ESANN'2003 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 23-25 April 2003, d-side publi., ISBN 2-930307-03-X, pp. 493-502

Digital Image Processing with Neural Networks

Axel Wismüller[1] and Udo Seiffert[2]

[1]Institut für Klinische Radiologie,
Ludwig-Maximilians-Universität München
Klinikum Innenstadt, Ziemssenstr. 1, 80336 München, Germany
e-mail: axel@wismueller.de
http://www.wismueller.de

[2]Department Cytogenetics - Pattern Recognition Group
Institute of Plant Genetics and Crop Plant Research
Corrensstr. 3, 06466 Gatersleben, Germany
e-mail: seiffert@ipk-gatersleben.de
http://www.ipk-gatersleben.de

Abstract. Real-world image processing systems frequently represent a chain of hierarchically organized, interacting components ranging from basic preprocessing to high-level image analysis and interpretation. Functional operations such as preprocessing, feature extraction, data reduction/compression, segmentation, object recognition, image understanding, and scene analysis have to be applied to different structural levels of data complexity ranging from pixel data, local features, structure and texture level data to objects, object arrangements, scene and context descriptions. Neural networks, as a special kind of learning and self-adapting data processing systems, have to offer considerable contributions to this field. Their abilities to handle noisy and high-dimensional data, nonlinear problems, large data sets etc. have led to a wide scope of successful applications in digital image processing.

Image processing is a challenge to neural network computation. As numerous application domains in science and industry are facing vast, rapidly growing amounts of digital image data, the need for advanced computer-assisted image processing and analysis techniques increasingly moves into the focus of attention. In this context, artificial neural networks, as a special kind of learning and self-adapting data processing systems, have to offer considerable contributions. Their abilities to handle noisy and high-dimensional data, nonlinear problems, large data sets etc. have lead to a wide scope of successful applications in image processing.

Typically, real-world image processing systems can be represented by a chain of hierarchically organized, interacting components ranging from basic preprocessing to high-level image analysis and interpretation, where the output obtained by each step serves as an input to the subsequent component. However, complex interactions and feedback loops between the different levels are frequently included in order to optimize the final results in the light of the specific application. In general, the various aspects of such image processing systems can be disentangled into two parallel threads:

The first thread refers to the *function* of each component: preprocessing, feature extraction, data reduction/compression, segmentation, object recognition, image understanding, and scene analysis. The second thread classifies the specific kind of *data* adopted by each component: pixel data, local features, structure and texture level data, objects, object arrangements, scene and context descriptions.

Although, at a first glance, the growing number of applications in the field may seem encouraging, there are still considerable unsolved problems: In particular, there is a need for continuous research emphasizing quality assessment including critical comparative evaluation of competing image processing algorithms with respect to specific constraints of given application domains. In this context, it increasingly becomes clear that knowledge about neural network theory alone is not sufficient for designing successful applications aiming at the solution of relevant real-world problems in image processing. What is required as well is a sound knowledge of the data, i.e. the underlying application domain. Although there may be methodological similarities, each application requires specific careful consideration with regard to algorithmic design, quality assessment, and optimization of each system component within the functional image processing 'thread' sketched above. This challenge can only be managed by close interdisciplinary cooperation of neural network theorists, engineers, and computer scientists. Hence, this subject can serve as an example for lively cross-fertilization between neural network computing and related research.

This special session is focussed on image processing based on neural networks as well as other advanced methods of computational intelligence. The contributions to this session put special emphasis on real-world applications combining original ideas and new developments with a strong theoretical background. In the following, we describe examples taken from our own research work that illustrate the applicability of neural network computation techniques to real-world digital image processing.

1. Image Analysis System for MRI Data of Patients with Multiple Sclerosis

Here, we present a complete digital image processing system for high-precision computer-assisted segmentation of multispectral MRI data sets in patients with Multiple Sclerosis which has been developed by A. W. and his group. A similar description will be published in [16]. The system comprises several interacting functional and structural components covering both unsupervised and supervised neural network learning. Therefore, it may serve as a lively example of how the conceptual power of neural machine learning techniques can contribute to the solution of advanced real-world image processing problems:

In the light of current scientific discussions on the clinical role of MRI for the evaluation of white-matter disease [4], the development of *flexible* innovative strategies for computer-assisted high-precision segmentation methods is a subject of topical interest in human brain imaging. Flexibility here refers to (i) the input, (ii) the output, and (iii) the level of human intervention required in such systems. As far as the input is concerned, the user should have the opportunity of freely choosing among different MRI sequences and various combinations thereof. As for the output, the system should not be restricted to lesion quan-

ESANN'2003 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 23-25 April 2003, d-side publi., ISBN 2-930307-03-X, pp. 493-502

tification alone, but should offer the potential to provide high-precision whole-brain or tissue-specific segmentation as well, in order to account for global brain atrophy measures, e.g. Percentage of Brain parenchyma Volume (PBV), which have recently moved into the focus of current basic and clinical research interest [5]. Finally, the system should offer different levels of human intervention: On one hand, the development and evaluation of computer-assisted segmentation systems can benefit from the superior image analysis capabilities of human beings which implies a higher degree of operator interaction. On the other hand, for large-scale clinical (e.g. multi-center) studies, however, a reduction of human intervention may sometimes be helpful in situations where user interaction could reduce reproducibility, i.e. could impose subjective bias on segmentation results. Thus, the development, test, and evaluation of a segmentation system aiming at the analysis of specific pathological changes in MS is a challenge that requires considerable effort w.r.t. integrating substantial human expertise in order to optimize computer-assisted decision support. Here we present a neural network-based segmentation system for multispectral MRI data sets of the human brain that has been specifically designed in order to provide a high degree of flexibility with regard to all three aspects mentioned above.

1.1. Methods

Data: Six patients with relapsing-remitting MS and EDSS [7] scores between 1.0 and 3.5 were included in the study. Image data were obtained on a 1.5 T MRI scanner General Electric, Signa™ employing a standardized MRI sequence protocol including *T1* and *T2* weighted, *Proton Density* (PD) weighted, *Fluid-attenuated Inversion-Recovery* (FLAIR), and *Magnetization Transfer* (MT) sequences in axial slice orientation. The T1 and MT sequences were repeated after intravenous contrast agent administration. Total scanning time was 27.4 min.

Image Analysis: The conceptual basis of single components of our system has been described in [17]. Here, we want to put special emphasis on the functional interplay between the various components in so far as it is relevant to brain segmentation in MS. An overview of the segmentation system is shown in fig. 1. Thick-lined boxes indicate interactive steps. Boxes with rounded corners refer to segmentation results. After co-registration and gray level rescaling (1) of the input data, the intracranial cavity (ICC) is pre-segmented interactively. For the data presented here, this step was performed manually by human expert readers, however, (semi-)automatic techniques may be used as well, such as the methods developed by our group [15]. In a second step, a training data set is obtained manually comprising small reference regions labeled as “Gray Matter (GM)”, “White Matter (WM)”, “Cerebrospinal Fluid (CSF)”, “White Matter Lesion (WML)”, and a “Residual Class (RC)”, representing other tissues such as meninges or larger vessels (2). Subsequently, gray level shift effects induced by magnetic field inhomogeneities and cross-talk effects can be corrected using the training data and the ICC masks (3, 4). For this purpose, we have developed a specific bootstrap algorithm based on iterative improvement of a preliminary neural network tissue classification, which will be published elsewhere. After these preprocessing steps, each voxel within the ICC mask is assigned to a feature vector $\mathbf{x}$ representing its MRI signal intensity spectrum. This set

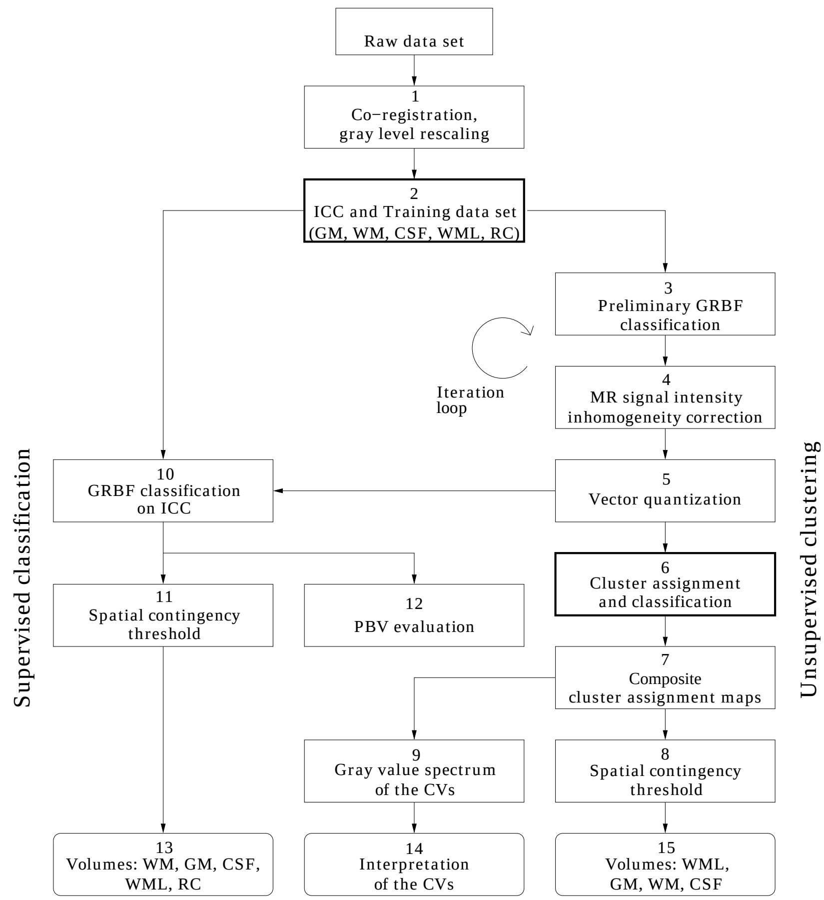
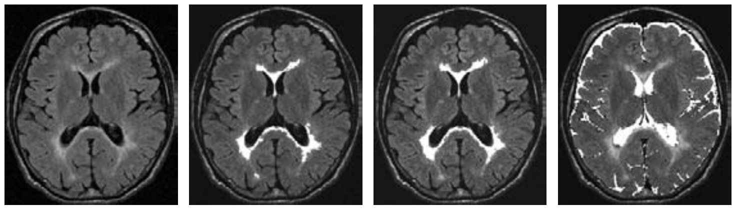


Figure 1: The image analysis system for high-precision segmentation of multi-spectral MRI data in patients with multiple sclerosis.

of feature vectors is partitioned into N clusters by unsupervised learning (5) based on minimal free vector quantization [17]. The resulting codebook can either be used for interactive visual tissue type classification based on cluster assignment maps (6), or automatic supervised segmentation can be obtained by subsequent training of a Generalized Radial-Basis Functions- (GRBF-) neural network (10), see [17].

For the interactive visual classification of cluster assignment maps, we developed a software system named CASCADE (Computer-Assisted Cluster Assignment Decision Environment) which enables quick and efficient screening of cluster assignment maps and underlying MRI data. Here, each feature vector $\mathbf{x}$ is uniquely attributed to its closest codebook vector $\mathbf{w}_j(\mathbf{x})$ according to a minimal distance criterion in the gray level feature space, and corresponding cluster assignment maps (6) are constructed for visual inspection. In a second step, each cluster j belonging to codebook vector $\mathbf{w}_j$ is interactively assigned to

ESANN'2003 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 23-25 April 2003, d-side publi., ISBN 2-930307-03-X, pp. 493-502



(a) (b) (c) (d)

Figure 2: (a) Axial slice (FLAIR MRI sequence) of a brain containing WML; (b) WML classification based on interactive cluster assignment using the CASCADE system; (c) supervised automatic WML classification using a GRBF neural network; (d) CSF segmentation by GRBF neural network classification. For explanation, see text.

a specific tissue class $\lambda \in \{0, \ldots, m\}$ by a human expert reader. Finally, all the clusters assigned to each specific tissue class λ are collected and merged yielding a composite cluster assignment map (7) representing the final segmentation result. Based on this tissue assignment, the isolated or merged codebook vectors representing prototypical gray level spectra may be plotted for further visual analysis and interpretation (9, 14). Finally, segmentation results may be used for tissue-specific volume measurements (15), where spatial smoothing techniques or geometric contingency thresholding (8) can be employed as optional post-processing steps. An example for WML segmentation results is presented in fig. 2b.

Alternatively, for automatic supervised classification by a GRBF neural network (10) the training data from step (2) and the resulting codebook from step (5) can be re-cycled [17]. Based on the respective tissue segmentation, the WML volume can be quantified as well (11, 13 – see fig. 2c). Furthermore, the GRBF segmentation approach can be used for PBV calculation based on automatic CSF identification (12 – see fig. 2d).

Table 1: Statistical analysis of WML and PBV quantification methods w.r.t. inter-observer agreement (univariate F-test, $N = 6$). The method yielding better results, i.e. higher inter-observer agreement is printed in bold face for each pairwise comparison.

Method A	Method B	p-value
	WML segmentation	
Region Growing	**GRBF**	0.075
Region Growing	**CASCADE**	0.029
GRBF	**CASCADE**	0.027
	PBV calculation	
GRBF	Angle Image	0.003

ESANN'2003 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 23-25 April 2003, d-side publi., ISBN 2-930307-03-X, pp. 493-502

1.2. Evaluation and Results

In order to perform a thorough quantitative evaluation of the described pre-processing and segmentation procedures w.r.t. all data processing steps involving human interaction, WML quantification and PBV computation were performed based on (i) interactive definition of training data sets by two different observers independently for supervised GRBF classification of WML and PBV computation, respectively, (ii) interactive reference region contour tracing for threshold definition of an observer-guided region growing technique [14] serving as a reference method, by two different observers independently, (iii) interactive cluster assignment using the CASCADE system by two different observers independently, and (iv) interactive threshold definition for the angle image method [5] serving as a reference method for PBV computation, by two different observers independently. The computation of inter-observer agreement levels was performed according to the statistical guidelines of the British Standards Institution [1].

In order to rank the methods w.r.t. their segmentation quality, the inter-observer agreements of CASCADE and GRBF neural network segmentation were compared to region growing, based on a univariate F-test. From the results presented in tab. 1, it can be concluded that (i) the mean inter-observer agreement in cluster assignment using the CASCADE segmentation procedure is higher than in both region growing ($p = 0.029$) and GRBF neural network classification ($p = 0.027$), i.e. there is a significant method effect; (ii) the mean inter-observer agreement in GRBF neural network classification is higher than in threshold-based region growing. However, statistical analysis reveals only a method effect of reduced significance for the comparison of GRBF neural network segmentation and region growing ($p = 0.075$). In conclusion, interactive cluster assignment using the CASCADE segmentation system performs significantly best in a comparison of the three methods, whereas supervised GRBF neural network classification is slightly better than conventional region growing serving as a reference method for WML quantification. For PBV computation, our GRBF neural network method outperforms the reference angle image technique w.r.t. inter-observer agreement at a significance level of $p = 0.003$.

1.3. Discussion

For WML quantification in MS we obtain the best segmentation results using the CASCADE approach, where human expert knowledge is incorporated at a “cluster level” instead of a “pixel level”, i.e. at an advanced, abstract level of knowledge representation within the pattern recognition process. We conjecture that this observation could be of particular interest in the light of ongoing discussions on “domain knowledge data fusion for decision support” in the machine learning community. Our study shows that computer-assisted image analysis using semi-automatic neural network segmentation outperforms conventional threshold-based techniques w.r.t. inter-observer agreement levels for both WML quantification and PBV calculation in MRI data of MS patients. At the same time, our segmentation system allows the radiologist and neuro-scientist to choose freely among different input MRI sequences and various combinations thereof in order to systematically explore their contribution to brain imaging in MS.

ESANN'2003 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 23-25 April 2003, d-side publi., ISBN 2-930307-03-X, pp. 493-502

2. Adaptive Image Compression within High-Throughput Screening using Auto-Associative Feed-Forward Networks

2.1. Motivation

Many High-Throughput Screening (HTS) methods [6] being used in almost all scientific fields, from life science to engineering, often lead to an immense quantity of data. Besides the actual processing of these data, which depends on the specific task of course, there is usually a problem of transmitting and storing these data. Especially when dealing with image based methods, existing problems of data handling may limit the throughput of the entire system. Assuming the number of images and their spatial and/or temporal resolution cannot be further optimized anymore, image compression algorithms [9], [13], [8] are frequently applied. As long as there are no strong demands for the details of image content on pixel level, lossy compression methods, i.e. JPEG (Joint Photographic Experts Group) [10], are appropriate. These methods are fast, rather effective and scalable. Thats why JPEG is the standard image format in the internet. On the other hand, there are lossless compression algorithms, i.e. LZW (Lempel-Ziv Welch) Compression or JPEG2000 [11], where the image is completely restored after compression/decompression. Depending on the image contents, these algorithms frequently are not as effective as many lossy compression methods, since information contained in the image is just rearranged or transformed but not reduced. For details see [13]. The dualism of lossy and lossless compression seems to fix a dilemma: either a high compression ratio with loss of possibly important information or keeping all information at a poor compression ratio. The reason for this dilemma is that all these methods process all image content the same way. Using an adaptive procedure which respects the image content can be built up using auto-associative feed-forward neural networks (Bottleneck neural networks) as shown in fig. 2.

2.2. Description of the Algorithm

This technology is not new. First applications date back to the late 80s/early 90s [3], [2], [12], but we can now notice a revival against the background of increasing HTS applications. During the learning phase of the neural network, relatively small blocks, usually 8x8 pixels, of images with typical content are presented to the net. At the end the auto-associative net is able to reflect the input data at its output with only a very small reconstruction error. Since the hidden layer is typically smaller than input/output layers, a compression (from input to hidden layer) and the inverse decompression (from hidden to output layer) is performed by the neural network. When the network is recalled now with previously unknown images respectively their blocks, the activation of the hidden layer, which was from the neural networks point of view just an intermediate result, is considered as the compressed image block. In order to restore the image, all compressed blocks are processed by the output layer of the network which effectively performs the decompression, and are reassembled in the correct order to form the reconstructed image. This procedure has been applied as a general purpose method for image compression with discouraging results. However, it has proven to work very well in those cases, where all images to

ESANN'2003 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 23-25 April 2003, d-side publi., ISBN 2-930307-03-X, pp. 493-502

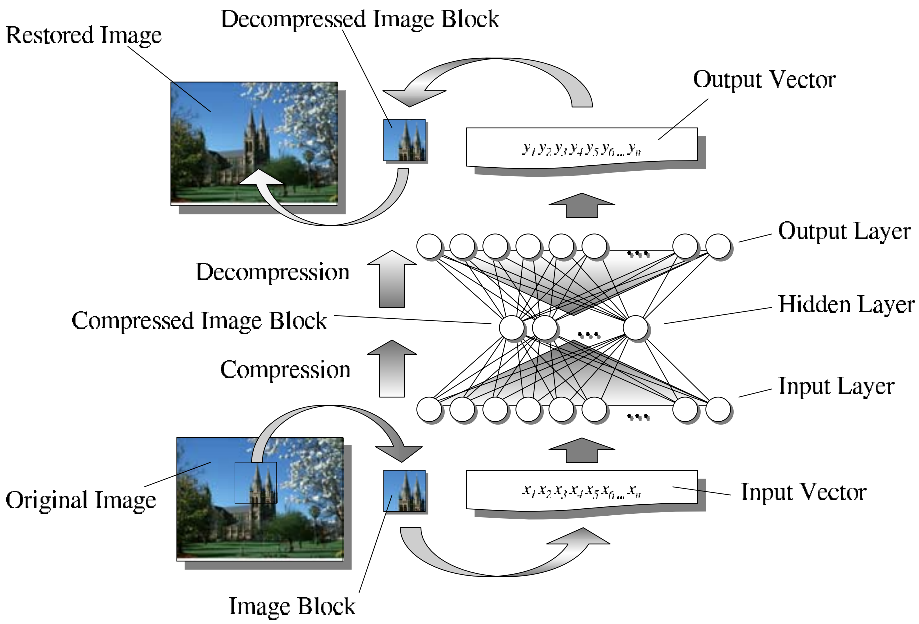


Figure 3: The original image is divided into equally sized blocks $(b_x \cdot b_y)$. All blocks are consecutively transformed into an input vector of length $n = b_x \cdot b_y$. The hidden layer is built up of $m < n$ neurons (m/n denotes the compression ratio) and contains the compressed image block.

be processed contain similar information, i.e. always a typical basic structure with small but important differences which would be equalized by conventional (non-adaptive) lossy compression methods. Once the neural network is trained with a set of typical and similar images, it can be used without retraining as long as the images do not change their basic features. Since this neural solution is rather slow, even in the recall phase, compared to the above mentioned compression methods, it is not suitable for on-line compression/decompression. However, in all applications with no hard schedule, such as storing/archiving images for later processing or occasional reference, it is a promising alternative to conventional compression methods.

Acknowledgments

We express our gratitude to Johannes Behrends, Oliver Lange, Helge Ritter, Frank Vietze, Gerda L. Leinsinger, Miriana Jukic, Klaus Hahn, Maximilian Reiser, and Dorothee Auer. MRI data in section 1. have been contributed by Dorothee Auer from the Max Planck Institute of Psychiatry in Munich.

References

[1] British Standards Institution, BS 5497. *British Standards. Precision of test methods. Part I. Guide for the determination of "reproducibility" for a standard test method by inter-laboratory tests*, 1987.

[2] S. Carrato. In *Neural Networks: Advances and Applications 2*. Amsterdam.

[3] G.W. Cottrell, P. Munro, and D. Zipser. In *Models of Cognition: A Review of Cognition Science*. Norwood, NJ.

[4] M. Filippi, M. Horsfield, P. Tofts, F. Barkhof, A. Thompson, and D. Miller. Quantitative assessment of MRI lesion load in monitoring the evolution of multiple sclerosis. *Brain*, 118:1601–1612, 1995.

[5] Y. Ge, R. I. Grossman, J. K. Udupa, L. Wei, L. J. Mannon, M. Polanski, and D. Kolson. Brain athrophy in relapsing-remitting multiple sclerosis and secondary progressive multiple sclerosis: longitudinal quantitytive analysis. *Radiology*, 214:665–670, 2000.

[6] W. P. Janzen. *High Throughput Screening: Methods and Protocols*. Humana Press, Totowa, NJ, 2002.

[7] J. F. Kurtzke. Rating neurologic impairment in multiple sclerosis: An expanded disability status scale EDSS. *Neurology*, 1983.

[8] J. Miano. *Compressed Image File Formats: JPEG, PNG, GIF, XBM, BMP*. Benjamin Cummings / Addison Wesley, San Francisco, Ca, 2002.

[9] J. D. Murray, W. Vanryper, and D. Russell. *Encyclopedia of Graphics File Formats*. O'Reilly UK, Cambridge, 1996.

[10] W. B. Pennebaker and J. L. Mitchell. *Jpeg: Still Image Data Compression Standard*. Kluwer International, Dordrecht, 1992.

[11] D. S. Taubman and M. W. Marcellin. *JPEG 2000: Image Compression Fundamentals, Standards and Practice*. Kluwer International, Dordrecht, 2000.

[12] L. Wang and E. Oja. Image compression by mlp and pca neural networks. In *8. Scandinavian Conference on Image Analysis*, pages 1317–1324, 1993.

[13] P. Wayner. *Compression Algorithms for Real Programmers*. Morgan Kaufman, San Francisco, Ca, 1999.

[14] D. Wicks, P. Tofts, D. Miller, C. du Boulay, A. Feinstein, and R. Sacares. Volume measurement of multiple sclerosis lesions with magnetic resonance images: a preliminary study. *Neuroradiology*, 34:475–479, 1992.

[15] A. Wismüller, J. Behrends, O. Lange, D.R. Dersch, G.L. Leinsinger, F. Vietze, and K. Hahn. Automatic segmentation of cerebral contours in multispectral MRI data sets of the human brain by self-organizing neural networks. *Radiology*, 221(P):461, 2001.

[16] A. Wismüller, J. Behrends, O. Lange, M. Jukic, K. Hahn, M. Reiser, and D. Auer. High-precision computer-assisted segmentation of multispectral MRI data sets in patients with multiple sclerosis by a flexible machine learning image analysis approach. In *Proceedings of the Workshop 'Bildverarbeitung für die Medizin' 2003*, Berlin, Heidelberg, New York, 2003. Springer-Verlag. In press.

[17] A. Wismüller, F. Vietze, and D.R. Dersch. Segmentation with neural networks. In Isaac Bankman, Raj Rangayyan, Alan Evans, Roger Woods, Elliot Fishman, and Henry Huang, editors, *Handbook of Medical Imaging*, pages 107–126, Johns Hopkins University, Baltimore, 2000. Academic Press. ISBN 0120777908.