

Cyclone Track Forecasting Based on Satellite Images Using Artificial Neural Networks

Rita Kovordányi*
*Corresponding author
Linköping University
SE-581 83 Linköping, Sweden
Tel: +46 13 281430
Fax: +46 13 142231
E-mail: ritko@ida.liu.se

Chandan Roy
Linköping University
SE-581 83 Linköping, Sweden
E-mail: charo@ida.liu.se

Many places around the world are exposed to tropical cyclones and associated storm surges. In spite of massive efforts, a great number of people die each year as a result of cyclone attacks. To mitigate the damages caused by cyclones, improved cyclone forecasting techniques must be developed. The technique presented here uses artificial neural networks to interpret NOAA-AVHRR satellite images. A multi-layered neural network, resembling the human visual system, was trained to forecast the movement direction of cyclones based on satellite images. The trained network produced correct directional forecast for 98 % of novel test images, thus showing a good generalization capability from training images to novel images. The results indicate that multi-layered neural networks could be further developed into an effective tool for cyclone track forecasting using various types of remote sensing data. Future work includes extension of the present network to handle a wide range of cyclones and to take into account supplementary information, such as wind speeds around the cyclone, water temperature, air moisture, and air pressure.

Keywords: Cyclone Track Forecasting, Artificial Neural Networks, Multi-Layer Networks, Receptive Field, Remote Sensing, Geohazards

1 Introduction

A tropical cyclone is an area of low pressure which develops over tropical or subtropical waters. These systems form over all tropical oceans with the exception of the South Atlantic and the eastern South Pacific, east of about 140° W longitude. There are seven tropical cyclone basins, where tropical cyclones form on a regular basis:

1. Atlantic basin
2. Northeast Pacific basin
3. North Indian basin
4. Southwest Indian basin
5. Southeast Indian/ Australian basin
6. Australian/ Southwest Pacific basin, and
7. Northwest Pacific basin

In their most intense state these storms are called hurricanes in the Atlantic, typhoons in the western North Pacific and cyclones in the Bay of Bengal. These low-pressure systems draw their energy from the very warm sea-surface waters. As the warm, moist air spirals counterclockwise

in toward the centre, the wind speeds increase, reaching their maximum values in the region surrounding the almost calm centre of the cyclone.

As most cyclones are formed in the tropical seas and as the density of population is greatest in the tropical regions, cyclones constitute a major hazard for a large number of people around the world (Cervený and Newman, 2000). Some of the more known examples include the Bangladesh cyclone of November 12, 1970 and May 24, 1985 (both crossed the coast at Chittagong), hurricane Camille that hit USA on August 17, 1969 and cyclone Tracy that swept over the Australian coast on December 25, 1974; and recently hurricane Katrina; these cyclones have each caused a large number of fatality and immense damage in property. Frequent attacks of less intensive cyclones (having less wind velocity and lower surge levels) continue to cause human casualties and considerable economic damage in tropical countries. For these reasons substantial resources have been devoted around the world to research, forecasting and socioeconomic preparedness for cyclones and cyclone generated disasters. The present paper focuses on automating the forecasting of cyclone track and thereby providing a more reliable basis for early warning systems.

1.1 Existing techniques for cyclone forecasting

Tropical Cyclone (TC) forecasting involves the prediction of several interrelated features, including the track of the cyclone, its intensity, resulting storm surge and rainfall and, of course, which areas are threatened. Among these features, forecasting the future track and intensity of tropical cyclones are considered to be the most important because the casualties and loss of property depend critically on these two features. Some of the more widely used cyclone track forecasting techniques are summarized in Table 1.

Table 1. A sample of techniques used in various meteorological offices to forecast cyclone tracks (adapted from McBride and Holland, 1987).

Office	Techniques					
	Subjective	Analogue	Steering	Statistical	Dynamical	Empirical
Japan	X			X	X	
Hong Kong	X		X	X		
Philippines	X			X	X	
Miami	X	X		X	X	
India	X	X		X		
Brisbane	X	X	X	X		
Fiji	X					
Reunion	X					
Mauritius	X					
Mozambique	X					
Madagascar	X					
Darwin	X	X	X	X		
Guam	X	X	X	X	X	X

Subjective assessment: This technique includes synoptic reasoning, evaluation of expected changes in the large-scale surrounding flow fields and subjective evaluation of the cyclone's steering current.

Analogue forecasts: Designated features of the cyclone, such as its latitude, longitude, intensity, maturity, and past motion, are compared to those of all previous cyclones in the same region to select one or more analogues. The cyclone movement is then derived from the previous development of tracks of the analogues.

Steering current: The cyclone's steering current is determined using analysis of winds at specified points and altitudes around the cyclone. The actual forecast can be based on simple regression analysis, or on analysis of the advection and propagation of winds, incorporating linear interactions between the vortex and the background absolute vorticity (Holland, 1984).

Statistical technique: All the statistical forecasting techniques are based on regression analysis. Here historical patterns of previous storms, such as cloud patterns, intensity development, and previous actual track are adapted and applied to the present storm.

Dynamical: These techniques are based on numerical integration of mathematical equations that approximate the physical behavior of the atmosphere. The technique looks different when it is applied on a regional or global scale.

Empirical: A skilled meteorologist has often developed an ability to detect overall patterns in climatological conditions and can assess how these may affect cyclone development. Manual forecasts made by a skilled meteorologist may therefore be a good complement to other forecasting techniques. The success of this technique critically depends on the experience of the forecaster.

The National Hurricane Center (NHC) of USA and Australia Bureau of Meteorology (2007) identify the following additional techniques for cyclone movement forecasting:

Persistence: This technique is useful for short-term forecasts, and assumes that the cyclone will maintain its recent track. This technique is often used in combination with other techniques.

Satellite-based techniques: In this technique track and intensity are forecasted based on the cloud pattern associated with the cyclone. Generally the outer cloud bands of cumulonimbus clouds indicate the future direction, and the cloud pattern surrounding the cyclone eye indicates the future intensity of the cyclone.

Hybrid: Meteorological centers round the world often employ a combination of techniques to get a more accurate track forecast, for example, combining elements of two or more of the above techniques. The elements are blended as a weighted sum, where the weights are based on past performance of each forecasting technique (Naval Research Laboratory, 1999).

1.1.1 Automated forecasting

Like other forecasts, tropical cyclone forecasts are not free from error. Error in registering the initial position and motion of the tropical cyclone can have an impact on the accuracy of subsequent forecasts. Errors can also arise from a lack of full understanding of the mechanisms behind the formation and growth of tropical cyclones and from the limitations of the forecasting techniques themselves. Mean track forecast error is typically smaller for lower latitude cyclones moving westward than for higher latitude cyclones in westerly winds and for those cyclones which are re-curving. In general, mean track forecast errors tend to increase with the forecast period and can be as much as 30 % of the cyclone movement within this same period. The forecasted track can deviate from the cyclone's actual track by as much as 20 degrees.

Due to the high complexity of the interaction between factors affecting cyclone development, meteorological offices around the world try to automate much of the work involved in cyclone forecasting. The Automated Tropical Cyclone Forecasting System (ATCF), developed by the Naval Research Laboratory (NRL) in Monterey, California, is an example of an automated forecasting system (Naval Research Laboratory, 1999). This computer-based application is intended to automate and optimize much of the tropical cyclone forecasting process. It provides a means for tracking, and intensity forecasting, as well as constructing messages, and disseminating warnings.

1.1.2 Satellite-based techniques

During the 1980s, forecasting of the track and intensity of tropical cyclones was mainly based on statistical (regression) methods using general meteorological data. Later, in the early 1990s,

remote sensing techniques were starting to be used successfully for cyclone forecasting (Marshall et al., 2002; Wells, 1987). Development of new techniques, such as the generation of high resolution atmospheric motion vectors (AMVs) from satellite images, and four dimensional variational assimilation (4D-VAR) have reduced the error in forecasting a cyclone's track and intensity (Marshall, et al., 2002).

Images of different channels obtained from weather satellites have their specialized use in track and intensity forecasting of tropical cyclones. Satellite images within the thermal infrared (IR) band can be used to forecast and analyze the cyclone's intensity (Kossin, 2003). In addition, Lau and Crane (1997) and Kishtawal et al. (2005) measured the intensity of tropical cyclones based on satellite images and data from a Thermal Microwave Imager using non-linear data fitting. Present techniques can be developed further, as data from Advanced Microwave Sounder Units having better horizontal resolution and vertical temperature sounding abilities provide an improved basis for temperature estimation compared to conventional Microwave Sounding Units and IR satellite images (Knaff et al., 2000).

Satellite-based techniques can be used for forecasting both cyclone intensity and cyclone track. TC development can be analyzed by studying the cloud patterns and determining how they change with time. Repeated observations of a TC provide information on the intensity and the rate of growth or decay of the storm. This method of intensity analysis is based on the degree of spiraling in the cloud bands. The more spiral the cloud pattern is the more intense is the cyclone (Dvorak, 1975). The Dvorak model of forecasting intensity is successful in most cases. However, the technique is based on subjective judgment and is unreliable in the sense that various weather stations around the world can arrive at discrepant results for the same cyclone. Velden et al. (1998) uses a computer based algorithm named Objective Dvorak Technique (ODT) to address this problem.

2 A new method for automated cyclone track forecasting

One technique that has not been used for cyclone track forecasting previously is artificial neural networks (ANN). This in spite the fact that ANN-techniques have been used in other remote sensing application areas, such as road network detection (Barsi and Heipke, 2003a), cloud detection (Jang et al., 2006), cloud motion detection (Brad and Letia, 2002a, 2002b), and precipitation forecasting (Hong et al., 2004; Rivolta, et al., 2006) based on aerial photographs and satellite images. One conceivable reason why ANN-techniques have not been applied to cyclone forecasting before is that it is difficult to achieve robust network performance in this complex domain.

Previous research suggests that cloud patterns surrounding the cyclone is a good indicator of the direction of cyclone movement. There is a tendency for TCs to move towards the downstream end of convective cloud bands in the outer circulation strips around the cyclone (Lajoie, 1976). Changes in the orientation of such cloud bands indicate that a similar change in cyclone movement direction may occur in the next 12-24 hours. Further, TCs do not continue towards, nor curve towards cumulonimbus free sectors in the outer circulation. Also Fett and Brand (1975) noted that rotation of gross cloud features (such as an elliptical cloud mass or a major outer band) provide a very good indication of cyclone direction changes during the next 24 hours.

The above results indicate that a cyclone's track is indicated by the shape, and relative position of surrounding cumulonimbus clouds. What is more, these features are visible in satellite images (Fig. 1). Detecting and categorizing these features using a neural network would therefore provide valuable input to automated cyclone forecasting. In this article we demonstrate that neural networks can learn to detect the surrounding clouds that are present in the satellite image and learn to recognize how these clouds are positioned relative to the cyclone center.

Thick outer cumulonimbus cloud bands in NOAA - AVHRR image

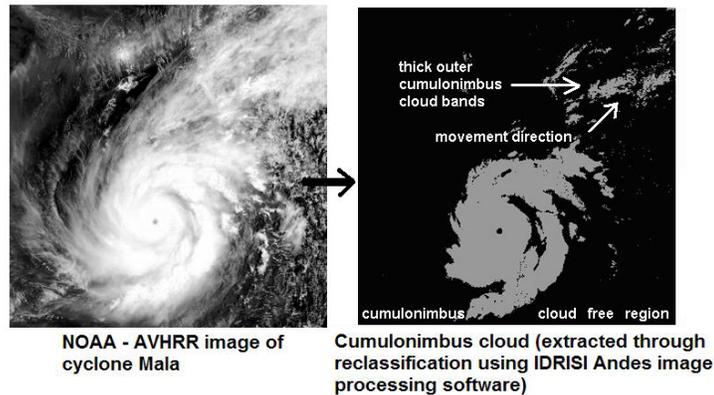


Fig. 1. Example of a satellite image showing a typical cloud pattern surrounding a cyclone. The dense cumulonimbus clouds, which are visible in the top right corner of the image, indicate a strong upstream of hot, moist air. This upstream creates a low pressure sector that fuels the cyclone and sucks it in a northeastern direction.

2.1 Robust pattern recognition inspired by human vision

Previous applications of neural networks in the area of remote sensing, for example, detection of the presence of clouds, and estimation of air temperature based on satellite images deploy standard three layer feed forward networks (e.g., Brad and Letia, 2002a, 2002b; Barsi and Heipke, 2003b; Hong et al., 2004; Jang et al., 2006; Rivolta et al., 2006).

However, the task for the neural network in the present study is to recognize the elongated shape present in the image, irrespective of random variations in size and position of this shape, and it is difficult to achieve position and size invariance in standard three-layer networks.

2.1.1 Biologically inspired image processing techniques

Research in computer vision suggests that position and size invariance can be achieved by using deep multi-layer neural network architectures. These multi-layer architectures are to varying extent inspired by human vision. Somewhat simplified, the human visual system is made up of a series of transformational steps, or processing levels. Each transformational step is reciprocally connected to both the predecessor and successor step (Ungerleider and Mishkin, 1982; Felleman and Van Essen, 1991). This recurrent (i.e., bi-directional) communication, together with a cascade-like spreading of activation across the system (McClelland, 1993; O'Reilly, 1998), allows the various processing levels to work in concert and arrive at a consistent interpretation of visual input.

Individual feature detectors in the human visual system have limited receptive fields, that is, they are sensitive to a given feature within a limited region of the image. Interpretation of the complete visual image is achieved through parallel feature extraction across the image and step-wise hierarchical combination of these features into a coherent pattern.

There are two sorts of computations going on during image processing: 1. Recognition of more and more complex features. 2. A stepwise increase of invariance. For this reason, biologically inspired networks for image processing often comprise pairs of layers containing simple cells and complex cells, respectively (Fukushima and Miyake, 1982; Spratling, 2005; Mutch and Lowe, 2008). Simple cell layers operate on a limited part of the image and combine

features from the previous processing step into more complex shapes. In contrast, complex-cell layers achieve increased invariance by reacting uniformly to shapes that can be present anywhere within an increasingly larger receptive field.

As a rule, the first processing step in these artificial neural networks is tuned for the detection of line segments. These detectors are not developed during learning, but are instead pre-programmed, for example, using Gabor-filters (Mutch and Lowe, 2008). In the present project, filled cloud segments turned out to be more appropriate as basic features. Detectors for these features were allowed to develop from the images through learning.

Existing mixed-cell multilayer networks have been demonstrated to achieve robust shape recognition, and are currently state of the art. The question is, however, if shape and position invariance could be achieved using a simpler architecture, involving only one type of cells, so that feature detection and spatial integration occurs at each processing level.

When constructing our network, we used this simplified approach, with a down-scaled network architecture comprising five reciprocally connected layers (Fig. 2). We used one layer (and one type of cells) at each processing step, instead of the commonly used intertwining of single-cell and complex-cell layers.

2.1.2 Combined unsupervised and supervised learning

Existing networks in the area of remote sensing have without exception been trained using the backpropagation of error learning algorithm (Barsi and Heipke, 2003b; Hong et al., 2004; Jang et al., 2006; Rivolta, et al., 2006). As opposed to this, state-of-the-art networks in computer vision and image processing often use unsupervised learning for the training of feature detectors.

Biologically-based—as opposed to biologically inspired—networks for image processing combine unsupervised and supervised learning, as also the human brain seems to use a combination of learning methods (O’Reilly, 1998). To process perceptual information, humans use unsupervised model-based learning, which takes into account the statistical regularities that occur across various experiences and trains the network to detect these regularities. In addition, in order to learn how to use regularities in the environment so that an appropriate response is elicited each time, humans also use a form of error-driven task learning, where the output (action) that is produced is compared to and slowly adapted towards the desired output.

We employ a combination of model-based learning and error-driven task learning for each processing step. While unsupervised, model-based learning is well suited for extracting cloud segments and more complex cloud patterns from the satellite images, error-driven task learning is necessary for training the network to associate these patterns with a correct forecast of cyclone track direction. One way of combining the two learning algorithms would be to let the first few layers in the network perform model-based learning, while subsequent layers would use error-driven task learning. However, if weights in the first few layers were changed independently of the task requirements, these weights could develop to work against the required input-output mapping, for example, learning to detect features that are not relevant for the task at hand.

An efficient way to combine the two learning algorithms is to let both take place simultaneously at each layer, and to combine the outcome of the two algorithms for each learning cycle. There are both practical (faster learning) and theoretical advantages of combining the two learning algorithms in this way (cf. O’Reilly and Munakata, 2000).

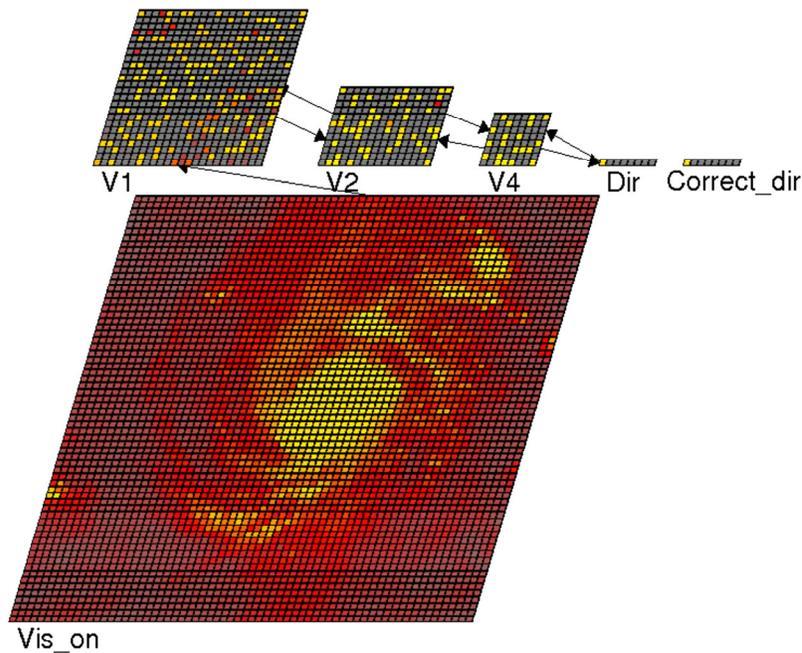


Fig. 2. The five-layer network used in this study, with bi-directional projections connecting subgroups of sending and receiving nodes. Hence, the network is constructed to detect constituent features in limited parts of the satellite image, and then integrate these features into an overall shape. Note that the layer called ‘Correct_dir’ did not partake in computation or error calculation; it was used as an aid, to display the correct answer for each image, enabling us to visually compare predicted versus correct direction of cyclone movement. Calculation of training error took place in the Dir-layer, independent of activation in Correct_dir.

2.2 Biological transformational steps

At the first transformational step in the human visual system (Fig. 3), the image captured by the retina is routed via the Lateral Geniculate Nucleus (LGN) to primary visual cortex (brain area V1). More exactly, the role of LGN is to bundle input from the two eyes and to mediate these signals to primary visual cortex V1.

Brain area V1 contains specialized neurons that are sensitive to specific features, such as variously oriented line segments.¹ V1 neurons are dedicated to the processing of input from a limited part of the retinal image. This part of the retinal image corresponds to a specific region in external space, which comprises the neuron’s receptive field (RF).

Feature detector neurons with the same RF cooperate in detecting various features at the same part of space. For example, neurons that cover the same part of space may be sensitive to line segments with various orientation; these neurons are organized into one hyper column.

¹ Visual features constitute the basic building block of human vision, and reflect simple visual properties that are present in limited parts of the image. For example, recognition of the letter ‘A’ builds on detection of the constituent line segments ‘/’, ‘-’, and ‘\’. These features are located in different parts of the input (the left, right, and middle part of the input). It is the combined presence of these features and their relative location that signals the presence of the letter ‘A’.

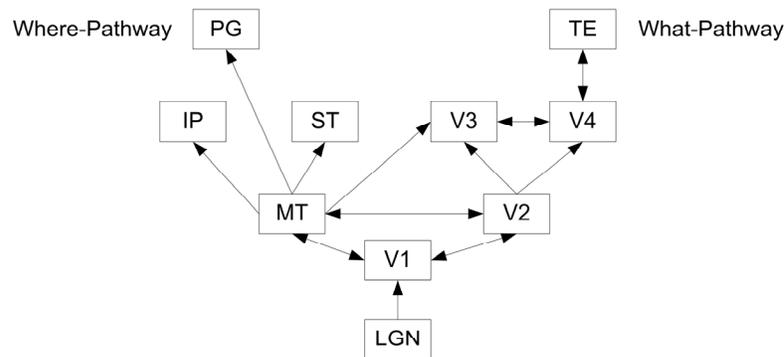


Fig. 3. Architecture of the human visual system. The what- and where-pathways depicted in the figure are physically separated in the brain. The where-pathway runs in the dorsal (upper) part of the brain, and is specialized for spatial processing of size, location and orientation of objects. The what-pathway is located in the ventral (lower) part of the brain, and implements color and shape detection (figure adapted from Felleman and Van Essen, 1991).

Neighboring hyper columns of feature detectors cover neighboring parts of space. Together, the matrix of hyper columns can process the entire input image. Logically, the simultaneous activation of a specific combination of neurons within these groups would reveal the presence of various features at specific parts of the image.

However, image interpretation is not accomplished in one step in the brain, as it would be extremely difficult to reliably recognize a complex pattern of features, when feature detection itself is underdetermined and has to be adjusted to variations in appearance (size, location, and orientation) of these features. There are simply too many degrees of freedom and these cannot be handled all at once. Instead, the brain implements a gradual transformation, cutting down the space of possible interpretations of the input image step by step. Hence, neighboring simple features represented in V1 are merged into slightly more complex (aggregate) features at the next step of processing that is, in brain area V2. More specifically, detectors for these aggregate features react to the combined activation of simple feature detectors in V1. Also V2 neurons receive input from a limited part of the retinal image, but these RFs now cover a slightly larger part of space.

Through further step by step transformations (in areas V3, V4 and TE), features are gradually aggregated into more and more complex object properties. For each transformational step, the aggregated features will be located in larger and larger parts of the image. At each transformational step, feature detection needs only to allow for variations in location within a limited part of space (i.e. within the RF of the feature detector). Processing of the image is simplified as variations in size and orientation apply to simpler constituent features (as opposed to complex overall patterns). In addition, variations in size are limited by the boundaries of the RF. Hence, step-by-step processing in a feature detector hierarchy makes handling of the many degrees of freedom inherent to image processing more tractable.

2.3 Architecture of our artificial neural network

Robust pattern recognition can thus be achieved through stepwise processing in a bi-directionally connected multi-layer network. At the first processing level, shape recognition is based on detection and integration of characteristic *constituent features*, such as small cloud segments. The

constituent features are smaller in size than the aggregate shape (the whole cloud pattern), and therefore can be detected by looking at limited parts of the image.

The stepwise integration of features that we have implemented in our artificial neural network resembles the part of the human visual system which is responsible for pattern detection and shape recognition, what is commonly called the what-pathway (Goodale and Milner, 1992; Kosslyn, 1994; Creem and Proffitt, 2001), comprising brain areas V1 to V4 (cf. Fig. 3).

By processing the image through a matrix of feature detector units, constituent features in different parts of the image can be extracted in parallel (O'Reilly and Munakata, 2000). For each feature that is detected, the network needs to abstract away from minor variations in position and size, but only within a small part of the image. Due to the task requirements, rotational variations are encoded as meaningful features at each stage of processing.

At subsequent stages of processing, the simpler constituent features are integrated into larger patterns that is, into more complex features. Recognition of these secondary features is based on the combined activation of feature detectors at the previous stage, and is thus not strictly dependent on which part of the image these activations has arisen from. For each additional transformational step, the more independency is achieved and the more complex patterns can be recognized.

At the final stages of processing, the constituent features can be integrated into an overall shape which is perceived to be rotated in a specific way. This information is then mapped onto a directional indicator showing the predicted movement of the cyclone.

2.3.1 Enhanced generalization ability

Through consecutive transformational steps, detection of shape and orientation becomes increasingly independent of the exact position and size of the features, and is instead based on how the constituent features are located and oriented *relative to each other*.

Image processing is based on simple features which are present also in novel satellite images. Novel satellite images that have not been presented to the network, would still elicit a correct response based on the detection of previously experienced and well learnt constituent features. Hence, step-by-step image processing enhances generalization capability to novel images.

3 Implementation of the network

The network presented in this paper was implemented and trained using the artificial neural network simulation tool Leabra++, which is specialized for modeling biologically-based (brain-like) artificial neural networks (O'Reilly and Munakata, 2000).

The network consisted of an input layer, three consecutive hidden layers, and an output layer (Fig. 2). Layers were bi-directionally connected, meaning that if a layer was sending input to another layer, it also received feedback from that layer. Connections were symmetrical, so that if unit x was sending input to unit y , then x also received feedback from y . This symmetry was preserved during training.

The first receiving layer V1 was designed to encode simple visual features in the input image, such as small pieces of cloud patterns. To be able to extract features from part of the input, nodes in this layer were grouped and connected to the input in a way so that each group would receive input only from a limited part of the image. Grouping of nodes and selective communication between groups of nodes across layers simulated the concept of limited receptive fields (RF) that can be found in the human visual system (Fig. 4). Likewise, each group in the sending layer received feedback signals from a small number of neighboring groups in the receiving layer. Neighboring receptive fields on the input side slightly overlapped, in order to maintain continuity of processing across various parts of the image. V1 consisted of 9 groups of 8 x 8 units. Each of these groups received input from a 26 x 26-pixel patch of the input image (overlaps disregarded).

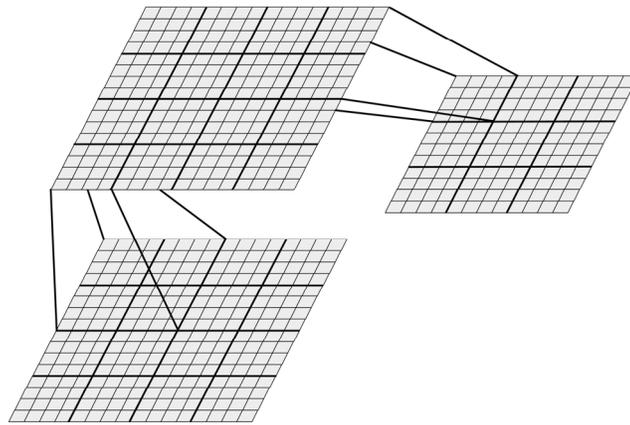


Fig. 4. Selective communication between corresponding groups of nodes across layers, resembling the limited receptive fields in the human visual system. For each transformational step, increasingly complex features can be extracted, and these features originate from increasingly larger parts of space.

Using this network architecture, local information from different parts of the input image could be processed in parallel. For each transformational step in the network, information was integrated, taking into account larger and larger parts of the input image. Hence, layer V2 consisted of 4 groups of 8 x 8 units, where each group received input from 4 groups in layer V1. Finally, layer V4 comprised of 8 x 8 = 64 units, where each unit received input from all units in V4. All feedback connections were symmetrical on a unit-to-unit level.

At the final transformational step (V4 → Dir), the shape and orientation of the overall cloud pattern was transformed into a directional prediction, indicating in which direction the cyclone was most likely to move. Output from the network was represented by an 8 x 1 vector, where each bit (i.e. activation of one individual node) signified a possible rotational direction ranging from 0° to 315° in 45° steps. The low resolution of output reflected the relatively low resolution of the input images (66 x 66).

3.1 Training data

Training data was prepared by downsizing the original satellite images to 66 x 66 pixels and then converting the resulting images into grayscale and linearly transforming this image from a format where each pixel contained a value between 0 and 255 into a representation where pixel values were limited to the interval [0 1]. This transformation would thus map an original pixel value of 255 to a new pixel value of 1. Similarly, an original pixel value of 127 would be transformed into a new pixel value of 127/255, and so on. Note that this transformation preserved the information contained in the original image.

The resulting image was then artificially rotated clockwise in 45° steps, which resulted in 8 rotated images.

In order to create variation in the cyclone's size relative to the input image frame, each of the rotated images was zoomed in three ways (using zoom factors 0.8, 1, and 1.2).

We were forced to use relatively low-resolution images (Fig. 5) due to computational resource limitations. We used commercially available standard PCs for our network simulations. We reasoned that if the network would learn to forecast cyclone tracks based on these low-

resolution images, this would be an excellent proof of concept, demonstrating the usefulness and robustness of ANN-techniques for satellite image interpretation.

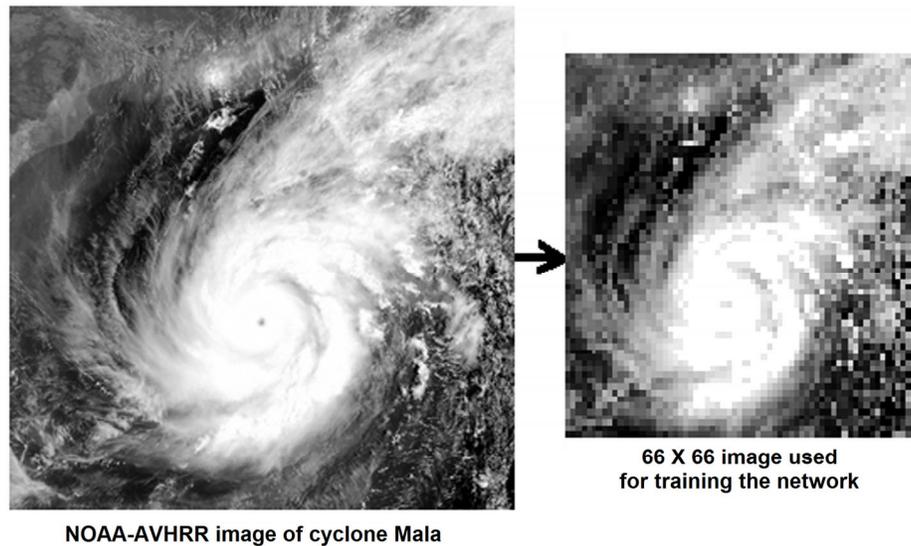


Fig. 5. Example of original image and corresponding downsized image that was used to generate input patterns to the network. The downsized image was artificially rotated clockwise in 45° steps. Each of the rotated images was zoomed and shifted in position, in order to create variation in the cyclone's orientation, size, and location in the image.

Finally, each image was shifted in position zero, one, or two steps. The size of the steps was varied across training sessions, so that we could compare network performance for step-size one pixel with step-size three pixels. Position shifts were generated in eight directions (north, south, east, west, northeast, et cetera). The resulting position of the cyclone could differ by as much as 12 pixels (when using 3-pixel steps for shifting), which amounted to almost one fifth of the total image size. As a result, important information about the cyclone's surroundings could disappear off the edge for some images. To avoid moving the most informative part of the cyclone's surroundings outside the image frame, certain combinations of image rotation in a particular direction and two-step position shifts in the same direction were automatically discarded. All in all, the transformations resulted in 1200 images.

From the total set of images, a random subset comprising approximately 5 % of the images was set aside for later testing of network performance. The remaining 95 % of the images were used for training the network.

Correct movement direction was determined by visual estimation for the original satellite image and was automatically calculated for the transformed images. Visual estimation of the relative position of cumulonimbus clouds in the surroundings the cyclone was based on the rule of thumb that the cyclone is most likely to move in a direction where there are cumulonimbus clouds in the outer circulation bands around the cyclone, as these signify a local area of low pressure (cf. Fig. 1). The network's task was to detect the same pattern that the human eye did, and to use this information to predict the cyclone's future movement direction. In this way, the artificial neural network was required to play a role similar to a human expert at a meteorological department.

3.2 Information processing in the network

For each input-output pair in the training set, the network went through two activation phases. During the *minus phase* the network was presented with the input and then was allowed to settle in an activation state consistent with both the input and with the weights in the network. During settling, weakly activated nodes were suppressed, allowing at most k nodes to be active within a group of nodes in the same group of units and/or the same layer. This mechanism is commonly referred to as k -winners-take-all (kWTA).

In the second *plus phase*, activation in the network was reset to zero and the network was presented with both the input and the correct output. During this second phase, the network settled into an activation state that reflected the correct (“should-have-been”) activation of each node, given the correct output. By comparing the activations between the minus and the plus phases, an error was computed for each node.

The incoming weights for each node were adjusted in a direction that decreased this error. In addition, a weight change was also computed based on co-activation of pair of nodes: Nodes that were activated together during the plus phase were assigned a small positive weight change to enforce this tendency for co-activation. Likewise, a negative weight change was assigned to connections going from an inactive to an active node, and vice versa, to decrease the dissonance between these nodes. The final weight update for each connection reflected the sum of the error-driven and correlation-based weight change (see Equation 4).

Training of the network can be described in mathematical terms as follows. Weights in the network were limited to the interval $[0\ 1]$. Also the activation of nodes was limited to the interval $[0\ 1]$. The activation function (describing how strongly a node should react to input) was given by the following biologically-based, sigmoid-like function (O’Reilly & Munakata, 2000):

$$y_j = \frac{\gamma[V_m - \Theta]_+}{\gamma[V_m - \Theta]_+ + 1}, \quad [z]_+ = \begin{cases} z & \text{if } z \geq 0 \\ 0 & \text{if } z < 0 \end{cases} \quad \text{Equation 1}$$

y_j = activation of receiving node j
 γ = gain
 V_m = membrane potential
 Θ = firing threshold

Learning in the network was based on a combination of Conditional Principal Component Analysis (CPCA), which is a form of unsupervised model-based learning algorithm (Hebbian learning) and Contrastive Hebbian learning (CHL), which is a biologically-based error-driven algorithm, an alternative to backpropagation of error (O’Reilly, 1998; O’Reilly & Munakata, 2000):

$$\text{CPCA: } \Delta w_{ij} = \epsilon y_j (x_i - w_{ij}) = \Delta_{hebb} \quad \text{Equation 2}$$

x_i = activation of sending node i
 y_j = activation of receiving node j
 w_{ij} = weight from node i to node j

$$\text{CHL: } \Delta w_{ij} = \epsilon (x_i^+ y_j^+ - x_i^- y_j^-) = \Delta_{err} \quad \text{Equation 3}$$

x_i = activation of sending node i
 y_j = activation of receiving node j
 x^+, y^+ = activations when both input and correct output is presented to the network
 x^-, y^- = activations when only input is presented to the network

Combined learning: $\Delta w_{ij} = \epsilon [c_{hebb} \Delta_{hebb} + (1 - c_{hebb}) \Delta_{err}]$ **Equation 4**
 ϵ = learning rate (initially set to 0.01, and then gradually decreased during training)
 c_{hebb} = proportion of Hebbian learning (set to 0.01 in the present study)

4 Results

Training was run in Epochs. Each epoch consisted of one round of presentation of the complete training set. The training set consisted of approximately 95 % of all images. A randomly chosen set of 5 % of the total images was saved for testing, and was therefore not presented to the network during training. Between epochs, the images were permuted, to avoid sequence learning effects.

After each epoch, that is, after the complete training set was presented to the network (each time in permuted order), a sum of summed squared error (SSE) was calculated. The average error was recorded and plotted for each epoch (i.e. after each presentation of the complete training set).

Error in network performance was defined as the difference between correct output and the output produced by the network. The difference was calculated for each output unit, and then squared and summed into an overall error measure. The output layer in the network consisted of 8 nodes, where each node represented a possible movement direction, ranging from 0° (north) to 315° (northwest) in 45° steps. Performance error was measured in each of these eight nodes, as the network was required to activate one of these nodes, but not the others. The individual errors were squared (to eliminate differences in sign) and summed for the eight units, producing a summed squared error (SSE) for each output produced by the network. Performance error for each epoch was measured as the SSE summed over all input-output pairs in the training set. We conducted multiple trainings to make sure that network training was stable, that is, independent of the initial random weights. In general for all training runs, training error (SSE) decreased

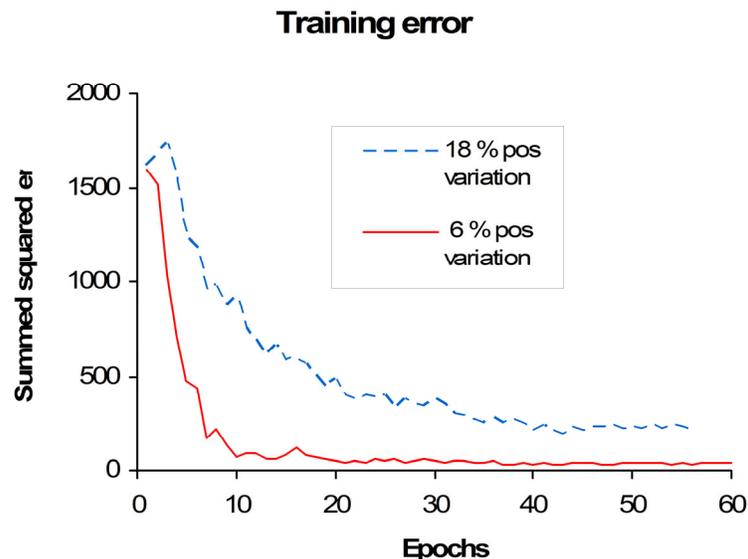


Fig. 6. Sample training curves. Typically, the network reached a training error of 30-40 after 10-15 epochs of training, when cyclone position was varied by 6 % of the image size. The dashed line shows a sample training curve for input images where the cyclone center varied in position by up to 18 % of the image size.

rapidly and leveled out at around 20, after 10-15 epochs of training (Fig. 6). This meant an average error of 0.02, divided among 8 nodes, for each output produced by the network.

These results were attained for images where the position of the cyclone center was varied by a maximum of 6 % of the total image size. This means that when using the network for cyclone track prediction, the cyclone needs not be relatively precisely positioned in the center of the image in order for the network to produce valid track predictions. To get a feeling for how well the network handled variations in the position of cyclone center, we experimented with training the network using larger variations in the position of cyclone center. A typical training curve for such a training set is depicted with dashed line in Fig. 6.

4.1 Network performance for novel images

About 5 % of the images were excluded from the training set and were used as test images. These test images were presented to the network for the first time during testing. During testing, we recorded which images were and were not interpreted correctly. We obtained about 99 % correct direction prediction for the novel cyclone images when cyclone center varied by 6 % of the

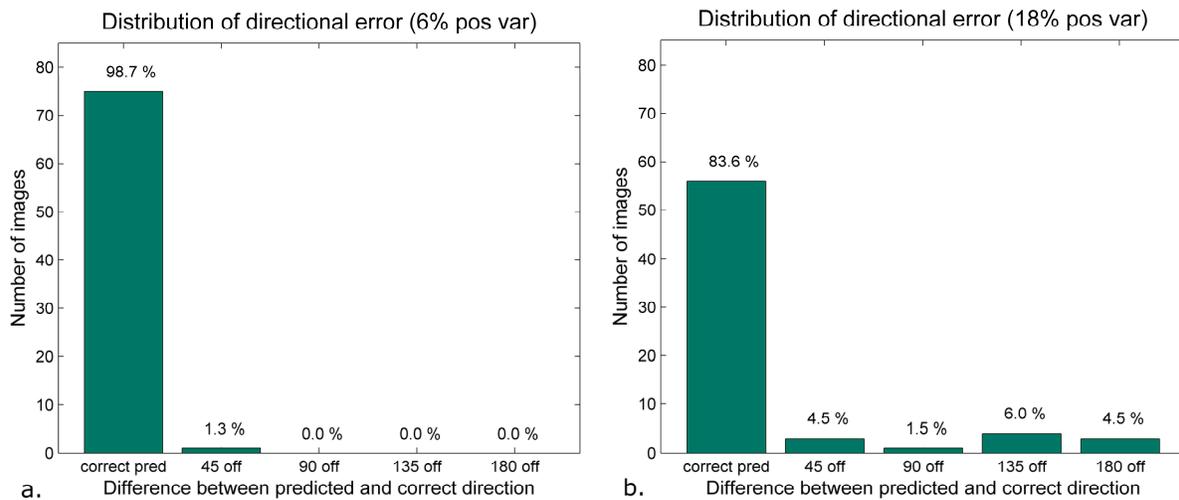


Fig. 7. Sample distribution of errors for the test images. **a.** Directional errors when cyclone center varied by 6 % of the total image size. The leftmost bar indicates correct cyclone track predictions for novel images (98.7 %). **b.** Distribution of errors when cyclone position varied by 18 %. A stable pattern across multiple runs was the relatively few orthogonally incorrect predictions (90° off).

image size (Fig. 7 a). When the position of the cyclone center varied more within the image frame (18 % of the image size), correct track predictions dropped to 84 % for novel images (Fig. 7 b).

For the image set that contained images with up to 18 % positional variance of the cyclone center relative to the image frame, several training and testing sessions were conducted using different starting weights, and we detected a reoccurring pattern. In about 5% of the cases that is, for 3-4 of the 70-75 testing images, the network mistakenly predicted that the cyclone was moving opposite to the correct direction. In other words, the prediction was 180° off (Fig. 7 b). For a couple of images, prediction was 45° off. Finally, for 1-2 images, predicted direction was perpendicular to the correct direction. Hence, mistaken predictions tended to be either close to the correct direction (45° off), or in the opposite direction (or directions close to this). In contrast,

incorrect predictions that were oriented perpendicular to the correct direction were comparatively rare.

5 Analysis of network performance

The pattern of errors produced by the network (Fig. 7) indicated that predictions were based on shape information, namely the elongated shape formed by the cyclone together with the cumulonimbus clouds on one side of the cyclone. The images fed into the network were relatively low-resolution (66 x 66), and it seems that for some of these images, the network mistakenly recognized the elongated shape as being oppositely oriented (e.g., upside-down, or mirrored left-to-right,) resulting in predictions that were approximately 180° off. In contrast, the network made relatively few errors where predictions were 90° off.

The internal workings of artificial neural networks are often considered to be impenetrable for the human analyst. An interesting question is therefore how the network performs cloud pattern recognition. For this, we need to understand which low-level features are extracted by the network when a satellite image is presented. What sort of low-level features the network has learnt to detect is in turn reflected by the weight structure that has developed during training between the first input layer (Vis_on) and the next receiving layer (V1).

The receiving (incoming) weights into individual receiving units in V1 can be plotted to see which pixels in the input image that this particular receiving unit has learnt to focus on. An example is shown in Figure 8.a., where the receiving weights have been visualized for one particular unit in V1. As can be seen in the figure, these weights indicate that the receiving unit has specialized in detecting a shape that resembles a cloud segment.

One way of visualizing the receiving weights for several units is to stack the weight matrices for neighboring receiving units (a receiving group) into a regular matrix-like pattern. Hence, for example, the weight matrix for unit (0, 0) in V1—the leftmost bottom unit—would be placed in position (0, 0) in the stacked weight matrices plot. Looking at the stacked weight matrices plot, it becomes evident that various receiving units in V1 have specialized for detecting different types of cloud segments. Depending on which of these features are present in the satellite image, the corresponding units in V1 will be activated simultaneously, indicating the presence of a particular combination of features. Based on this activation pattern, subsequent layers (e.g. V2) will be able to recognize the overall cloud shape.

6 Discussion

The network was able to learn the task quickly (in 15-20 epochs), and this learning behavior was stable, which indicates that the network architecture (size and structure of layers and connection structure) was sound.

Because we had limited computational resources available, we were forced to use relatively low-resolution input images (cf. Fig. 5). At this level of resolution, there is a risk that elongated shapes are perceived as equally thick at both ends, making it impossible to distinguish the cyclone itself from the adjacent cumulonimbus cloud. It thus becomes difficult to determine which end of the elongated shape is occupied by the cyclone, and which part constitutes the surrounding cumulonimbus clouds.

In spite of the low-resolution images (66 x 66 pixels) that were used, it was possible to train the network and achieve correct cyclone track predictions for 99% of the novel images, when small variations in the cyclone position were allowed. According to our analysis, the network was able to learn to interpret the shape of cloud patterns in the cyclone, and map the orientation of this overall shape to a correct movement direction. Most importantly, the network could achieve this irrespective of variations in position and apparent size of the cyclone.

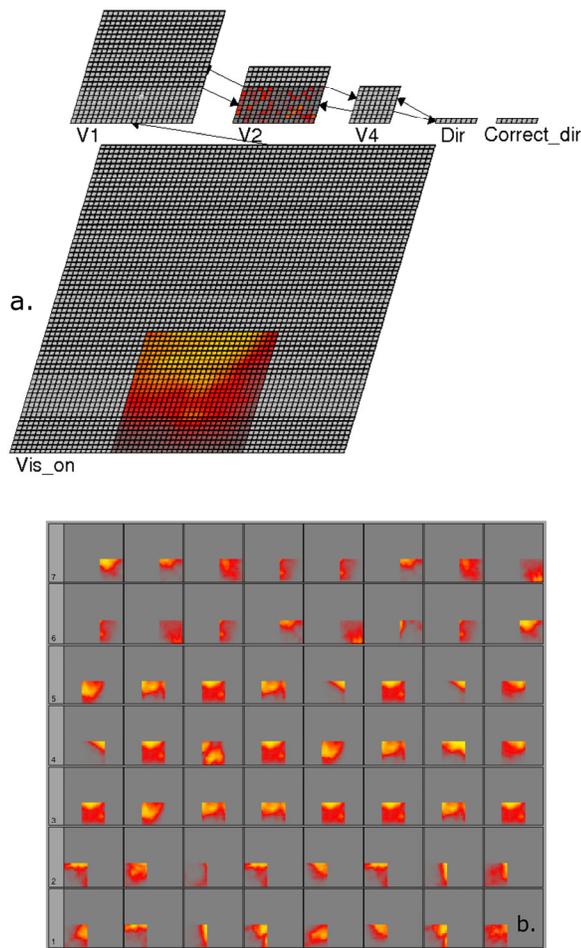


Fig. 8. a. Receiving weights for to a particular node in V1 developed during training. This node receives input from part of Vis_on (comprising the RF for this node), and also receives feedback signals from V2. Note that the weight pattern from Vis_on resembles a cloud segment—this is fact the feature that this particular receiving node specializes for. **b.** Receiving weights for a group of neighboring nodes in V1. Square (x, y) shows the weight pattern from part of Vis_on into node (x, y) within the receiving group in V1.

When larger variations in cyclone position were allowed, for about 10 % of the novel images, the network mistakenly interpreted the image as being upside-down or mirrored, making predictions that were approximately 180° off. This type of error entails that the network successfully recognized the elongated shape that was formed by the cyclone and the surrounding clouds, but that the orientation of this overall shape was mistakenly recognized as being upside-down and/or left-right mirrored. These mistakes are most likely caused by the low-resolution images that were used as input.

In a relatively few cases (1.5 % when cyclone position was varied by 18 % of total image size), the direction predicted by the network was orthogonal to correct cyclone movement direction. This result probably reflects the complex training situation that arose when individual features were allowed to move outside the receptive field of individual feature detectors.

The present study is a proof of concept and is meant to demonstrate the applicability of neural network techniques for cyclone track prediction. This work constitutes a first step towards fully automated interpretation of satellite images for the purpose of cyclone track and intensity prediction. Among other things, the technique remains to be tested on higher resolution images of a wider variety of cyclones.

We see several avenues for future work. First of all, it would be interesting to develop a scaled-up version of the network, and to compare its performance with the present network to see if higher resolution images contain additional useful information. Although our present training results indicate that low-resolution images may be sufficient, we expect that some cases, especially at the early stages of cyclone development, may require recognition of subtle cloud patterns, which may not be extractable from low-resolution satellite images.

Second, the kind of deep recurrent (bi-directionally connected) network that we have been using could easily be redesigned to take into account also other non-visual factors, such as previous cyclone track and current movement direction of the cyclone, as well as temperature and air pressure data. Inclusion of supplementary input to the network would impose additional constraints on the activation state of the network. Provided that the constraints are intrinsically compatible, they would probably improve activation settling in the network, and would probably boost learning performance—at the expense of slower network simulations.

7 References

- Barsi, A., and Heipke, C., 2003a. Artificial neural networks for the detection of road junctions in aerial images. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences* 34 (Part 3/W8), 113-118.
- Barsi, A., and Heipke, C., 2003b. Detecting road junctions by artificial neural networks
- Brad, R., and Letia, I.A., 2002a, Extracting cloud motion from satellite image sequences, *Seventh International conference on control, Automation, Robotics and Vision (ICARCV'2002)*, Singapore.
- Brad, R., and Letia, I.A., 2002b, Cloud motion detection from infrared satellite images, *The International Society for Optical Engineering (SPIE), Second International Conference on Image and Graphics*, 4875, 408-412.
- Bureau of Meteorology Research Centre, 2006, Global guide to Tropical Cyclone forecasting, Australia. http://www.bom.gov.au/bmrc/pubs/tcguide/globa_guide_intro.htm Visited 2007-01-12.
- Cerveny, R. S. and Newman, L. E., 2000, Climatological relationships between tropical cyclones and rainfall. *Monthly Weather Review*, 128, 9, 3329–3336.
- Creem, S. H., and Proffitt, D. R., 2001, Defining the cortical visual systems: “What”, “Where”, and “How”, *Acta Psychologica*, 107, 1-3, 43-68.
- Dvorak, V.F., 1975, Tropical Cyclone Intensity Analysis and Forecasting from Satellite Imagery, *Monthly Weather Review*, 103, 5, 420-430.
- Felleman, D. J., and Van Essen, D. C., 1991, Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex*, 1, 1-47.
- Fett, R.W. and Brand, S., 1975, Tropical cyclone movement forecasts based on observations from satellites, *Journal of Applied Meteorology*, 14, 4, 452-465.
- Fukushima, K., and Miyake, S. (1982). Neocognitron: A new algorithm for pattern recognition tolerant of deformations and shifts in position. *Pattern Recognition*, 15, 6, 455-469.
- Goodale, M. A., and Milner, A. D., 1992, Separate visual pathways for perception and action. *Trends in Neuroscience*, 15, 1, 20-25.
- Hong, Y., Hsu, K.-L., Sorooshian, S., Gao, X., 2004, Precipitation Estimation from Remotely Sensed Imagery using an Artificial Neural Network Cloud Classification System, *Journal of Applied Meteorology*, 43.
- Jang, J.-D., Viau, A.A., Anctil, F., and Bartholomé, E., 2006, Neural network application for cloud detection in SPOT vegetation images, *International Journal of Remote Sensing*, 27, 4, 7 19-736

- Kishtawal, C. M., Patadia F., Singh R., Basu S., Narayanan M. S., and Joshi P. C., 2005, Automatic estimation of tropical cyclone intensity using multi-channel TMI data: A genetic algorithm approach, *Geophysical Research Letters*, vol 32.
- Knaff, J. A., Zehr, R. M., Goldberg, M. D., 2000, An example of temperature structure differences in two cyclone systems derived from the advanced microwave sounder unit. *Weather and Forecasting: 15*, 4, 476–483.
- Kossin J., 2003, *A user's guide to the UW-CIMSS Tropical Cyclone Intensity Estimation (TIE) Model*. Cooperative Institute for Meteorological Satellite Studies (CIMSS), University of Wisconsin-Madison.
- Kosslyn, S. M., 1994, *Image and Brain: The resolution of the imagery debate*. Cambridge MA, MIT Press.
- Lau, N. C. and Crane, M. W., 1997, Comparing satellite and surface observations of cloud patterns in synoptic-scale circulation systems. *Monthly Weather Review*, American Meteorological Society, 125, 12, 3172–3189.
- Lajoie, F. A., 1976, On the direction of movement of tropical cyclones, *Australian Meteorological Magazine*, 24, 95-104.
- Marshall, J. F. L., Leslie, L. M., Abbey Jr. R. F., and Qi, L., 2002, Tropical cyclone track and intensity prediction: The generation and assimilation of high-density, satellite-derived data. *Meteorology and Atmospheric Physics*, 80, 1-4, 43-57.
- McBride, J.L., Holland, G.J., 1987, Tropical-Cyclone Forecasting: A Worldwide Summary of Techniques and Verification Statistics, *Bulletin of the American Meteorological Society*, 68, 10, 1230-1238.
- McClelland, J. L. (1993) The GRAIN model: a framework for modeling the dynamics of information processing. In Meyer, D. E. and Kornblum, S., (eds): *Attention and Performance XIV: Synergies in Experimental Psychology, Artificial Intelligence, and Cognitive Neuroscience*, 655–688. Lawrence Erlbaum.
- Mutch, J, and Lowe, D. G. (2008). Object class recognition and localization using sparse features with limited receptive fields. In press. *International Journal of Computer Vision*. DOI 10.1007/s11263-007-0118-0.
- Naval Research Laboratory, 1999, Tropical Cyclone Forecasters Reference Guide, <http://www.nrlmry.navy.mil/~chu/> Visited 2007-01-12.
- O'Reilly, R. C., & Munakata, Y., 2000, *Computational explorations in cognitive neuroscience: Understanding the mind by simulating the brain*. Cambridge, MA: MIT Press.
- O'Reilly, R. C., 1998, Six principles for biologically-based computational models of cortical processing. *Trends in Cognitive Sciences*, 2, 11, 455-462.
- Rivolta, G., Marzano, F.S., Coppola, E., and Verdecchia, M., 2006, Artificial neural-network technique for precipitation nowcasting from satellite imagery, *Advances in Geosciences*, 7, 97-103.
- Spratling, M. W. (2005). Learning Viewpoint Invariant Perceptual Representations from Cluttered Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27, 5, 753-761
- Ungerleider, L. G., and Mishkin, M., 1982, Two cortical visual systems. In Ingle, D. J., Goodale, M. A., and Mansfield, R. J. W. (eds): *The analysis of visual behavior*. Cambridge, MA: MIT Press.
- Velden, C.S., Olander, T.L., and Zehr, R.M., 1998, Development of an Objective Scheme to Estimate Tropical Cyclone Intensity from Digital Geostationary Satellite Infrared Imagery, *Weather and Forecasting*, 13, 1, 172-186.
- Wells, F. H., 1987, Tropical cyclone intensity estimates using satellite data: the early years. *Proceedings of 26th Conference on Hurricanes and Tropical Meteorology*.