can affect a location (in 2 hours). Additionally, no two vortices will be exactly alike as far as in what portion of the system any severe thunderstorms will occur. Variations in the location of the warmest air, the air with the greatest moisture supply, the location and direction of the jet stream and any low-level jet structures will result in variations in the location and movement of severe thunderstorms that develop relative to the center of a particular vortex.

Lastly, we only assigned labels for vortices inside the CONUS (for the year of 2008). For vortices outside of the

[8]We noticed that NWS revised the criteria for hail reporting from 3/4 to 1 inch in diameter in 2010 so less hails were reported after that. The criteria change affect the absolute storm numbers for different locations. However, the change has little effect on the severe thunderstorms' relative spatial distribution used in the machine learning system. The training and testing data used in the experiment are labeled by the same criteria too.

[9]We count severe thunderstorms around February 28th in non-leap years for storm statistics on February 29th.

[10](u, v) is the geometric center of the vortex.

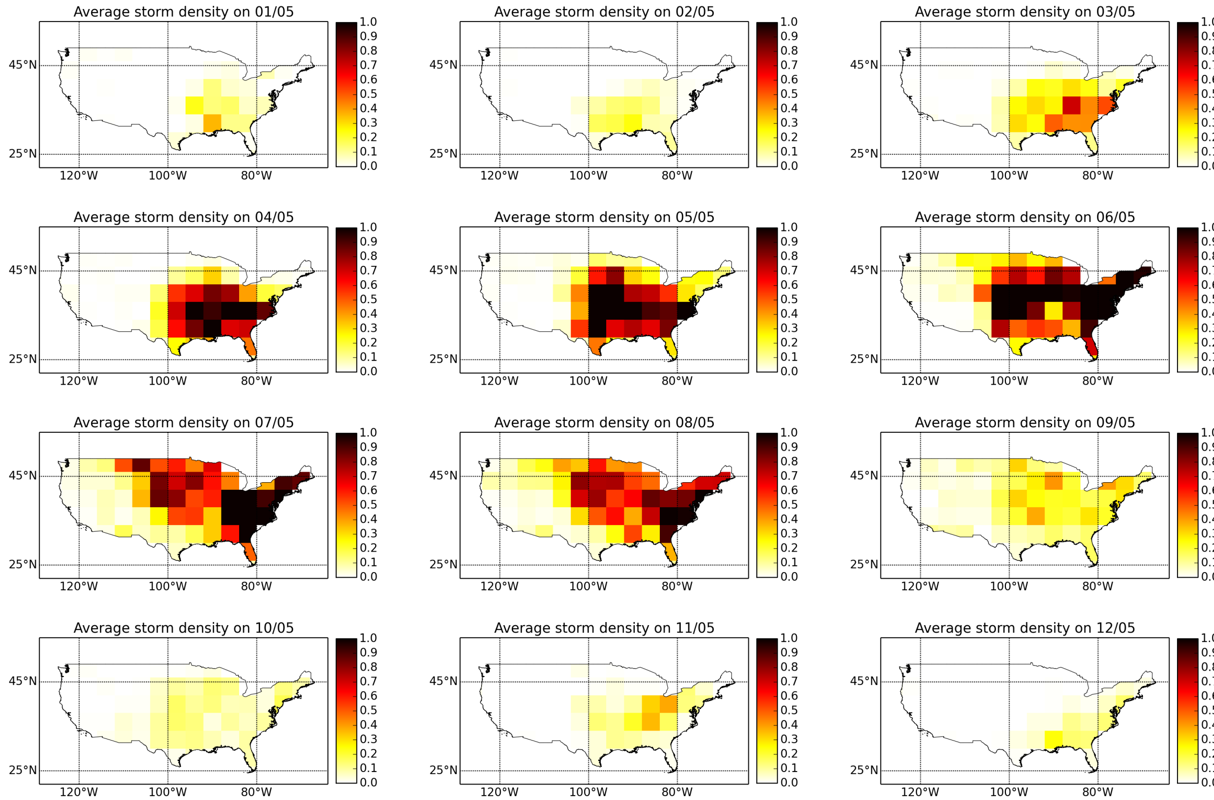


Fig. 4. Average storm densities (number of severe thunderstorms per 30,000 km^2) for the fifth day of each month in US continent. Darker color means higher storm density. The statistics are based on the historical records of severe thunderstorms from 2000 to 2013. Around each given date (+/-5 days), we count the total number of severe thunderstorms reported in each state across the past 14 years. The numbers are then divided by the total number of days and the areas of corresponding states to calculate the average densities.

CONUS, we had no historical data indicating whether or not they are storm-related, so we did not use them for training and testing. Applying our algorithm in real case requires training with specific data.

On the 120 selected days (10 for each month), a total of 1,379 severe thunderstorm-related vortices were discovered by the above method. We randomly selected the same number of vortices in the CONUS that were not storm-related from these days. A binary classifier is then trained on the 2,758 training vortex samples.

B. Random Forest Classifier

We trained a random forest classifier [41] using the training data to distinguish severe thunderstorm-related vortices from vortices not affiliated with a storm. The features for each vortex are described in Section II-C and the ground-truth is defined in Section III-A.

Using the training data, the random forest learns rules to determine whether a vortex of certain descriptor is related to the occurrence of severe thunderstorms. The random forest is essentially an ensemble learning algorithm that makes predictions by combining the results of multiple individual decision trees trained from random subsets of the training data. We chose this approach because it has been found to outperform a single classifier (*e.g.*, SVM). In fact, this strategy resembles the *ensemble forecasting* [42] approach in numerical weather prediction, where multiple numerical results with different initial conditions are combined to make a more reliable weather forecast. For both numerical and statistical weather forecasting approaches, the ensemble approaches help to improve the prediction qualities.

C. Testing Vortex Samples and Testing Methods

To develop a benchmark for the performance of the vortex classifier, we used image sequences from the 18th to the 22nd days of each month in 2008 as the testing data. On each selected day, the image sequence (with a 30-minute interval) from 12:00 GMT to 22:00 GMT is used to extract dynamic cloud vortices. The test dataset contains a total of 21,064 vortices, among which 563 are labeled as positive, and the rest 20,501 samples are negative. The descriptors of these vortices are classified by the classifier. We then evaluate the classification results in two ways.

Firstly, we assigned the testing vortices with ground-truth labels in the same way as how the training ground-truth labels are assigned. The predicted storm status for each test vortex is compared with the ground-truth. In this way the classification accuracy for single vortices are evaluated. Table I presents the classification performance of both the training data (by cross validation) and the testing data. The classifier shows consistent performances on both the training set and the test set. To demonstrate the effect of including geographic storm

TABLE I
CLASSIFICATION PERFORMANCE AND CONTRIBUTIONS OF VISUAL FEATURES AND HISTORICAL PRIORS

		All features	Visual only	Prior only
Training set (cross validation)	Overall	85.9%	68.2%	79.3%
	Sensitivity	88.4%	60.5%	78.0%
	Specificity	83.5%	75.9%	80.6%
Testing set	Overall	78.2%	75.5%	76.1%
	Sensitivity	78.1%	35.7%	70.3%
	Specificity	79.9%	76.5%	76.3%

Note: Training set contains 1,379 severe thunderstorm-related and 1,379 no-storm vortex regions from 120 days in 2008. 10-fold cross validation is performed in the evaluation. Testing set contains 20,501 storm-related cells and 563 no-storm cells from 60 days far away from the training data. The feature vector for each vortex is composed by both visual features and storm priors (see Section II-C). Beside the merged features (results shown in the first column), we test the two types of features separately (results shown in the second and third columns).

priors in the classification, we also trained and tested the random forest classifiers only on the visual features and the storm density separately. Clearly none of the visual features and the historical storm density standalone performs as well as combining them together. Though the storm density contributes more in the classification, including the visual features further increases accuracy. Combining these two types of features significantly enhances the classification performance. The amount of improvement seen in the classification performance using both visual features and the historical density of severe thunderstorms is similar to the degree of improvement gained by various enhanced iterations of the NWP models used by meteorologists in recent years [43]–[45].

Secondly, we evaluated the classifier's ability to predict the future locations of severe thunderstorms. Instead of comparing each vortex's classification result with its ground-truth label, which reflects its status within the period of $(t-0.5hr, t+2hr)$ (t is the timestamp of the vortex), we observed how much time in advance a severe thunderstorm can be detected by our algorithm. Given a testing vortex extracted at the location (u, v) and time t, we find the earliest time $t' = T(u, v, t)$ that a neighboring storm $(u', v', t') \in \mathscr{D}$ developed since two hours before t, *i.e.*,

$$t' = \min_{t_i}\{(u_i, v_i, t_i) \in \mathscr{D} \text{ and } t_i > t - 2hr \text{ and } \|u_i - u\| < 2.5^\circ \text{ and } \|v_i - v\| < 2.5^\circ\} .$$

The value $\Delta t = T(u, v, t) - t$ indicates the time difference between an observed vortex and the first severe thunderstorm to form nearby. For a vortex with $\Delta t > 0$, a future severe thunderstorm is predicted if the classifier can label a vortex as severe thunderstorm-related. The percentages of severe thunderstorms predicted for vortices with different $\Delta t > 0$ are shown in Fig. 5. The results show that the proposed algorithm can effectively predict future severe thunderstorms within several hours of the detection of vortices and the prediction reliability increases as Δt decreases. Particularly, it shows a reasonable prediction rate for the development of nearby

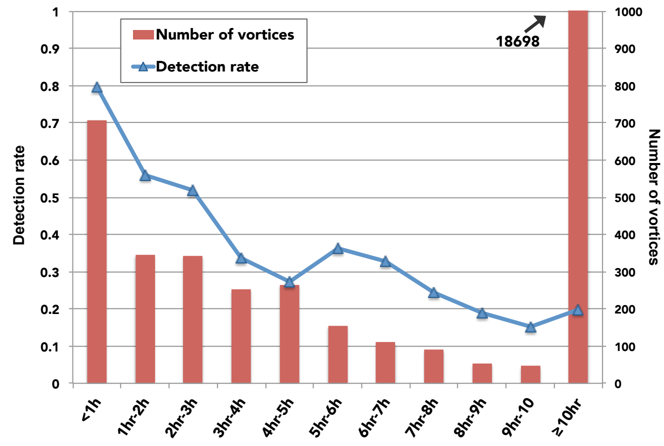


Fig. 5. Severe thunderstorm detection rate with different amounts of time in advance. The distribution of the time differences (Δt) between the testing vortex samples and storms are plotted in red bars (the last bar is not fully shown because it is much higher than others). And the percentages of vortices categorized as severe thunderstorms with different Δt are plotted in a blue curve. The figure shows that the predictivity of the classifier is generally reliable within few hours of time in advance. The detection rate decreases as the time difference becomes longer.

storms within a period of three hours after the observation of visual vortices in satellite images.

It is necessary to emphasize that the classification is performed on already extracted candidate vortex regions where local clouds are rotating. Most of the cloud-covered regions on the satellite images without obvious rotational motions have been already eliminated before the algorithm proceeds to the classification stage. Therefore in practice the system achieves a good overall precision for detecting the locations of severe thunderstorms. In addition, the classification results for individual vortex regions can be further corrected and validated based on their spatial and temporal neighboring relationships.

D. Case Study

To demonstrate the usefulness of the proposed vortex detection and classification algorithm in real scenarios, we present several case studies where our algorithm is applied for automatic storm detection from satellite images.

Three different cases are presented and the results are visualized in Figs. 6, 7, and 8[11]. An active severe thunderstorm reporting day, a clear storm-free day, and a cloudy storm-free day are selected to demonstrate the adaptiveness of our algorithm. These three image sequences are not included in the training process of the storm classifier. In these figures, automatically detected vortices are filled with colors on each image. vortices classified as storms are colored in red, and the green vortices are categorized as storm-free cases by the machine learning module. The US state boundaries are drawn on each image to show the alignment between the satellite images and the geo-coordinate system. In particular,

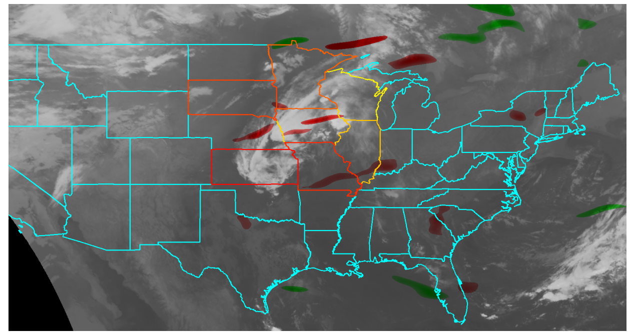
[11] We applied the algorithm on all available data and selected three representative cases to present in the paper. Readers can check the thunderstorm detection results in other cases online at http://riemann.ist.psu.edu/docs/projects/weather/weather_visualizer/.

boundaries of states affected by severe thunderstorms are highlighted by warm colors from red to yellow. The highlighting color indicates how long after the observation storms were reported in the corresponding states. As the color turns to yellow, the highlighted states are affected by storms further in the future. Specifically, the red color means an ongoing severe thunderstorms and the yellow color means a severe thunderstorm six hours after the observation time. Only the first three image frames are displayed on the figures due to space limitation. Each test image sequence in fact contains more than three frames, and the storm detection results are consistent in the subsequent frames for each case.

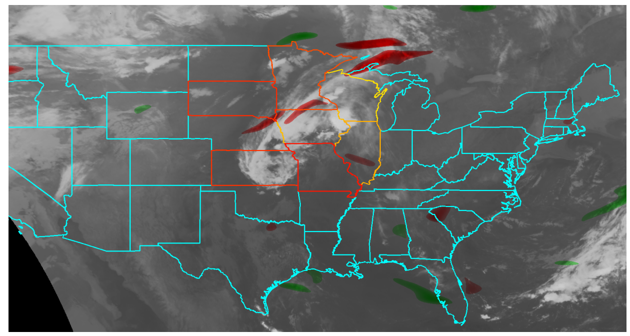
The first case (Fig. 6) is from May 25, 2008, which was an active severe thunderstorm day, when several severe thunderstorms hit the Midwest US. An obvious cloud vortex system is visible on the satellite images and it is related to severe thunderstorms reported in nearby states. Our algorithm captures all the major vortices from the satellite image sequence, and correctly categorizes most of them as severe thunderstorms. The result also shows the capability of the proposed algorithm to assist in short-term storm forecasting. In most cases there will be severe thunderstorms detected around regions where such storms were reported shortly thereafter.

Different from the first case, the second case is strongly negative in that few clouds and vortices can be perceived from the images. Results on Fig. 7 shows that no storm is predicted because few clouds and vortices can be captured by the feature extraction stage of our algorithm. In this case, it is easy for the algorithm to identify features correctly.

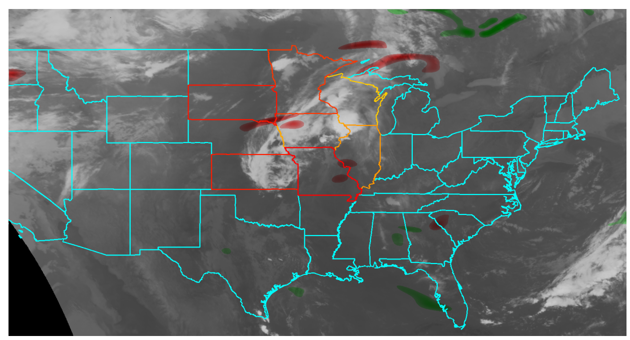
In the third case, several cloud systems can be observed above the CONUS, and several vortices are detected by the optical flow analysis. In reality, even though some cloud vortex systems look dominant on the image sequence, there is no severe thunderstorm reported on that day. Results shown on Fig. 8 demonstrate that generally the machine learning module successfully avoids labeling the detected cloud vortices as severe thunderstorms, though some classification errors occur. This shows that the system does not over-predict the occurrence of storms. This benefit is also validated by the standalone test for the classifier showing that specificity of the classifier is as high as its sensitivity.



(a) 2008/05/25 12:15 GMT



(b) 2008/05/25 12:45 GMT



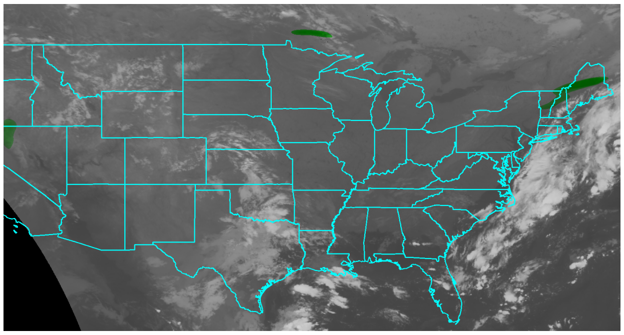
(c) 2008/05/25 13:15 GMT

Fig. 6. The example case of applying the proposed algorithm to satellite image data from May 25, 2008, when Major severe thunderstorms occurred in the central, eastern, and southeastern states of US. Vortices associated with severe thunderstorms are colored in red and rejected vortices are colored in green. States hit by severe thunderstorms are highlighted in warm colors. A red state boundary indicates the state is hit by an ongoing storm based on the historical report. As the highlighting color turns to yellow, the corresponding state is affected by more future storms as far as 6 hours away from the observation. Three successive image frames are displayed. Multiple vortices are detected and classified as severe thunderstorm systems in each frame and the locations of the vortices well match the ground-truth.

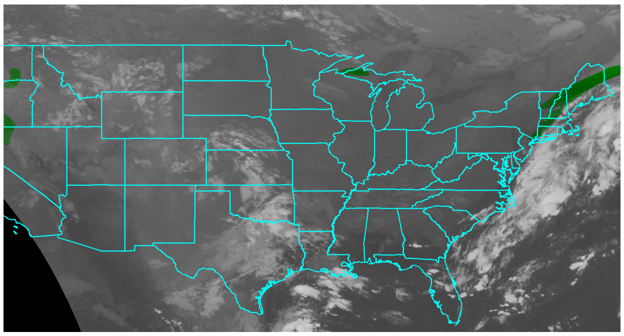
IV. Conclusions and Future Work

In this paper, we present a severe thunderstorm detection algorithm that locates storm visual signatures from satellite images. The algorithm automatically analyzes the visual features from satellite images and incorporates the historical meteorological records to make predictions of severe thunderstorm locations. As opposed to traditional ways of weather forecasting, our approach primarily relies on the visual information from images and tries to extract high-level storm signatures as meteorologists usually do manually. Without using any sensory measurement (*e.g.*, temperature, air pressure, etc.) as numerical weather forecasting does, the algorithm captures global cloud patterns solely from the satellite images, which is helpful in synoptic-scale storm prediction. Additionally, the algorithm takes advantage of big data, using historical severe thunderstorm reports in the past years together with the satellite image archives to learn the correspondence between visual image signatures and the occurrences of current and future severe thunderstorms. Experiments and case studies show that the algorithm is effective and robust.

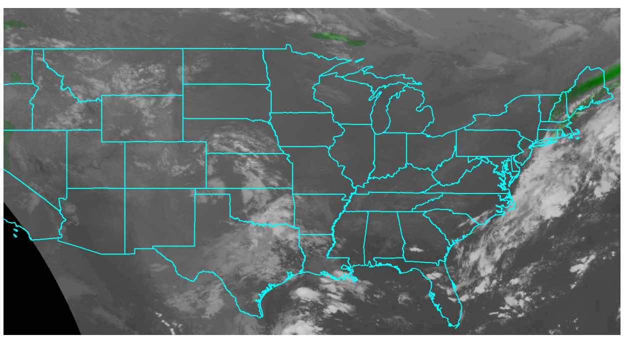
Properties of optical flows between every two images in a sequence are the basic visual clues adopted in the work. The algorithm can estimate robust and smooth optical flows between two images and determine the rotations in the flow field. Unlike numerical weather forecasting methods that are sensitive to noise and initial conditions, our approach can



(a) 2008/03/25 12:15 GMT

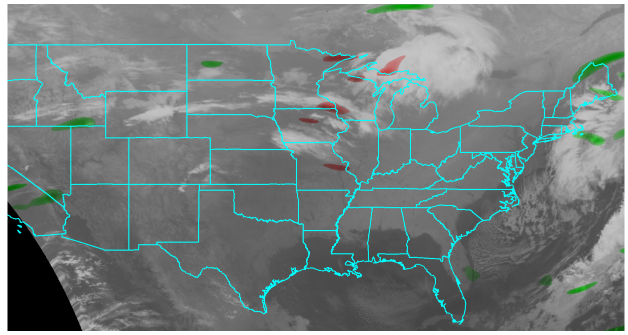


(b) 2008/03/25 12:45 GMT

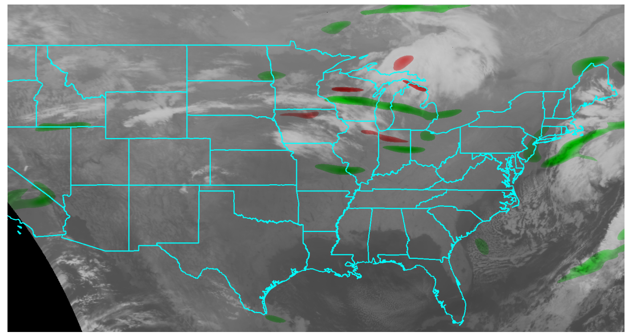


(c) 2008/03/25 13:15 GMT

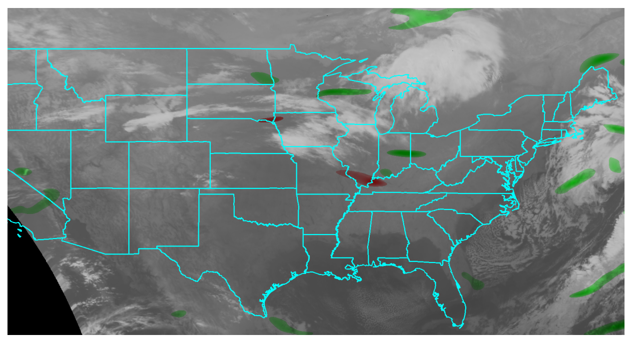
Fig. 7. The example case of applying the proposed algorithm to satellite image data from March 25, 2008, when the US continent is mostly clear. The algorithm does not produce any false alarm in this case.



(a) 2008/09/18 12:15 GMT



(b) 2008/09/18 12:45 GMT



(c) 2008/09/18 13:15 GMT

Fig. 8. The example case of applying the proposed algorithm to satellite image data from Sep 18, 2008. No storm is reported on the day. Many vortices are detected on the images, and quite a few of them are rejected (colored in green), though some are recognized as false alarms (colored in red).

consistently detect reliable visual storm clues hours before the occurrence of severe weather. Therefore, the results are useful for short term weather forecasts, especially storm forecasts.

The application of historical severe thunderstorm records and machine learning boosts the performance of the storm extraction algorithm. Standalone vortex detection from the optical flow is not sufficient to make reliable predictions. The statistical model trained from historical meteorological data together with the satellite image visual features further selects the extracted vortices and removes vortices not related to severe thunderstorms. In our current algorithm, storm detection is based on individual vortex regions. In reality, multiple vortices tend to appear near each other in synoptic scale weather systems both temporally and geographically. Therefore, taking into account nearby vortices within a frame and tracking synoptic scale weather systems across a sequence can improve the overall reliability of the system. In particular, tracking the development of a synoptic scale weather system will be helpful for analyzing its future trend. This problem is nontrivial because the clouds with synoptic scale weather systems are non-rigid and highly variant, and the same vortices within such a system are not always detectable over time. Future work needs to tackle the problem of non-rigid object tracking to better make use of the temporal information in the satellite image sequences.

Lastly it should be emphasized that weather systems are highly complex and chaotic, so it is always a challenging task to make accurate weather forecasts. The proposed algorithm based on the satellite image visual features is not completely accurate and needs to be further validated. Also, the algorithm only performs short term predictions with a 30-minute resolution, though many severe weather phenomena occur in less than 30 minutes. We expect the algorithm or similar approaches to play a part in supplementing weather forecasts rather than making the forecasts alone. It could be incorporated with other existing weather prediction techniques (*e.g.*, NWP) and used as an automatic tool to refine the current forecasts. It is useful and valuable because it provides forecasts in a different aspect from the numerical approaches. The purpose for developing the new algorithm is not to replace the current weather forecasting models, but to produce additional independent predictions to be combined with other approaches. The proposed approach and experiment is a starting point in exploring a new paradigm to be incorporated as an additional tool into an operational forecasting environment. As a result, we will focus the future study on how to integrate information from multiple data sources and predictions from different models to produce more reliable and timely forecasts of severe thunderstorms.

Acknowledgment

This material is based upon work supported by the National Science Foundation under Grant No. 1027854. Shared computational infrastructure was provided by the Foundation under Grant No. 0821527. Part of the work was done when J. Z. Wang and J. Li were with the Foundation. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the Foundation. We also thank the US National Oceanic and Atmospheric Administration (NOAA) for providing the data used in this research. Siqiong He assisted with data collection. The discussions with Jose A. Piedra-Fernandez of the University of Almeria, Spain, has been very helpful. We would also like to acknowledge the comments and constructive suggestions from anonymous reviewers and the associate editor.

References

[1] P. Lynch, "The origins of computer weather prediction and climate modeling," *Journal of Computational Physics*, vol. 227, no. 7, pp. 3431–3444, 2008.

[2] G. W. Platzman, "The ENIAC computations of 1950-gateway to numerical weather prediction," *Bulletin of the American Meteorological Society*, vol. 60, no. 4, pp. 302–312, 1979.

[3] J. Charney, "The use of the primitive equations of motion in numerical prediction," *Tellus*, vol. 7, no. 1, pp. 22–26, 1955.

[4] F. Yang, "Review of GFS forecast skills in 2013," *National Centers for Environmental Prediction*, 2013.

[5] S. D. Burt and D. A. Mansfield, "The great storm of 15–16 october 1987," *Weather*, vol. 43, no. 3, pp. 90–110, 1988.

[6] R. L. Winkler, "The importance of communicating uncertainties in forecasts: Overestimating the risks from winter storm juno," *Risk Analysis*, vol. 35, no. 3, pp. 349–353, 2015.

[7] A. C. Lorenc and T. Payne, "4D-Var and the butterfly effect: Statistical four-dimensional data assimilation for a wide range of scales," *Quarterly Journal of the Royal Meteorological Society*, vol. 133, no. 624, pp. 607–614, 2007.

[8] J. D. Doyle, C. Amerault, C. A. Reynolds, and P. A. Reinecke, "Initial condition sensitivity and predictability of a severe extratropical cyclone using a moist adjoint," *Monthly Weather Review*, vol. 142, no. 1, pp. 320–342, 2014.

[9] V. F. Dvorak, "Tropical cyclone intensity analysis and forecasting from satellite imagery," *Monthly Weather Review*, vol. 103, no. 5, pp. 420–430, 1975.

[10] R. A. Scofield, "The NESDIS operational convective precipitation-estimation technique," *Monthly Weather Review*, vol. 115, no. 8, pp. 1773–1793, 1987.

[11] X. Peng, Y. Che, and J. Chang, "A novel approach to improve numerical weather prediction skills by using anomaly integration and historical data," *Journal of Geophysical Research: Atmospheres*, vol. 118, no. 16, pp. 8814–8826, 2013.

[12] G. Liu, "Satellite microwave remote sensing of clouds and precipitation," *Observation, Theory and Modeling of Atmospheric Variability: Selected Papers of Nanjing Institute of Meteorology Alumni in Commemoration of Professor Jijia Zhang*, vol. 3, p. 397, 2004.

[13] A. N. Srivastava and J. Stroeve, "Onboard detection of snow, ice, clouds and other geophysical processes using kernel methods," in *Proceedings of the ICML 2003 Workshop on Machine Learning Technologies for Autonomous Space Applications*, 2003.

[14] Y. Hong, "Precipitation estimation from remotely sensed imagery using an artificial neural network cloud classification system," *Journal of Applied Meteorology*, vol. 43, pp. 1834–1852, 2004.

[15] A. Behrangi, K. Hsu, B. Imam, and S. Sorooshian, "Daytime precipitation estimation using bispectral cloud classification system," *Journal of Applied Meteorology and Climatology*, vol. 49, pp. 1015–1031, 2010.

[16] J. C. Price, "Using spatial context in satellite data to infer regional scale evapotranspiration," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 28, no. 5, pp. 940–948, 1990.

[17] K. Bedka, J. Brunner, R. Dworak, W. Feltz, J. Otkin, and T. Greenwald, "Objective satellite-based detection of overshooting tops using infrared window channel brightness temperature gradients," *Journal of Applied Meteorology and Climatology*, vol. 49, no. 2, pp. 181–202, 2010.

[18] K. M. Bedka, "Overshooting cloud top detections using MSG SEVIRI infrared brightness temperatures and their relationship to severe weather over europe," *Atmospheric Research*, vol. 99, no. 2, pp. 175–189, 2011.

[19] S.-S. Ho and A. Talukder, "Automated cyclone discovery and tracking using knowledge sharing in multiple heterogeneous satellite data," in *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2008, pp. 928–936.

[20] T. Zinner, H. Mannstein, and A. Tafferner, "Cb-TRAM: Tracking and monitoring severe convection from onset over rapid development to mature phase using multi-channel Meteosat-8 SEVIRI data," *Meteorology and Atmospheric Physics*, vol. 101, no. 3-4, pp. 191–210, 2008.

[21] C. Keil and G. C. Craig, "A displacement and amplitude score employing an optical flow technique," *Weather and Forecasting*, vol. 24, no. 5, pp. 1297–1308, 2009.

[22] D. Merk and T. Zinner, "Detection of convective initiation using meteosat SEVIRI: implementation in and verification with the tracking and nowcasting algorithm cb-tram," *Atmospheric Measurement Techniques Discussions*, vol. 6, pp. 1771–1813, 2013.

[23] A. N. Evans, "Cloud motion analysis using multichannel correlation-relaxation labeling," *IEEE Geoscience and Remote Sensing Letters*, vol. 3, no. 3, pp. 392–396, 2006.

[24] A. McGovern, A. Kruger, D. Rosendahl, and K. Droegemeier, "Open problem: dynamic relational models for improved hazardous weather prediction," in *Proceedings of ICML Workshop on Open Problems in Statistical Relational Learning*, vol. 29, 2006.

[25] NOAA, "Earth location users guide (ELUG). rev. 1, NOSS/OSD3-1998-015R1UD0," no. NOAA/NESDIS DRL 504-11, 1998.

[26] J. P. Snyder, *Map projections–A working manual*. USGPO, 1987, no. 1395.

[27] V. Zwatz-Meise and G. Mahringer, "SATMOD: An interactive system combining satellite images and model output parameters," *Weather and Forecasting*, vol. 5, no. 2, pp. 233–246, 1990.

[28] T. N. Carlson, "Airflow through midlatitude cyclones and the comma cloud pattern," *Monthly Weather Review*, vol. 108, no. 10, pp. 1498–1509, 1980.

[29] J. M. Wallace and P. V. Hobbs, *Atmospheric science: an introductory survey*, vol. 92.

[30] C. A. Doswell III, "Weather forecasting by humans-heuristics and decision making," *Weather and Forecasting*, vol. 19, no. 6, pp. 1115–1126, 2004.

[31] B. K. Horn and B. G. Schunck, "Determining optical flow," in *1981 Technical Symposium East*, 1981, pp. 319–331.

[32] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision." in *IJCAI*, vol. 81, 1981, pp. 674–679.

[33] J. S. Pérez, E. Meinhardt-Llopis, and G. Facciolo, "TV-L1 optical flow estimation," *Image Processing On Line*, vol. 2013, pp. 137–150, 2013.

[34] J. Bouguet, "Pyramidal implementation of the affine lucas kanade feature tracker description of the algorithm," *Intel Corporation*, vol. 5, 2001.

[35] T. Acharya and A. K. Ray, *Image processing: principles and applications*, 2005.

[36] R. Temam, *Navier–Stokes Equations*. American Mathematical Soc., 1984.

[37] J. Stam, "Stable fluids," in *Proceedings of the 26th annual conference on Computer graphics and interactive techniques*, 1999, pp. 121–128.

[38] S. I. Green, *Fluid Vortices: Fluid Mechanics and Its Applications*. Springer, 1995, vol. 30.

[39] G. B. Arfken and H. J. Weber, *Mathematical Methods for Physicists*, 4th ed. Academic Press: San Diego, 1995.

[40] J. C. Hunt, A. Wray, and P. Moin, "Eddies, streams, and convergence zones in turbulent flows," in *Studying Turbulence Using Numerical Simulation Databases, 2*, vol. 1, 1988, pp. 193–208.

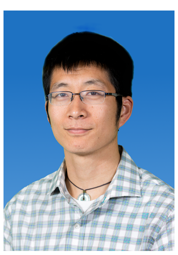
[41] L. Breiman, "Random forests," *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.

[42] F. Molteni, R. Buizza, T. N. Palmer, and T. Petroliagis, "The ECMWF ensemble prediction system: Methodology and validation," *Quarterly Journal of the Royal Meteorological Society*, vol. 122, no. 529, pp. 73–119, 1996.

[43] W. Zheng, H. Wei, Z. Wang, X. Zeng, J. Meng, M. Ek, K. Mitchell, and J. Derber, "Improvement of daytime land surface skin temperature over arid regions in the NCEP GFS model and its impact on satellite data assimilation," *Journal of Geophysical Research: Atmospheres*, vol. 117, no. D6, 2012.

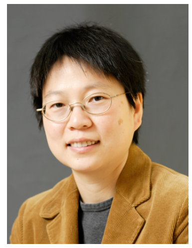
[44] H. Yoo and Z. Li, "Diagnosis and improvement of cloud parameterization schemes in ncep/gfs using multiple satellite products," in *Proceedings of the International Radiation Symposium*, vol. 1531, no. 1, 2013, pp. 268–271.

[45] F. Vitart, "Evolution of ecmwf sub-seasonal forecast skill scores," *Quarterly Journal of the Royal Meteorological Society*, vol. 140, no. 683, pp. 1889–1899, 2014.



Yu Zhang received the bachelor's degree (2006) and the MS degree (2009) from Tsinghua University, and the PhD degree (2014) in Information Sciences and Technology from The Pennsylvania State University. He specializes in computer vision, machine learning, and geo-informatics. He worked at HERE Maps at Chicago as Senior Engineer from 2015 to 2016 on geospatial feature learning. He is currently a Senior Software Engineer at DeepMap Inc. at Palo Alto, California, working on generating high-definition maps for autonomous vehicles.

Stephen Wistar is a Certified Consulting Meteorologist (CCM) and Senior Forensic Meteorologist. He has worked on numerous cases involving Hurricane Katrina, building collapses, flooding and slip and falls. His work involves explaining meteorology to the non-scientist. At AccuWeather, Steve constructs impartial scientific weather evidence for use in courts of law to support the prosecution, which has aided Steve in testifying over 125 times in courtrooms or depositions. Furthermore, he coordinates numerous past weather research projects using forensic meteorology, like the 250 reports Steve wrote on the impacts of Hurricane Katrina at specific locations on Gulf Coast.



Jia Li is a Professor of Statistics at The Pennsylvania State University. Her research interests include statistical learning, data mining, image processing, retrieval/annotation, and composition analysis. She also worked as a program director at the National Science Foundation from 2011 to 2013, a visiting scientist at Google Labs in Pittsburgh from 2007 to 2008, a researcher at the Xerox Palo Alto Research Center from 1999 to 2000, and a research associate in the Computer Science Department at Stanford University in 1999. She received the M.Sc. degree in Electrical Engineering (1995), the M.Sc. degree in Statistics (1998), and the Ph.D. degree in Electrical Engineering (1999), all from Stanford University.

Michael A. Steinberg received the BS in atmospheric sciences from Cornell University and the MS degree in meteorology from The Pennsylvania State University. He is an Expert Senior Meteorologist, Senior Vice President and Emeritus member of the Board of Directors of AccuWeather, Inc., where he has been employed since 1978. In this role he interacts in a wide variety of scientific, tactical and strategic areas. He is a Fellow of the American Meteorological Society (AMS) and was a recipient of their 2016 Award for Outstanding Contribution to the Advance of Applied Meteorology for numerous, visionary innovations and accomplishments in meeting public and industrial needs for weather information. He has been a recipient of research grants from the NASA Small Business Innovation Research program and the Ben Franklin/Advanced Technology Center of Pennsylvania, and is the inventor or co-inventor for numerous patents related to weather indices and location-based services.

James Z. Wang received the bachelor's degree in mathematics and computer science *summa cum laude* from the University of Minnesota, the MS degree in mathematics and the MS degree in computer science, both from Stanford University, and the PhD degree from Stanford University. He has been a faculty member at The Pennsylvania State University since 2000 where he is a Professor in the College of Information Sciences and Technology. His main research interests are image modeling and retrieval, with applications in weather forecasting, aesthetics, emotions, art, and biology. He was a visiting professor at the Robotics Institute at Carnegie Mellon University (2007-2008) and a program manager at the National Science Foundation (2011-2012). He has been a recipient of a US National Science Foundation Career award and the endowed PNC Technologies Career Development Professorship.