

A hybrid method based on CNNs and edge-based active contour models for medical image segmentation

A. Bendaoud* and F. Hachouf

*Laboratoire d'Automatique et de Robotique. Faculté des Sciences de la Technologie
Université Constantine1. Constantine, Algérie.*

*Corresponding author, emails : aa.bendaoud@gmail.com & hachouf.fella@gmail.com

Selected paper of JSSE'18, received in final form: March 4, 2019 ; accepted date: March 23, 2019

Abstract

Edge-based active contour models have been one of the most prominent and influential approaches in image segmentation. It has been proven that they are very effective when they are applied on images with inhomogeneous intensity. Traditional edge-stop functions (ESFs) are usually used when edges are defined by the image gradient. They often produce weak edges because they fail to stop at the precise boundary. In this work, a new approach integrating machine learning algorithm with edge-based model using a level set method (LSM) is proposed. The ESF is constructed from a convolutional neural network. Then it is applied to an edge-based active contour model. The proposed method has been applied on medical images. Obtained results have been compared to those given by k-nearest neighbors and support vector machine to confirm the effectiveness of the proposed method.

Keywords: Convolutional neural network; edge-based active contour; image segmentation; medical images

1. Introduction

Image segmentation is one of the main steps in pattern recognition and computer vision. It is a process which divides an image into different regions and extracts out the interested targets. These targets represent some kind of information, for example: color, intensity or texture, etc. For decades, a number of very inspiring and pioneering image segmentation algorithms have been developed. These algorithms have been successful in many applications; however none of them is generally applicable to all images [1]. The snake model introduced by Kass et. al. [2], was the first approach using energy minimization since then it got the interest of scientists and it has been deeply studied. Among energy minimization approaches, the level set method (LSM) has received considerable attention. It has been widely adopted in medical image segmentation applications [3, 4].

In general, existing image segmentation models using LSMs can be divided into two categories: edge-based models and region-based models. While the former uses local edge information to fit the boundaries of the shape [5], the latter uses a region descriptor to manipulate the motion of the active contour [6, 7]. Edge-based techniques [3], [5], [8], carry out the curve evolution geometrically. The stopping criteria are controlled by the extracted edge features from the image. Chan and Vese [7] proposed a level set formulation to minimize the Mumford Shah functional [6] for segmentation. Edge

based techniques are better suited for heterogeneous images, although they are unable to tackle fragmented or blurred edges, which occur frequently in medical applications.

To overcome the limitations of edge-based models, some studies have been carried out by integrating machine learning methods with classic models [9]. In this work convolutional neural networks are used to build the edge stop function which constitutes the main problem in traditional edge based active contour algorithms. The framework of the proposed method is given by Fig.1.

This paper is organized as follows: section 2 presents the proposed approach. In this section, level sets for image segmentation are presented. Then, the convolutional neural networks are detailed because they are hybridized with edge based methods to build the proposed method. Edge stop function construction is explained. In section 3, the obtained results are analyzed and discussed. They are compared with results produced by k-NN and SVM algorithms. Section 4 concludes this work.

2. Proposed approach

In this section, different parts of the proposed approach are detailed.

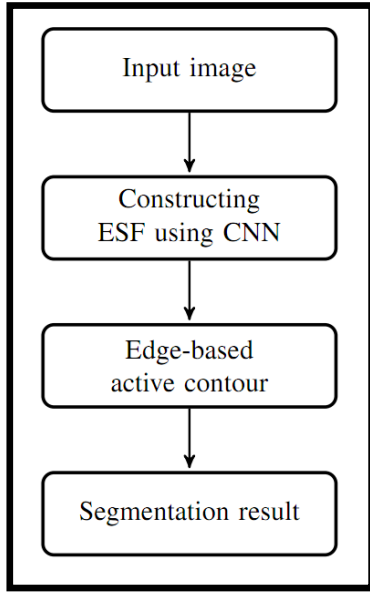


Figure 1. Block diagram of the proposed method

2.1. Level set Methods for image segmentation

The zero level set contour of $\emptyset$ is defined by:

$$C(t) = \{(x, y) | \emptyset(x, y, t) = 0\} \quad (1)$$

where $\emptyset(x, y, t)$ is the zero level set function. The evolution of the curve C can be written [3] as

$$\frac{\partial \emptyset}{\partial t} + F|\nabla \emptyset| = 0 \quad (2)$$

where ∇ is the gradient operator and F is the speed of the evolution. Caselles et al. [10] proposed a geometric active contour model. Later, they presented a geodesic active contour model [5] given by:

$$\frac{\partial \emptyset}{\partial t} = |\nabla \emptyset| \operatorname{div} \left(g \frac{\nabla \emptyset}{|\nabla \emptyset|} + \vartheta g |\nabla \emptyset| \right) \quad (3)$$

where div is the divergence operator, g is the edge stop function and ϑ is a constant.

Let I be an image on a domain Ω . Traditional ESF is represented by:

$$g = \frac{1}{1 + |\nabla(G_\sigma * I)|^p} \quad p = 1, 2, \dots \quad (4)$$

G_σ is a gaussian kernel with standard deviation σ . The inconvenient of the LSF is that it requires a re-initialization [11], [12] to avoid induced numerical errors. This problem was solved by adding a distance regularization term to the geometric active contour model [13].

$$\frac{\partial \emptyset}{\partial t} = \mu \operatorname{div} \left(d_p(|\nabla \emptyset|) \nabla \emptyset \right) + \lambda \delta_\epsilon(\emptyset) \operatorname{div} \left(g \frac{\nabla \emptyset}{|\nabla \emptyset|} \right) + \alpha g \delta_\epsilon(\emptyset) \quad (5)$$

Where μ, λ and α are constants, d_p is a function defined by $d_p(s) = \frac{p'(s)}{s}$ and δ_ϵ is the Dirac delta function.

2.2. Convolutional Neural Network

Convolutional Neural Networks, Convnets or CNNs forshort, are a type of multilayer perceptron (MLP). They are designed for the recognition of two-dimensional visual patterns (Fig.2). There are various CNN architectures in the literature. Yet, their basic components are very similar.

CNNs have been used in several applications including hand-written digit recognition, face detection, and face recognition. The first successful application of convolutional networks has been developed by Yann Le Cun [14]. The break through success of Alex Net for Image Net Large-Scale Visual Recognition Challenge (ILSVRC) in 2012 was the start of the evolution in CNN architectures. Afterwards many efforts have been made in developing various CNN models, including VGGNet [15] GoogLe Net [16], and ResNet [17].

CNNs are a type of feed forward network; however the training process for the feature learning works backwards. To adjust the feature weights and minimize the error, the error is computed between the obtained results and expected results. Then it is distributed back through the network, layer by layer.

A typical CNN consists of three types of layers, namely convolutional, pooling, and fully-connected layers.

2.2.1 Convolutional layer:

The convolutional layer aims to learn feature representations of the inputs. It is composed of several convolution masks which are used to compute different feature maps. A single set of weights is being shared by all the nodes within the network. Afterwards, each output value is mapped to a different numerical value through a transformation. This transformation is called an Activation Function. Its function is to introduce a non-linearity in the network. There exist different activation functions. The commonly used ones are: sigmoid, hyperbolic tangent and Rectified Linear Unit (ReLU) as shown in Fig.3.

a) Sigmoid:

The sigmoid function is defined by:

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (6)$$

Where all values are mapped in the range [0,1]

b) Hyperbolic Tangent (tanh)

tanh is defined by:

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (7)$$

The output values of tanh differ between -1 and 1. tanh is a scaled sigmoid, i.e:

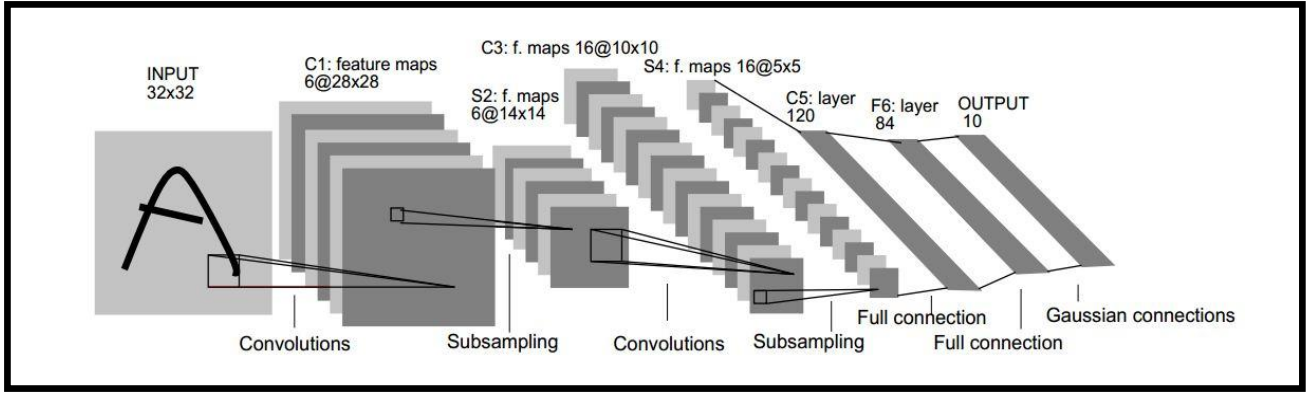


Figure 2. The architecture of the original convolutional neural network, as introduced by LeCun et al. [14]

$$\tanh(x) = 2\sigma(2x) - 1 \quad (8)$$

c) ReLU

The ReLU activation function is simple to implement. It is defined by:

$$f(x) = \max(0, x) \quad (9)$$

ReLU gained its popularity from (ILSVRC) in 2012 after being used as an activation function in AlexNet [18].

fitting. There exist two types of pooling: maxpooling and average pooling. Fig.4 shows both options. While the first

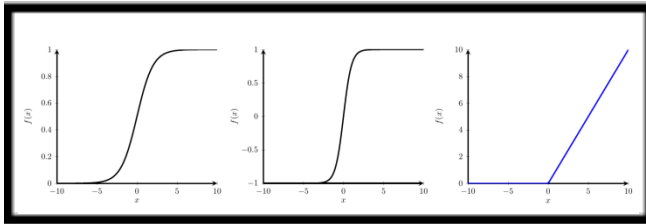


Fig. 3: Popular activation functions: to the left sigmoid, in the middle tanh, to the right ReLU.

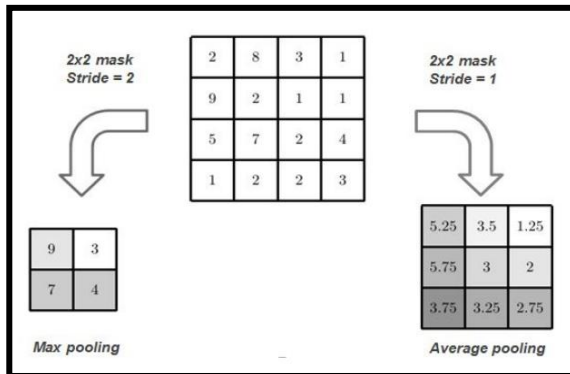


Figure 4. Popular pooling options.

2) *Pooling layer*: Pooling is an important concept of CNN. A Pooling layer is inserted between successive convolutional layers in a Convnet architecture. Its function is to speed up the network by progressively reducing the number of connections between convolutional layers. It leads to reduce the amount of parameters and computation in the network, hence to also control over

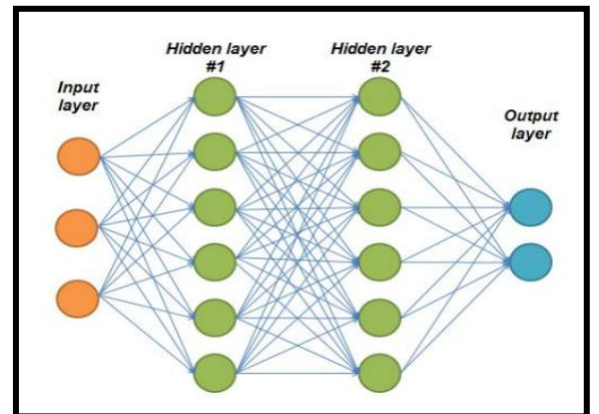


Figure 5. Fully Connected Layers (Multilayer Perceptron)

pools the max value within the mask, the latter pools the average value (the mean of the mask).

3) *Fully connected layer*: After several convolutional and pooling layers, there (FCLs). fig. 5 show a common MLP. FCLs connect every neuron in one layer to every neuron in another layer. Its principle is the same as the traditional multi-layer perceptron neural network. The last one is the output layer. For classification tasks, the softmax operator is commonly used. It is a normalized exponential function, in which all value are in the range [0, 1] and all values add to 1. The softmax is defined by:

$$f(z)_j = \frac{e^{z_j}}{\sum_{k=1}^K e^{z_k}} \text{ for } j = 1, \dots, K \quad (10)$$

where K is the number of classes and z is the data.

2.3. Convolutional neural networks for ESF construction

The main downside of the traditional ESF is that it often fails to stop at the precise target, due to the slow change in gradient. To overcome this inconvenience, the authors in [9] proposed a new ESF. It computes the probability scores by applying a classification algorithm to all the pixels in the image. The obtained scores change in the range $[0, 1]$. In case of a classification score equal to 0.5, there is a high chance of a boundary. The energy at this value is also minimized to the lowest. While for values equal to 0 or 1 the energy is at the highest. The ESF used is defined by:

$$\rho(s) = (2(s - 0.5))^2 \quad (11)$$

s represents the foreground probability score. Another ESF used is:

$$\rho(s) = (\cos \pi s)^p, 0 \leq s \leq 1, p = 2, 4, 6, \dots \quad (12)$$

Fig.6 shows the plot of Eq. (11) and Eq. (12). Both plots have similar characteristics, which is monotonically decreasing in $[0, 0.5]$ and monotonically increasing in $[0.5, 1]$.

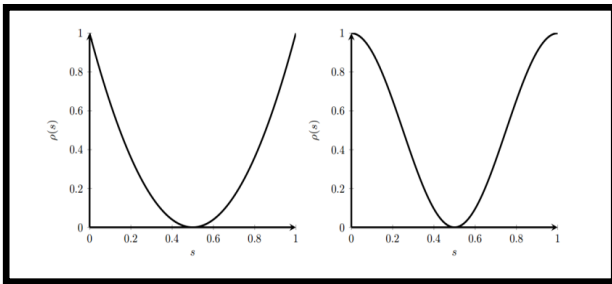


Figure 6. Representation of Eq. (11) and Eq. (12)

3. Results

A. Experimental Setup

The proposed method has been tested on the MRI brain tumor data base [18]. It contains various images with their ground truth segmentation. The images have been segmented by the active contour algorithm using the ESF computed by the convolutional neural network. For evaluation, we compare the obtained results with those of k -NN and SVM. The segmentation process is the same as in [9]: the segmentation is initialized by manually drawing two marks that serve as the training data for the classifier. The red mark refers to the foreground and its boundary is the initial contour while the blue mark refers to the background. The proposed method is implemented in Matlab. An image patch of 3×3 is used to generate the feature vector for all the experiments. Our CNN model is composed of four layers without counting the input; three

convolutional layers plus one fully connected layer. The number of convolution masks (or filters) is randomly set while the weights are initialized manually. In conclusion, both architecture and hyper-parameter tuning are done by using a heuristic approach. To evaluate the segmentation accuracy, two standard measurements are used, i.e., the Jaccard index (JI) and the similarity index (SI) which is also known as the Dice coefficient.

Let A and B be the segmentation result and the ground truth, respectively, the accuracy is defined by:

$$JI = \frac{|A \cap B|}{|A \cup B|} \quad (13)$$

$$SI = \frac{2|A \cap B|}{|A| + |B|} \quad (14)$$

where $|\cdot|$ is the number of elements.

B. Segmentation results

Fig.7 and fig.8 show some of the results from the experiments conducted on a number of MRI images. Table 1 shows the accuracy using different ESFs. Applying CNN gives the best accuracy in both ESFs. It outperformed the k -NN and SVM which gives similar results for JI, with a slight difference for SI. That is when the parameters are well tuned. However, if these parameters are randomly set, the accuracy drops considerably in all methods; but still with CNN giving the best accuracy, as it is shown in Table 2.

4. Conclusion

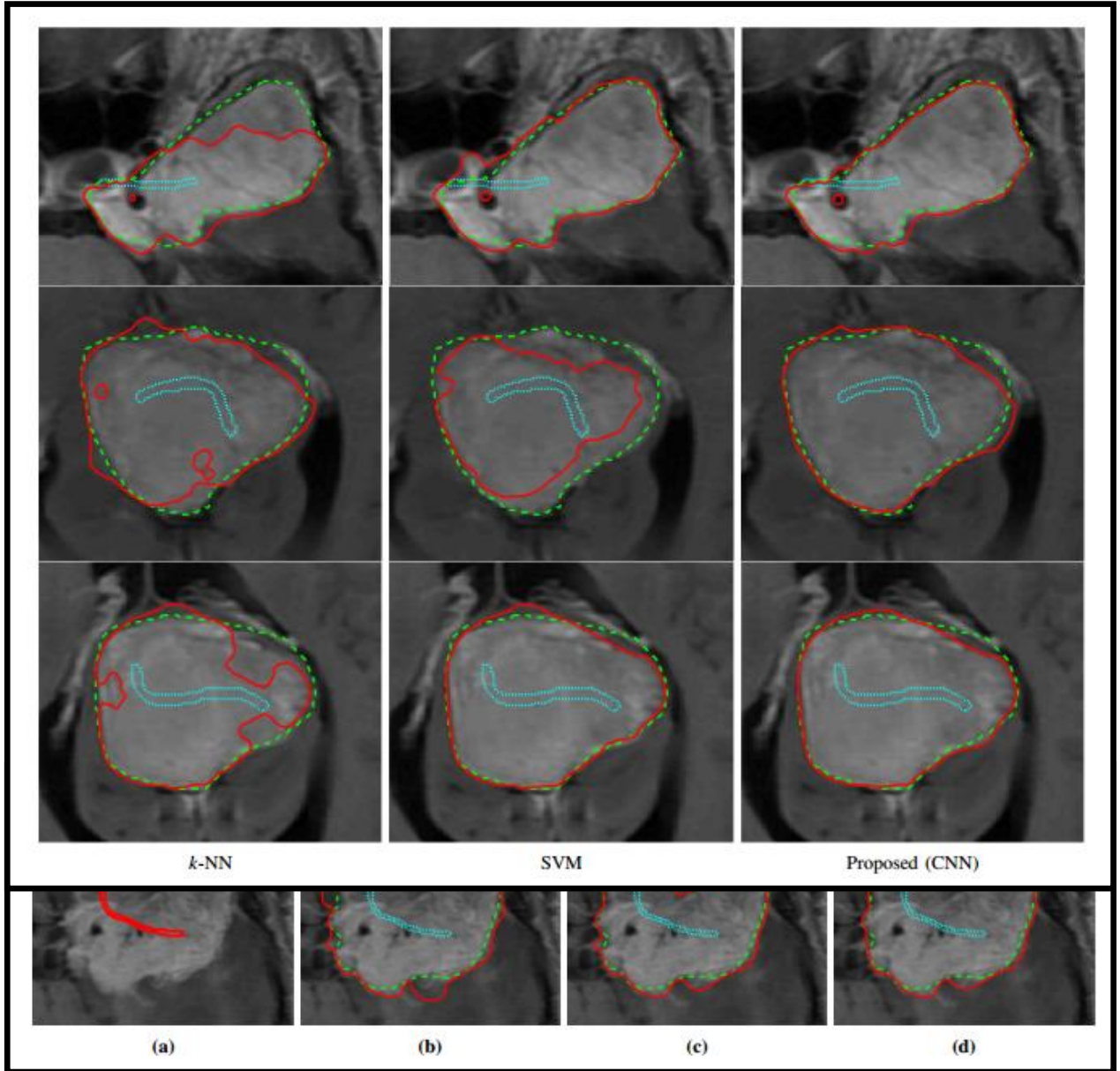
In this paper, a hybrid method that integrates a convolutional neural network with an edge-based active contour model has been proposed. To confirm the effectiveness of this method, an experimental implementation of k -NN and SVM integrated with active contour models were utilized. The proposed method has proven its efficiency. It gives a higher accuracy in comparison to k -NN and SVM integrated with edge-based active contour model. It is neither sensitive to the choice of the regularization function nor to parameter tuning, i.e. setting λ randomly decreases the accuracy but still it gives the higher score.

Table 1: Segmentation accuracy using different ρ

Methods	$\rho(s) = 4(s - 0.5)^2$		$\rho(s) = \cos(\pi * s)^2$	
	JI	SI	JI	SI
k-NN	88.86	94.10	88.32	93.80
SVM	86.43	92.72	86.52	92.77
CNN	91.25	95.72	92.77	96.25

Table 2: Comparison of Segmentation Accuracy

Methods	$\lambda = 1$		$\lambda = 3$		$\lambda = 5$	
	Jl	SI	Jl	SI	Jl	SI
<i>k</i> -NN	49.43	66.16	83.67	91.11	67.33	80.47
SVM	75.12	85.80	67.67	80.71	67.71	80.75
CNN	77.33	87.22	90.54	95.03	78.71	88.09

Figure 8. Brain tumor with: (a)user initialization, the segmentation result for: (b) *k*-NN, (c) SVM, (d) CNN

References

- [1] W. Khan, "Image segmentation techniques: A survey," *Journal of Image and Graphics*, vol. 1, no. 4, pp. 166-170, 2013.
- [2] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *International journal of computer vision*, vol. 1, no. 4, pp. 321-331, 1988.
- [3] S. Osher and J. A. Sethian, "Fronts propagating with curvature-dependent speed: algorithms based on hamilton-jacobi formulations," *Journal of computational physics*, vol. 79, no. 1, pp. 12-49, 1988.
- [4] X. Yang, X. Gao, D. Tao, and X. Li, "Improving level set method for fast auroral oval segmentation," *IEEE transactions on Image Processing*, vol. 23, no. 7, pp. 2854-2865, 2014.
- [5] V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic active contours," *International journal of computer vision*, vol. 22, no. 1, pp. 61-79, 1997.
- [6] D. Mumford and J. Shah, "Optimal approximations by piecewise smooth functions and associated variational problems," *Communications on pure and applied mathematics*, vol. 42, no. 5, pp. 577-685, 1989.
- [7] T. F. Chan and L. A. Vese, "Active contours without edges," *IEEE Transactions on image processing*, vol. 10, no. 2, pp. 266-277, 2001.
- [8] R. Malladi, J. A. Sethian, and B. C. Vemuri, "Shape modeling with front propagation: A level set approach," *IEEE transactions on pattern analysis and machine intelligence*, vol. 17, no. 2, pp. 158-175, 1995.
- [9] A. Pratondo, C.-K. Chui, and S.-H. Ong, "Robust edge-stop functions for edge-based active contour models in medical image segmentation," *IEEE Signal Processing Letters*, vol. 23, no. 2, pp. 222-226, 2016.
- [10] V. Caselles, F. Catte, T. Coll, and F. Dibos, "A geometric model for active contours in image processing," *Numerische mathematik*, vol. 66, no. 1, pp. 1-31, 1993.
- [11] J. A. Sethian, *Level set methods and fast marching methods: evolving interfaces in computational geometry, fluid mechanics, computer vision, and materials science*. Cambridge university press, 1999, vol. 3.
- [12] S. Osher and R. Fedkiw, *Level set methods and dynamic implicit surfaces*. Springer Science & Business Media, 2006, vol. 153.
- [13] C. Li, C. Xu, C. Gui, and M. D. Fox, "Distance regularized level set evolution and its application to image segmentation," *IEEE transactions on image processing*, vol. 19, no. 12, pp. 3243-3254, 2010.
- [14] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278-2324, 1998.
- [15] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.
- [16] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich et al., "Going deeper with convolutions." *Cvpr*, 2015.
- [17] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770-778.
- [18] J. Cheng, "brain tumor dataset," 4 2017. [Online]. Available: https://figshare.com/articles/brain_tumor_dataset/1512427
- [19] X. Yang, X. Gao, J. Li, and B. Han, "A shape-initialized and intensity-adaptive level set method for auroral oval segmentation," *Information Sciences*, vol. 277, pp. 794-807, 2014.
- [20] X. Yang, X. Gao, D. Tao, X. Li, and J. Li, "An efficient mrf embedded level set method for image segmentation," *IEEE transactions on image processing*, vol. 24, no. 1, pp. 9-21, 2015.
- [21] C. Li, R. Huang, Z. Ding, J. C. Gatenby, D. N. Metaxas, and J. C. Gore, "A level set method for image segmentation in the presence of intensity inhomogeneities with application to mri," *IEEE Transactions on Image Processing*, vol. 20, no. 7, pp. 2007-2016, 2011.
- [22] Y. Wang, L. Liu, H. Zhang, Z. Cao, and S. Lu, "Image segmentation using active contours with normally biased gvf external force," *IEEE Signal Processing Letters*, vol. 17, no. 10, pp. 875-878, 2010.
- [23] S. Mukherjee and S. T. Acton, "Region based segmentation in presence of intensity inhomogeneity using legendre polynomials," *IEEE Signal Processing Letters*, vol. 22, no. 3, pp. 298-302, 2015.
- [24] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097-1105.