

A Ground Truth Tool for Synthetic Aperture Radar (SAR) Imagery

I. Pavlidis¹ D. Perrin²

N.P. Papanikolopoulos² W. Au¹

S. Sawtelle³

¹Honeywell Technology Center
3660 Technology Drive
MN65-2500
Minneapolis, MN 55418

²University of Minnesota
Department of Computer Science
200 Union Street S.E.
Minneapolis, MN 55455

³WL/AACR
2241 Avionics Circle
Wright Pat. AFB
OH 45433-7318

Abstract

The performance of Computer Vision algorithms has made great strides and it is good enough to be useful in a number of civilian and military applications. Algorithm advancement in Automatic Target Recognition (ATR) in particular, has reached a critical point. State-of-the-art ATRs are capable of delivering robust performance for certain operational scenarios. As Computer Vision technology matures and algorithms enter the civilian and military marketplace as products, the lack of a formal testing theory and tools becomes obvious. In this paper, we present the design and implementation of a Ground Truth Tool (GTT) for Synthetic Aperture Radar (SAR) imagery. The tool serves as part of an evaluation system for SAR ATRs. It features a semi-automatic method for delineating image objects that draws upon the theory of deformable models. In comparison with other deformable model implementations, our version is stable and is supported by an extensive Graphical User Interface (GUI). Preliminary experimental tests show that the system can substantially increase the productivity and accuracy of the Image Analyst (IA).

1 Introduction

Until recently, the emphasis in the Computer Vision community was more on the development of algorithms and less on their experimental evaluation. Researchers used to demonstrate the experimental validity of their algorithms on an arbitrary and limited set of images. As some of these algorithms mature and move towards the productization phase, they necessitate the use of a formal evaluation method and tools. This may involve the testing of the algorithm against a representative sample of images. Since in product evaluation we need to establish statistical variability, the imagery sample is usually in the order of thousands.

Computer Vision algorithms could be considered as transformations from the image domain to the annotated image domain. In other words, a Computer Vision algorithm produces a segmented image, the various parts of which are labeled. Therefore, the test imagery needs to be segmented and labeled by an *objective agent* (ground truth) in order to evaluate the quality of the algorithm's output. In statistical and neural network algorithms, part of the annotated imagery is also used as a training sample.

As mentioned earlier, the testing imagery can be quite voluminous and therefore, its ground truth can prove labor- and cost-intensive. This is especially true in the case of model-based ATRs. Model-based ATRs should perform reliably under a variety of operating conditions and hence should be subjected to extensive testing. In addition, model-based ATRs usually employ SAR sensor phenomenology which relates to images far different than the visible range images. This means that the ground truth is performed by trained IAs which increases the cost even further.

The most time consuming and demanding part of image ground truth is the segmentation part. Since in this paper we concentrate rather on this aspect and not on image labeling, when we refer to ground truth we mean image segmentation only. So far, in the ATR domain, ground truth of images is performed usually by hand. The user delineates the objects of interest in the image using the pointing device (usually a computer mouse) and a suitable GUI.

There is a clear need for a semi-automatic ground truth method that will minimize human input, increase speed and accuracy, and work for all images irrespectively of operating scenarios. We have designed and implemented a ground truth tool that performs along these guidelines. It draws upon the theory of "snakes" first introduced by Kass *et al.* [4].

The objects of interest in military SAR images are targets and their shadows. The targets appear very different from their shadows. Targets feature jagged bright areas interspersed with small gaps. Shadows feature a single smooth

and continuous dark area. Accordingly, we developed two snake varieties for the ground truth of SAR images. The first one works as an elastic contour that contracts and slithers from outside in and is more suitable for target delineation. The second one works as a balloon that expands from inside out and is more suitable for shadow delineation. A key element of our approach is a well thought user interface that frees the user from the intricacies of the task. Another key element, is the effort we put to achieve stable behavior from the tool. Stability is the major problem of “snakes,” and is what usually prevents them from becoming practical. Some preliminary use of the ground truth tool shows significant advantages over the traditional manual method.

We have structured the rest of the paper as follows: In Section 2 we present an overview of the ground truth system. In Section 3 we present in some detail the snake variety for segmenting shadows. Then, in Section 4 we present the snake variety for segmenting targets. Comparative experimental results between the manual delineation method and the snake delineation method are presented in Section 5. The paper concludes in Section 6.

2 System Overview

The design of the GTT presented algorithmic as well as systems challenges. The algorithmic challenges were addressed through the development of semi-automatic object extractors based on “snakes.” The systems challenges were addressed through object-based design and methodology.

GTT was developed in response to the needs of the Wright Laboratory of the U.S. Air Force for evaluating the Moving and Stationary Target Acquisition and Recognition (MSTAR) system. MSTAR is a state-of-the-art model-based ATR with unprecedented testing needs [5]. It requires the ground truth of several thousands of SAR images just for evaluation purposes. MSTAR represents a paradigm shift in the ATR domain and gives us a flavor of things to come. Ground truth consists of two main operations: image segmentation and image labeling. Image segmentation for ground truth purposes is traditionally performed by hand. This is not a cost effective solution, however, for the case of MSTAR because of the sheer volume of imagery involved. Image segmentation is one of the primary research problems in computer vision for which no satisfactory solution yet exists. Therefore, a fully automated segmentation tool that would perform reliably is still not within our reach. In contrast, semi-automatic segmentation methods, when designed appropriately and coupled with a good GUI produce robust and attractive systems.

We based our algorithmic approach for the semi-automatic object extractor to the *statistical snakes* developed by Ivins *et al.* [3]. We chose this type of snakes because they were applied successfully in medical images

(Nuclear Magnetic Resonance and Computer Tomography) which bear some phenomenology similarities with military SAR images. We appropriately modified the statistical snakes to fit the task at hand. In particular, we developed two varieties: one for capturing targets and another for capturing target shadows. The background can result from simple image differencing once the positions of the targets and their shadows are known. A very important component of our approach is the careful design of the user interface. The user interface allows the user to complete the object extraction in two steps; one for denoting the approximate position of the object in the image and one for controlling the termination of the snake.

At the system level, we had to build an open system, amenable to expansion and with minimal maintenance requirements. We had also, to incorporate in our design, and modify if appropriate, some legacy code related to the traditional hand segmentation tool. To meet the above challenging specifications we opted for an object based approach. Fig. 1 shows the object diagram of the GTT. The development methodology and the notation conform to the Booch method [1]. Object diagrams show the existence of objects and their relationships and provide a trace of the system’s behavior.

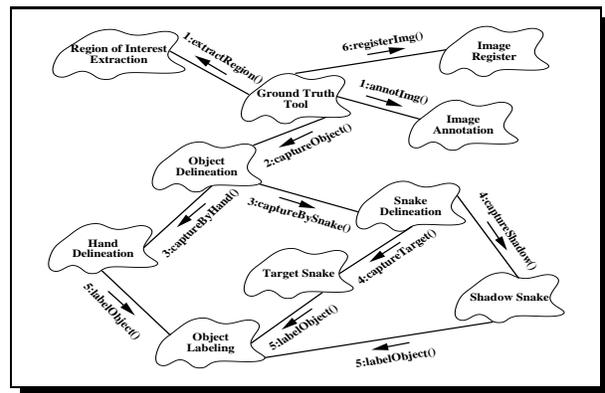


Figure 1: Object diagram of the GTT.

The scenario of a typical GTT use starts with the extraction of a Region of Interest (ROI) or with Image Annotation. We define as ROI in a wide area image, a portion that contains a single target and its shadow. Image Annotation refers to the specification of the operational conditions and the image type. Either of these operations could be executed first, at the discretion of the user, and this is why they are preceded by the same sequence numeral in Fig. 1. Object Delineation and Labeling should follow. Finally, the Ground Truth Tool invokes a registerImg() operation upon the Image Register object that registers the ground truth products in the Labeled Image Base.

3 Shadow Snake

Target shadows in SAR images usually feature a single connected region of dark color degradations. It is the region from which the SAR sensor does not get almost any returns due to the target geometry. Since the region is usually connected and almost uniform in the center, it is a legitimate candidate for the application of a *balloon*. Balloons are snakes that inflate from the inside out until they hit strong edges. We based our balloon approach on the *statistical snake* proposed by Ivins *et al.* [3] because in comparison to the original balloon model [2] features three major improvements:

1. Adaptive control of the pressure force.
2. Insertion and deletion of boundary elements as the model expands and contracts.
3. Control of boundary intersections.

The above improvements worked well on medical images, an indication that would rather work well on SAR images too, since the relevant sensor phenomenologies bear similarities. We will elaborate on the above issues and describe our shadow snake implementation, emphasizing where appropriate, the innovations we introduced in the original method of Ivins *et al.* [3].

We define as a *shadow snake* a two dimensional curve $s(\lambda) = (x(\lambda), y(\lambda))$. The curve can be considered as a band the mechanical properties of which are specified by the following energy functional:

$$E_{shadow} = E_{tension} + E_{stiffness} + E_{pressure}, \quad (1)$$

where,

$$E_{tension} = \frac{1}{2} \oint \alpha \left| \frac{\partial \mathbf{s}}{\partial \lambda} \right|^2 d\lambda, \quad (2)$$

$$E_{stiffness} = \frac{1}{2} \oint \beta \left| \frac{\partial^2 \mathbf{s}}{\partial^2 \lambda} \right|^2 d\lambda, \quad (3)$$

$$E_{pressure} = \frac{\rho}{2} \oint \frac{\partial \mathbf{s}}{\partial \lambda} \times \mathbf{s} d\lambda. \quad (4)$$

The energy term $E_{tension}$ makes the snake behave like an elastic band by introducing tension. The energy term $E_{stiffness}$ makes it resist bending by introducing stiffness. Finally, the energy term $E_{pressure}$ is an isotropic pressure potential that controls the evolution of the area enclosed by the snake. Contrary to the implementation in [3] we have not included an energy term that is generated by highlighted image edges (E_{edge}). We found that the pressure term alone produces satisfactory behavior. The E_{edge} term plays a crucial role in establishing a stopping condition. Since the

snake expansion is controlled by the user through the pointing device, the E_{edge} energy factor is not necessary. This makes the method more computationally efficient. Giving the user the power to stop the snake expansion at his/her discretion, was a design decision we arrived at after interviewing several IAs. IAs consistently felt that they would like GTT to do most of the work for them, but they expected to have some control over the final object shape. This is a reasonable attitude in the SAR image domain, because sometimes edge segments give false impression about the extent of the target or its shadow. Therefore, image edges in this sensor modality cannot always be considered a reliable termination point.

Starting from an initial circular arrangement the snake is subjected to an iterative energy minimization process that is defined from the differentiation of Eq. (1). In particular, the snake moves in the image plane according to the following equation:

$$\frac{\partial \mathbf{s}}{\partial t} = \alpha \frac{\partial^2 \mathbf{s}}{\partial \lambda^2} - \beta \frac{\partial^4 \mathbf{s}}{\partial \lambda^4} + \rho \left(\frac{\partial \mathbf{s}}{\partial \lambda} \right)^\perp \left(\frac{|I(\mathbf{s}) - \mu|}{\kappa \sigma} - 1 \right). \quad (5)$$

The first term on the right-hand side of the equation represents the tension force $F_{tension}$ in the snake that is produced by the energy term $E_{tension}$. The second term represents the stiffness force $F_{stiffness}$ and is produced by the energy term $E_{stiffness}$. The last term represents the pressure force $F_{pressure}$ and is produced by the energy term $E_{pressure}$. $I(\mathbf{s})$ is the original image. μ and σ are the mean and standard deviation of the pixel values from a portion $I_{seed}(\mathbf{s})$ of the image $I(\mathbf{s})$. $I_{seed}(\mathbf{s})$ is defined interactively by the user to be within the shadow area and is called the *seed region*. The terms α , β , ρ , and κ are user defined constants.

Four alternate pressure forces are proposed in [3]: binary, linear, quadratic, and sigmoid. As can be seen in Eq. (5), we chose the linear pressure force model. The reason is that it expresses best the pixel variance of shadows in SAR images: almost zero in the center, but growing about linearly towards the outer shadow boundary. The snake expands rapidly when the image pixels that it encounters have approximately the same value as the mean μ of the seed region. In our case, this happens within the target shadow area. When $|I(\mathbf{s}) - \mu| = \kappa \sigma$ the pressure force will be zero, bringing the model to a virtual stop. In our case, this happens after the snake has overstepped the outer boundary of the target shadow and started encountering pixel values that are κ standard deviations away from the mean μ . Some oscillation may take place at this point but since the user controls the minimization process through the GUI can stop it at will.

The values of the user defined constants have been specified experimentally to be as follows: $\alpha = 1.0$, $\beta = 1.0$,

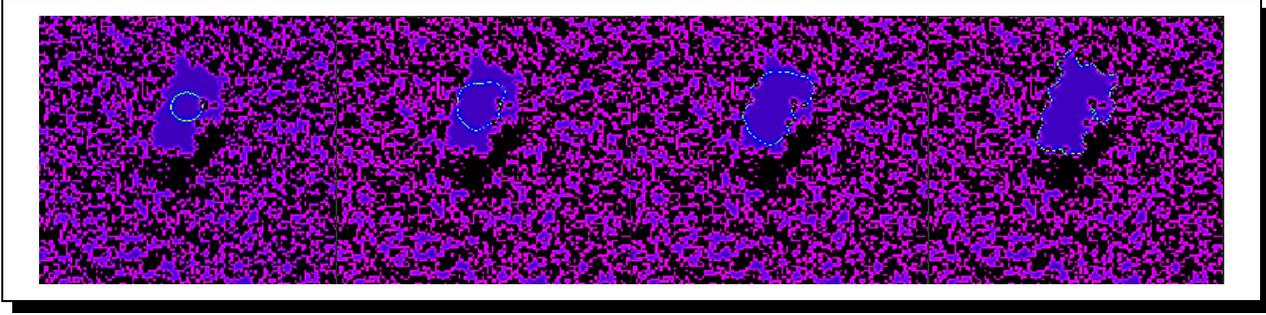


Figure 2: Snapshots of the snake evolution inside the shadow area of a target. The figures do not depict the original SAR image but its corresponding pressure field.

$\rho = 1.0$, and $\kappa = 0.02$. These values are valid for typical SAR images only. In case of another sensor modality, like IR, the above values have to be re-specified and entered through the GUI. The user can capture and label the target shadow of a SAR image by performing three steps. In the first step (mouse click), he/she has to specify the seed region to be somewhere within the shadow area. In the second step (mouse click), he/she has to specify the initial placement of the snake to be also somewhere within the shadow area. In the third step, he/she controls the iterative minimization process which can stop at will.

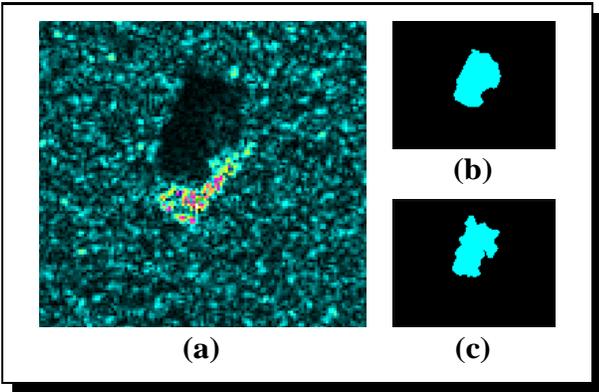


Figure 3: (a) Original SAR image. It depicts a Russian armored personnel carrier (BMP) shot at a 15° depression angle. (b) Labeled shadow chip segmented by hand. (c) Labeled shadow chip segmented by snake.

Fig. 2 presents a visualization of the snake motion in one of the SAR images we used in our experiments. Actually, the figure depicts not the original SAR image but its corresponding pressure field as it was computed after the specification of the seed region by the user. Brighter spots represent areas of lower pressure while darker spots represent areas of higher pressure. The snake expands from the center portion of the target shadow (solid gray area). The expan-

sion is primarily driven from the pressure force $F_{pressure}$ which pushes the snake boundary from the central higher pressure area towards the outer lower pressure areas. The forces $F_{tension}$ and $F_{stiffness}$ play a regularization role. The end result in the form of a labeled (color coded) chip is shown in Fig. 3(c). Fig. 3(a) shows the original SAR image upon which the shadow snake was applied. Fig. 3(b) shows the labeled chip of the target shadow segmented manually by a SAR expert. The reader may notice that the labeled chips (b) and (c) are very similar and in conformance with the general shadow shape of the original image (a). The only difference appears to be in the greater amount of shape detail captured by the snake.

4 Target Snake

Target regions in SAR images differ fundamentally from their shadows. They feature much greater variance of pixel values than the corresponding shadows. Parts of the metal skeleton of the target function as corner reflectors to the radar beam and show as bright spots in the image. Other parts scatter somewhat the radar beam and appear as spots of moderate intensity. Spots of high and moderate intensity are intermixed in the target area producing a distinct non-uniform texture. If we used part of the target area as the seed region, the resulting mean μ pixel intensity would be far less useful than the case of the shadow. Interspersed corner reflectors would most likely lie outside the $\kappa\sigma$ range prescribed by Eq. (5). Hence, they will function as faulty stopping points in the expansion of the snake curve. Also, the rather thin and concave shape of many SAR targets does not facilitate the initial placement of the balloon snake in its interior.

We addressed the above problems by modifying the shadow snake as follows: We introduced thresholding to binarize the pressure field. The user selects a threshold pixel value that results in a mostly positive (+1) target region. Automatically, the area surrounding the target region assumes

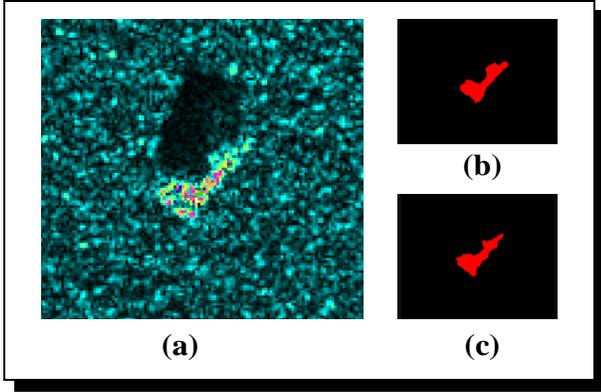


Figure 4: (a) Original SAR image. It depicts a Russian armored personnel carrier (BMP) shot at a 15° depression angle. (b) Labeled target chip segmented by hand. (c) Labeled target chip segmented by snake.

a mostly zero value (0). These binary values are used in the pressure term of Eq. (5) in the place of the linear model. Specifically, the motion equation for the target snake becomes:

$$\frac{\partial \mathbf{s}}{\partial t} = \alpha \frac{\partial^2 \mathbf{s}}{\partial \lambda^2} - \beta \frac{\partial^4 \mathbf{s}}{\partial \lambda^4} + \rho \left(\frac{\partial \mathbf{s}}{\partial \lambda} \right)^\perp (T(\mathbf{s}) - 1). \quad (6)$$

where $T(\mathbf{s}) = 0$ or 1 and represents the thresholded image functional. The user defined constants α , β , and ρ maintain the same values as in the case of shadow snake. The pressure force in Eq. (6) is maximal when $T(\mathbf{s}) = 0$. In other words, the snake moves rapidly when it encounters thresholded pixels that belong mostly to the background. For this, the user places initially the snake not inside the target but around the target. The snake then moves rapidly towards the mostly positive pressure area of the target. Once it oversteps the target boundary it comes to a virtual stop because $T(\mathbf{s}) = 1$ for the most part which means minimal pressure force. This arrangement of the target snake results mostly in a contraction (and not an expansion) of the initial snake curve during the iterative minimization process. The user freezes the snake once he/she is satisfied with the approximation of the target outline. During the contraction, the snake may stumble upon positive pressure outliers due to imperfect thresholding. This may result temporarily to curve self-intersections which are untangled following the method in [3].

The user can capture and label the target of a SAR image by performing three steps. In the first step (mouse click) he/she selects the threshold value. The selection of the threshold value takes place in a very intuitive manner. The user moves a threshold slider while he/she views in real time how the pressure field of the image changes. He stops when he feels that the pressure field makes justice to the shape

of the target. In the second step (mouse click) the user positions the initial snake curve around (and not inside) the target. In the third and final step the user controls the iterative minimization process which can stop at will. Fig. 4 mirrors Fig. 3 for the case of the target.

5 Experimental Results

After some preliminary use of the GTT and especially its snake facilities, IAs at our lab found that they can perform image ground truth at a fraction of the time they spent in the past. They also found, that ground truth through mouse clicks is considerably less tiresome than ground truth through hand delineation. The question, however, is how accurate the snake tools are comparatively with the manual method. Some relevant performance measures have been reported in the literature [6]. Admittedly, there is no objective way to define the “true” boundaries of the target and its shadow in a SAR image. The snake algorithms are to be used to aid IAs who would otherwise extract the boundaries manually. Therefore, we chose to evaluate the target and shadow snake algorithms by checking the consistency of their results with the corresponding manual measurements. As manual measurements we took the hand delineations of targets and shadows from 9 typical military SAR images. The delineations were carefully performed by a SAR expert and they can be safely considered as templates. Snake delineations were performed by two different users on the same set of images. Our method of comparison checks the amount of overlap between the corresponding labeled chips in the case of hand delineation (the templates) and in the case of snake delineation.

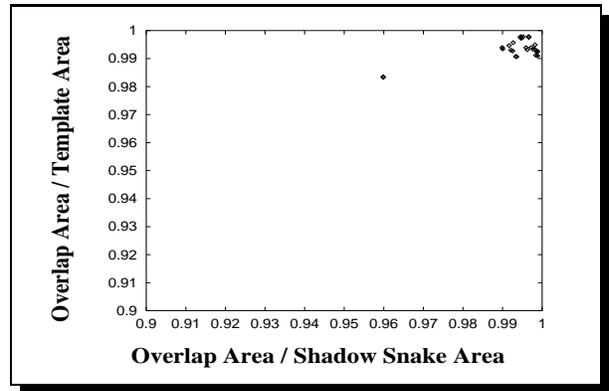


Figure 5: Comparison of manually generated shadow curves (templates) with shadow snake curves. The Overlap Area corresponds here to the intersection of the Template Area and the Shadow Snake Area.

Fig. 5 shows the experimental evaluation diagram for the case of the shadow snake and Fig. 6 for the target snake. If

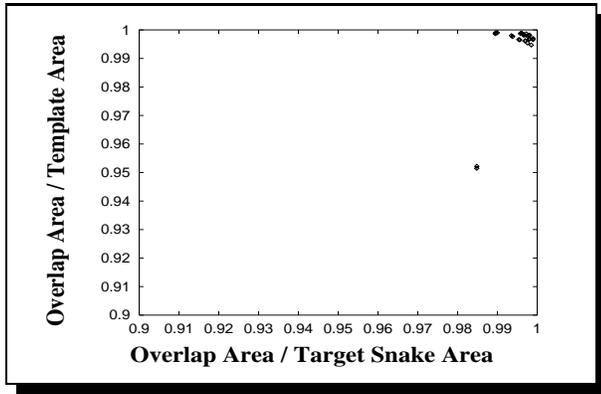


Figure 6: Comparison of manually generated target curves (templates) with target snake curves. The Overlap Area corresponds here to the intersection of the Template Area and the Target Snake Area.

the plotted point is (1, 1) the target or shadow segmented by the snake is exactly the same with the corresponding target or shadow delineated manually by the SAR expert. Therefore, the closer the points are to the (1, 1) corner, the better the performance of the snake. As can be seen in the figures the performance of both snakes is almost identical to the SAR expert.

6 Summary and Conclusions

We have designed and implemented a new system (GTT) for performing ground truth operations on SAR images. The two most important components of our system are two snake tools: one for delineating targets and one for delineating shadows. The snakes have been designed in a way that addresses the challenges present in the SAR imagery. IA feedback verified that ground truth with the snake tools is considerably faster and easier than manual ground truth. This is especially important for the mission of GTT which is the ground truth of thousands of SAR images for the evaluation of MSTAR. Comparative experiments, have also demonstrated that shadow and target snake segmentations are as good as careful manual delineations by SAR experts. In the near future, we plan on enhancing the labeling mechanism of GTT by indexing object attributes such as geometry and texture.

Acknowledgements

We would like to especially thank Dr. Peter Symosek and Mr. Joe Keller for their valuable comments and help. This research project was supported partially by the Wright Labs of the U.S. Air Force under contract #F33615 – 96 –

C – 1819. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the funding agency.

References

- [1] G. Booch. *Object-Oriented Analysis and Design*. The Benjamin/Cummings Publishing Company, Redwood City, California, 1994.
- [2] L. Cohen. “Note: On Active Contour Models and Balloons”. *CVGIP (Image Understanding)*, 53(2):211–218, 1991.
- [3] J. Ivins and J. Porrill. “Active Region Models for Segmenting Medical Images”. In *Proceedings IEEE International Conference on Image Processing*, volume 2, pages 227–231, 1994.
- [4] M. Kass, A. Witkins, and D. Terzopoulos. “Snakes: Active Contour Models”. *International Journal of Computer Vision*, 1(4):321–331, 1988.
- [5] E. Keydel, S. Lee, and J. Moore. “MSTAR Extended Operating Conditions: A Tutorial”. In *Proceedings SPIE*, volume 2757, pages 228–242, 1996.
- [6] R. Samadani, D. Mihovilovic, C. R. Clauer, G. Wiederhold, J. Craven, and L. Frank. “Evaluation of an Elastic Curve Technique for Finding the Auroral Oval from Satellite Images Automatically”. *IEEE Transactions on Geoscience and Remote Sensing*, 28(4):590–597, 1990.