ESANN'2002 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 24-26 April 2002, d-side publi., ISBN 2-930307-02-1, pp. 25-38

Exploratory Data Analysis in Medicine and Bioinformatics

Axel Wismüller[1] and Thomas Villmann[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]Klinik für Psychotherapie,
Universität Leipzig
Karl-Tauchnitz-Str. 25, 04107 Leipzig, Germany
e-mail: villmann@informatik.uni-leipzig.de

Abstract. Biomedical research is a challenge to neural network computation. As medical doctors and bioscientists are facing vast, rapidly growing amounts of data, the need for advanced exploratory data 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 led to a wide scope of successful applications in biomedicine.

1. Neural Networks in Biomedicine

Biomedical research is an important application domain for neural computation techniques. Vice versa, neural network modeling is influenced by concepts from experimental and theoretical biology. This cross-fertilization aspect has been a continuous source of inspiration since the early days of biomedical neural network computing (e.g. [38],[44],[25],[26]).

A major focus in exploratory data analysis, data mining, and knowledge discovery is to visualize high-dimensional data and to analyze its intrinsic structure. In the biomedical domain, respective methods should provide tools for data assessment and diagnostic decision support for the domain experts, e.g. medical doctors or bioscientists. In this context, neural networks can be utilized as powerful instruments providing techniques for adaptive and, in particular, non-linear data processing. The latter aspect makes an important difference in comparison to conventional methods in linear statistics, such as linear regression, factor analysis, principal component analysis etc.

For exploratory data analysis, unsupervised neural network *clustering* or Vector Quantization (VQ) algorithms have been applied successfully to a wide scope of biomedical data analysis problems. VQ procedures map a data space onto a finite set of prototypical feature vectors, a so-called codebook. Examples of this class of algorithms are elementary VQ methods such as LBG [15] and

k-means clustering [1],[5],[8] or more refined, 'soft-competing' algorithms such as Kohonen's Self-Organizing Maps (SOMs) [14], minimal free energy VQ [20], [4], or the 'neural gas' algorithm [16].

One of the most powerful neural network techniques in this context is the SOM which provides a robust method to visualize essential properties of the data [41]. Under certain conditions, it represents a topographic mapping of high-dimensional input data onto a low-dimensional output space usually chosen as a hypercube. Various extensions for the basic SOM are known for faithful data modeling and analysis. One of the most important is the Growing SOM (GSOM) approach which allows a successive adaptation of the output space structure to prevent violations of topography. In fact, the GSOM realizes a *non-linear principal component analysis (PCA)*. A second feature is the control of magnification induced by the SOM which is closely related to the problem of optimal information transfer [19],[55]. A magnification control scheme can be implemented by *local* learning [35]. Other extensions relate to learning using auxiliary data [47], probability density estimation [42], or kernel methods [37].

Once the training of a SOM is completed, further data analysis is possible using the information already acquired by the network. In particular, if the topography for a given SOM is proven, one is able to investigate the low-dimensional model instead of the original data. For this purpose, conventional clustering methods such as Ward-clustering or single linkage clustering can be applied to SOM [48],[53]. For a comprehensive overview we refer to [48].

Further methodological achievements have emerged as valuable extensions to the classical SOM, such as the Double SOM [21] or the Deformable Feature Map [30]. These advances have not only been shown to be conceptually interesting, but have proven their practical applicability in biomedicine, such as medical image segmentation and registration by the Deformable Feature Map [27], gene expression data analysis by the Double SOM [54], or the use of adaptive metrices based on auxiliary data [40].

In the following, we sketch several examples taken from our own research work that illustrate the applicability of both supervised and unsupervised neural network computation techniques to a wide range of biomedical applications.

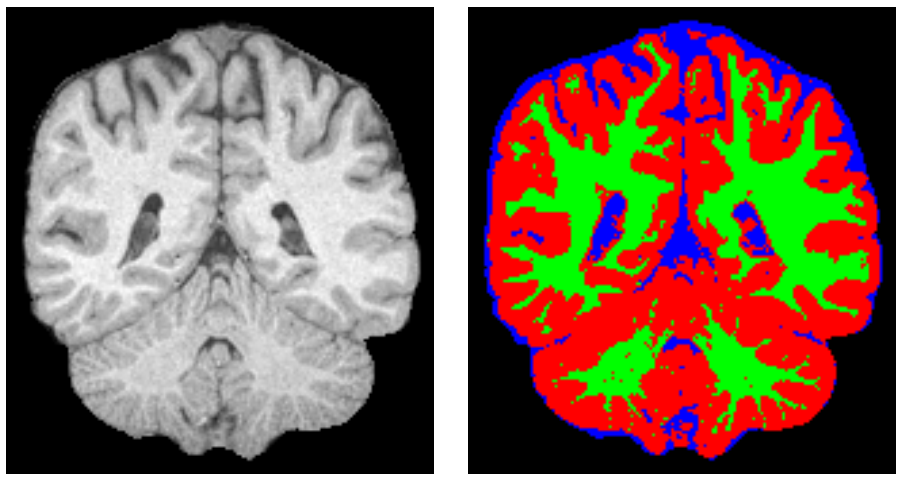
2. Biomedical Image Analysis

2.1. Segmentation of Multispectral MRI Data Sets

A classical problem of neural network computing in biomedicine is image segmentation. In [29], an algorithmic approach has been presented that aims to combine *Unsupervised Clustering* (UC) and *Supervised Classification* (SC) for image segmentation, where the information obtained during UC is not discarded, but is used as an initial step towards subsequent SC. Thus, the power of both image analysis strategies can be combined in an integrative computational procedure. This is achieved by applying Generalized Radial-Basis-Functions-(GRBF-) neural networks [10],[18].

As an example, we sketch the segmentation of multispectral 3D MRI data sets of the human brain consisting of 4 different MRI acquisition sequences with respect to the tissue classes "gray matter", "white matter" and "cerebrospinal fluid", which is a classical problem of brain image analysis with a wide scope of relevant clinical applications in neurology, psychiatry, and neuroradiology.

ESANN'2002 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 24-26 April 2002, d-side publi., ISBN 2-930307-02-1, pp. 25-38



(a) (b)

Figure 1: Segmentation. (a) *T1* weighted image of a 3D data set; (b) Corresponding segmented image with gray-level representation of the tissue classes (medium gray $\hat{=}$ "gray matter", light gray $\hat{=}$ "white matter", dark gray $\hat{=}$ "cerebrospinal fluid").

After correct anatomical alignment of the 4 data sets, the relevant region of interest, i.e. the brain, has to be extracted, see below. Finally, for each voxel, we obtain a 4-dimensional feature vector representing the signal intensities of the different MRI sequences. In [29], UC image segmentation is performed by minimal free energy VQ [20],[4], enabling unsupervised exploratory analysis of the image data on different scales of feature space resolution. For subsequent SC analysis, the UC results are not discarded, but can be re-utilized for the training of a GRBF neural network. Fig. 1 presents typical segmentation results for the brain tissue classification problem sketched above.

Meanwhile, this segmentation approach has shown to be useful for a wide scope of clinical applications ranging from plaque characterization in atherosclerosis (fig. 2) to high-precision segmentation and volumetry of white-matter lesions in demyelinating diseases such as multiple sclerosis [24].

2.2. The Deformable Feature Map

In biomedical pattern analysis, an important issue is to exploit apparent similarities of data sets when comparing different, but similar objects. In biomedical research data sets, this phenomenon can be observed frequently (see e.g. [26]). For example, one may think of the interindividual variability of anatomical features: there are no completely identical biological individuals, but there may be obvious anatomical "resemblances" (see e.g. the brains of two different individuals in fig. 4 a,b). Thus the question arises whether neural network training should be repeated for each data set separately, or knowledge could be re-utilized for new data sets in order to reduce the expense w.r.t. computation and human intervention time.

ESANN'2002 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 24-26 April 2002, d-side publi., ISBN 2-930307-02-1, pp. 25-38

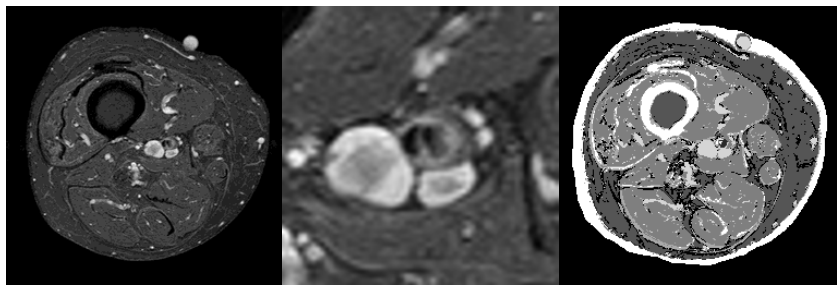


Figure 2: Tissue classification based on multispectral MRI data: Left: *T1* weighted MRI cross-section of the thigh of a patient with atherosclerosis. The region of the femoral vessels is marked. Middle: Magnification of the vessel region as indicated on the left image. Right: Tissue segmentation based on minimal free energy VQ of the gray level spectra enabling a distinction between different tissue classes such as fat, muscle, fibrotic tissue etc.

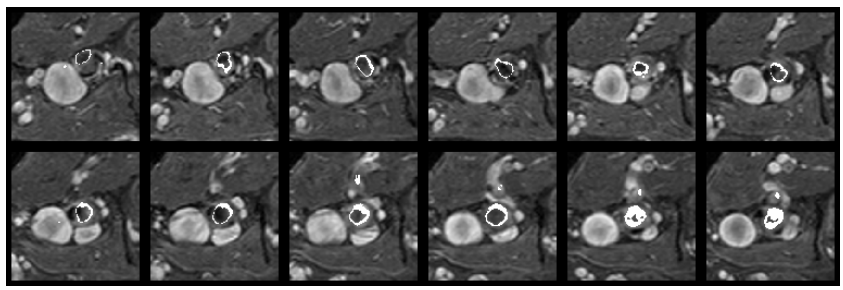


Figure 3: Cluster specializing on voxels representing the arterial vessel wall (A. femoralis superficialis) indicated by highlighted pixels. The images represent subsequent cross-sections of the vessel region (such as indicated by fig. 2 (middle)) at different levels. The lumen of the vessel is reduced by regional plaque-induced thickening of the vessel wall in the last few images.

Motivated by the efforts for optimizing the image analysis system sketched above, an algorithm has been discovered as a conceptual extension of self-organizing maps that provides adaptive plasticity in function approximation problems: the Deformable Feature Map [31],[30]. This approach reduces a class of similar function approximation problems to the explicit supervised one-shot training of a *single* data set. This is followed by a subsequent, appropriate similarity transformation which is based on a self-organized deformation of the underlying multidimensional probability distributions.

Applications of this algorithm have been presented to both automatic non-linear image registration [30] and segmentation [31],[32], see also fig. 4.

2.3. Conceptual Extensions and Additional Applications

Several conceptual extensions of the image analysis system comprise:

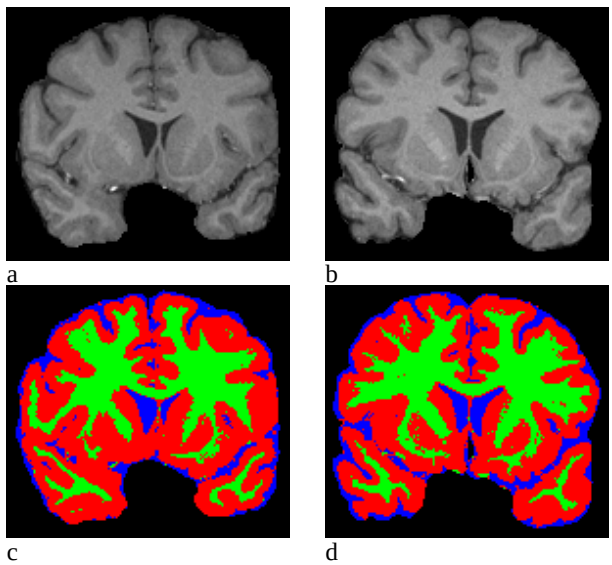


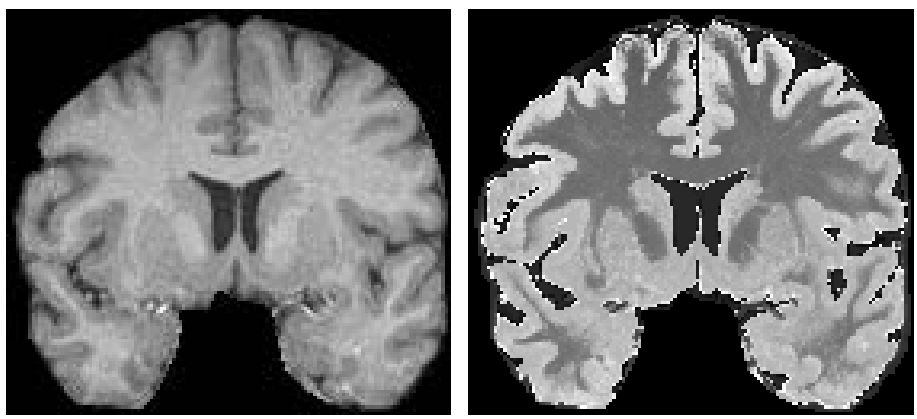
Figure 4: Results of fully automatic segmentation of multispectral magnetic resonance (MR) imaging data sets of the human brain using the Deformable Feature Map approach. The upper line (a,b) shows so-called T1-weighted MR images. The lower line (c,d) shows the corresponding segmentations with respect to three classes "white matter" (light gray), "gray matter" (middle gray), and "cerebrospinal fluid" (dark gray). The images of the left column (a,c) belong to an individual Y, the images of the right column (b,d) belong to a different individual X. The segmentation of Y (c) served as a reference data set for a fully automatic segmentation of X, shown in (d). From [31].

Automatic pre-segmentation: As mentioned above, the relevant regions of interest have to be extracted prior to segmentation. As manual contour tracing by human observers is very time-consuming, this is an important issue for the practical applicability of the system. For this purpose, a neural network based pre-segmentation system for the definition of brain contours in multispectral MRI data sets was developed and evaluated [23]. Image data were represented in a 63-dimensional feature space consisting of 3 spatial and 60 gray level coordinates of each voxel and its neighborhood. The segmentation quality achieved by GRBF neural network classification was comparable with respect to results obtained by human observers: The variability between manual and automatic contour definition was in the range of the inter-observer variability of different human expert readers.

Multispectral image synthesis: For some applications, it is useful to compress the image information of multispectral data into a single data set. For this purpose, a method called SOMSIS (Self-organized Multispectral Image Synthesis) was developed that performs a nonlinear principal component analysis by data projection onto a parameterized one-dimensional SOM. An example is presented in fig. 5. The method has shown to be helpful as a preprocessing step to nonlinear image registration of multispectral MRI data of

the human brain by the Deformable Feature Map, see above [30].

Analysis of phonation using MRI: For the evaluation of the functional anatomy of the oro-pharyngeal cavity during phonation based on MRI data sets, a conceptually similar data analysis technique was applied: The goal was to define a nonlinear principal component of the vocal tract in order to define cross-sectional "area functions" that play an important role in speech production modeling. For this purpose, healthy professionally trained volunteer speakers performed a prolonged emission of sounds of the German phonemic inventory during MRI acquisition. After semi-automatic segmentation of the vocal tract, the nonlinear principal component was determined by a one-dimensional SOM. In contrast to the SOMSIS application sketched above, spatial coordinates are used instead of signal intensity coordinates. An example is presented in fig. 6.



(a) (b)

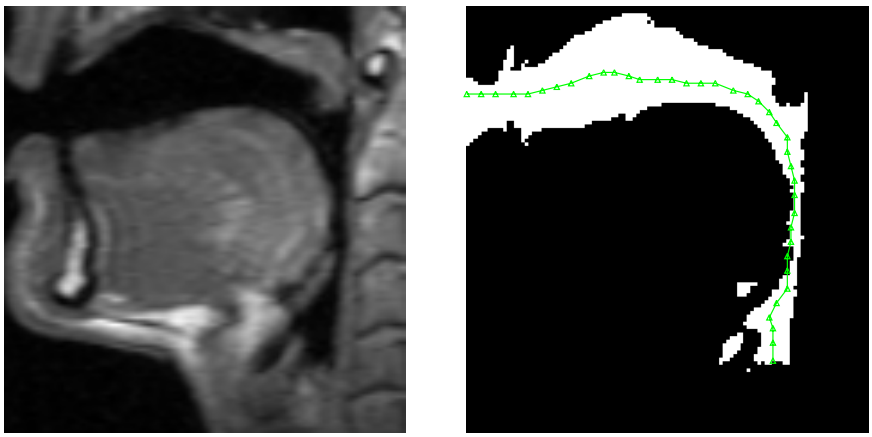
Figure 5: Self-organized multispectral image synthesis (SOMSIS). (a) $T1$-weighted image from a multispectral MRI data set composed of 4 MRI acquisition sequences; (b) Compressed image using the SOMSIS projection.

3. Time-Series Analysis

3.1. Image Time-Series

Besides the classical domains of biomedical time-signal analysis such as EKG or EEG processing, the analysis of biomedical image time-series has become an issue of growing importance for both basic research and clinical application. Neural network clustering by deterministic annealing has been presented as a powerful strategy for self-organized functional segmentation of biomedical image time-series data identifying groups of pixels sharing common properties of local signal dynamics [28]. Successful applications include (i) functional MRI data analysis for human brain mapping, (ii) dynamic contrast-enhanced perfusion MRI for the diagnosis of cerebrovascular disease, and (iii) magnetic resonance mammography for the analysis of suspicious lesions in patients with

ESANN'2002 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 24-26 April 2002, d-side publi., ISBN 2-930307-02-1, pp. 25-38



(a) (b)

Figure 6: Evaluation of the functional anatomy of the oro-pharyngeal cavity during phonation using MRI: (a) Midsagittal MRI section of the vocal tract. (b) Segmented vocal tract with nonlinear principal component as defined by a 1D-SOM.

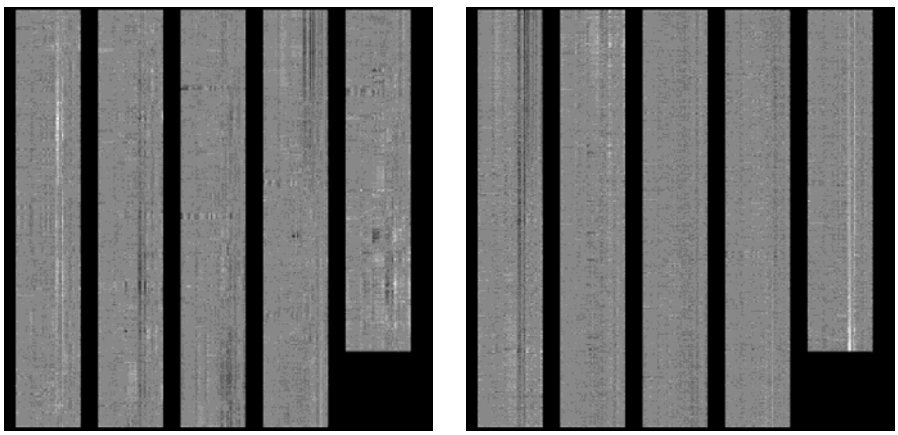
breast cancer. Alternative approaches to exploratory biomedical image time-series analysis range from k-means type clustering (e.g. [2]) and SOMs (e.g. [7]) to global optimization techniques (e.g. [9]). For further bibliographical data and details referring to the applications mentioned above, we refer to [28].

3.2. Functional Genomics and Bioinformatics

Exploratory analysis of gene expression profiles has emerged as an important field of bioinformatics. For data partitioning and visualization, various clustering approaches have been proposed in the literature as such as agglomerative clustering [6] or several variants of 2D-SOMs [22],[39]. An application of the Double SOM [21] to gene expression clustering is presented in this volume [54].

An alternative approach for data visualization in functional genomics is proposed in fig. 7. It is based on a one-dimensional parameterized SOM that enables fast and convenient projection of high-dimensional microarray gene expression profiles onto a one-dimensional manifold. In contrast to the 2D-SOM approaches cited above, this provides a graphical output intuitive for biologists, where the gene expression profiles are ordered along a linear list of subsequent entries in analogy to the classical agglomerative clustering approach by Eisen et al. [6]. However, the linear arrangement here is solely data-driven by self-organized alignment, i.e. it does not require meta-knowledge such as chromosomal position or heuristic ordering criteria e.g. average expression level or time of maximal induction used in [6] in order to enforce graphically appealing visualization. Fig. 7 (a) shows the yeast genome data set published in [6] in the original alignment based on agglomerative clustering, heuristic criteria and meta-knowledge as mentioned above. Fig. 7 (b) presents the self-organized alignment obtained by the one-dimensional parameterized SOM. In contrast to (a), discontinuities between neighboring expression profiles are markedly

ESANN'2002 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 24-26 April 2002, d-side publi., ISBN 2-930307-02-1, pp. 25-38



(a) (b)

Figure 7: Analysis of microarray gene expression profiles of the budding yeast *saccharomyces cerevisiae*. The data published in [6] consist of 2479 79-dimensional gene expression profiles. (a) Original order proposed by [6]. (b) Order induced by topology-preserving mapping using a parameterized 1D-SOM as described in the text.

reduced.

4. Clustering Based on PET Image Features

The image analysis applications discussed above are focussed on local features such as signal intensity or geometrical position. However, there are successful applications of exploratory data analysis based on high-level features extracted from biomedical image data as well. For illustration we consider the analysis of *Positron- Emission-Tomography* (PET) data of patients suffering from Wilson's disease (WD). WD is a rare autosomal-recessive disorder of copper metabolism which shows disturbances in liver function and basal ganglia leading to hepatic and extra-pyramidal motoric symptoms [46]. PET represents an advanced nuclear medicine imaging technique which allows non-invasive functional evaluation of metabolic processes which was applied here to detect glucose metabolism in the brain. Relative glucose consumption [%] is calculated for three-dimensional regions of interest. Thus, *five*-dimensional data vectors (components of thalamic region, putamen, caput nuclei caudati, cerebellum and midbrain) are obtained for each subject. Fig. 8 gives paradigmatic examples of the PET images in three WD patients. In addition, the severity of the neurological symptoms was evaluated using a clinical scoring system. Application of GSOM to the PET data generates a two-dimensional output space. The distribution of component values of the input dimensions according to the generated output space is depicted in fig. 9. A subsequent SOM-Ward-clustering [53] yields a three-cluster solution depicted in fig. 9. A deeper medical analysis of the cluster solution shows that normal subjects and patients with non-neurological WD are mainly attributed to cluster II, a high percentage of

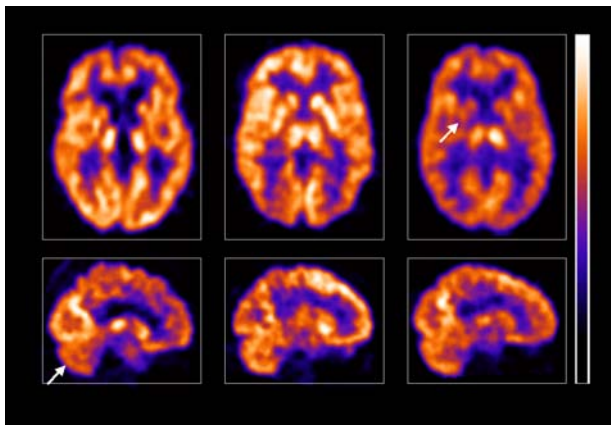


Figure 8: Paradigmatic PET Images (transverse slices in upper row and sagittal slices in lower row) of 3 patients with WD. SOM analysis revealed three patterns of glucose metabolism: normal (middle); decreased in stem brain which consists of cerebellum (white arrow) and midbrain (left column); decreased in striatothalamic areas (right column, white arrow in striatum).

the patients with pseudo-parkinsonian type of disease (TOD) is represented by cluster I. The patients with pseudo-sclerosis TOD, however, are not unequivocally attributed (cluster II (54%), cluster III (46%)). Arrhythmic-hyperkinetic TOD, which represents a mixture of symptoms of the former two types, correspondingly fits to cluster I and III. Comparing this analysis with the pattern of glucose consumption (fig. 9, component planes), cluster II can be assigned as "*normal*", cluster I as "*deficit in thalamostriatal areas*" and cluster III as "*deficit in stem brain* (cerebellum and midbrain)". With respect to the WD patients, scoring of the severity of neurological symptoms was 0.5 ± 1.1 in cluster II, 2.1 ± 1.5 in cluster I and 0.6 ± 0.9 in cluster III ($p = 0.002$ after one-factorial ANOVA). From the results can be concluded that the GSOM analysis was able to initially identify *three different patterns of brain glucose metabolism in WD* [51].

5. Analysis of Non-metric Data

The investigation and analysis of non-metric data frequently occurs in medical applications and social sciences. Therefore, a large set of methods based on classical statistics exist. Besides these traditional methods new approaches were developed in the community of soft-computing as well. Frequently, the innovative procedures provide the advantage of greater robustness and adaptivity with the possibility of online learning if new data are available.

A frequent problem is clustering of data only based on similarities. For this purpose neural network approaches were developed by HOFMANN&BUHMANN (Pairwise Data Clustering by Deterministic Annealing – PDCDA) [12] and GRAEPEL&OBERMAYER (Stochastic Self-Organizing Map for Proximity Data – SSMPD) [11]. Both approaches are based on a stochastic gradient descent

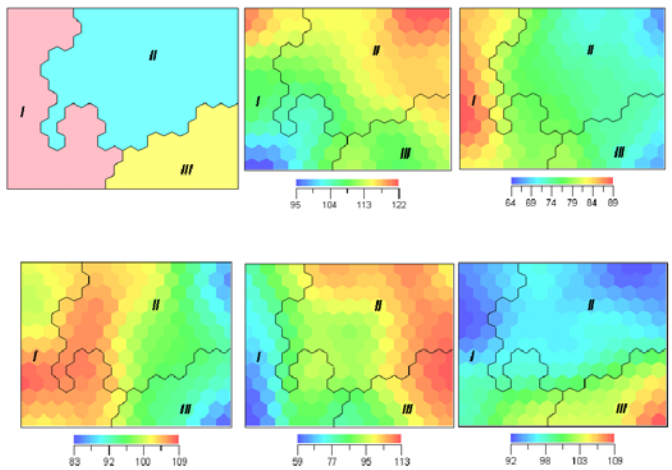


Figure 9: GSOM-analysis of the WD-probands and resulting cluster solution (upper row - left) together with the component planes of the 5 input variables of the GSOM training: components for thalamic region, putamen, caput nuclei caudati, cerebellum and midbrain (see text).

of an energy function. The advantage of the latter one is the incorporation of neighborhood learning into the adaptation process (adapted from neural maps) which should avoid local minima [36]. However, due to critical phase transitions during the adaptation process, a faithful parameter regime for the involved annealing strategies is needed [11]. On the other hand, these energy functions could serve as a fitness function in an evolutionary algorithm (EA) approach which is more robust. Furthermore, EAs are easy to implement and, hence, practicable without deep insight. However, in EAs there is no guaranty to find the optimum. Therefore, only large scaled EA applications with advanced strategies (special genetic operators [13], subpopulation approaches [3],[50], selection strategies [17]) are a real alternative to the mentioned neural network approaches. A further important advantage of EAs is that additional knowledge can be included for assessment. This aspect frequently is of special interest in medical applications where some external information is available.

As an example study, we perform a rough comparison of the methods in a practical application: One of the mostly used methods for the acquisition of structures of interpersonal relationships in the area of psycho-dynamic psychotherapy research is the method of the *'Core Conflictual Relationship Theme'* (CCRT) developed by LUBORSKY [43] investigating so-called *relationship-episodes*. In each of such episodes the components *wish of the subject* (W), *response of the object* (RO) and *response of the subject* (RS) were encoded which are taken to perform the CCRT using a classifier system of categories for each component. Yet, the categories are often correlated in meaning. Therefore, they are collected in a set of clusters [34] which are then used in further considerations instead of the categories. The number and the interpretation of the clusters as well as the assignment of the categories result from the ex-

ESANN'2002 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 24-26 April 2002, d-side publi., ISBN 2-930307-02-1, pp. 25-38

$\tilde{\kappa}$–coefficient	meaning
$\tilde{\kappa} < 0.1$	no agreement
$0.1 \leq \tilde{\kappa} < 0.4$	weak agreement
$0.4 \leq \tilde{\kappa} < 0.6$	clear agreement
$0.6 \leq \tilde{\kappa} < 0.8$	strong agreement
$0.8 \leq \tilde{\kappa}$	nearly complete agreement.

Table 1: Different values for the weighted concordance coefficient $\tilde{\kappa}$ and the respective meaning for intra-cluster agreements of the considered observables

database	$\tilde{\kappa}$ original clusters	$\tilde{\kappa}$ according PDCDA	$\tilde{\kappa}$ according SSMPD	$\tilde{\kappa}$ according EA
P^W	0.334	0.406	0.421	0.397
P^{RO}	0.323	0.427	0.466	0.423
P^{RS}	0.479	0.511	0.527	0.513

Table 2: Weighted concordance coefficients $\tilde{\kappa}$ for the original clusters and the improved cluster solutions.

periences of several psychotherapists which are subsequently evaluated using conventional statistical methods. This cluster solution should be improved using the above mentioned techniques. For clustering we used the conditional probabilities of coupled occurence of the respective categories in large data bases P^W, P^{RO}, P^{RS} [49]. The clustering results of several methods are compared via an independent measure of cluster agreement which is widely used in psychotherapy research – the so-called $\tilde{\kappa}$-coefficient [45]. The $\tilde{\kappa}$-coefficients may be interpreted according to tab. 1 [45]. The results are depicted in tab. 2.

As one can see, the SSMPD yields the best solutions whereas the PDCDA and EA show slightly less significant improvements of the original cluster solution. If one adds expert knowledge into the EA fitness, one yields for the different cluster problems $\tilde{\kappa}_W = 0.435$, $\tilde{\kappa}_{RO} = 0.421$, $\tilde{\kappa}_{RS} = 0.504$. Hence, the expert knowledge yields a better accuracy for the wishes which suggest that the expert knowledge helps to find minima in a complicated fitness landscape. This result is in agreement with medical experts stating that the classification of wishes is the most crucial point in CCRT [33]. For the other cases, the additional knowledge seems to be an additional restriction.

6. Conclusion and Outlook

Although, at a first glance, the growing number of applications in the field of biomedicine 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 biosignal

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 biomedicine. 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 data preprocessing, postprocessing, interpretation, and quality assessment. This challenge can only be managed by close interdisciplinary cooperation of medical doctors, biologists, engineers, and computer scientists. Hence, this subject can serve as an example for lively cross-fertilization between neural network computing and related research.

Acknowledgements: We thank Oliver Lange, Johannes Behrends, Dominik R. Dersch, Frank Vietze, Gerda L. Leinsinger, Dorothee Auer, Johannes Rieger, Oliver Meissner, Phil Hoole, Christian Kroos, Anja Geumann, Hans G. Tillmann, W. Herrmann, and H. Barthel for their support.

References

[1] M. Anderberg, editor. *Cluster analysis for applications*. Academic Press, New York, 1973.

[2] R. Baumgartner, C. Windischberger, and E. Moser. Quantification in functional magnetic resonance imaging: fuzzy clustering vs. correlation analysis. *Magnetic Resonance Imaging*, 16(2):115–125, 1998.

[3] E. Cantu-Paz. Topology, migration rates and multi-population parallel genetic algorithms. Technical Report 99007, Illinois Genetic Algorithms Laboratory, University of Illionois at Urbana-Champaign, Urbana IL, 1999.

[4] D. Dersch. *Eigenschaften neuronaler Vektorquantisierer und ihre Anwendung in der Sprachverarbeitung*. Verlag Harri Deutsch, Reihe Physik, Bd. 54, Thun, Frankfurt am Main, 1996. ISBN 3-8171-1492-3.

[5] R. Duda and P. Hart. *Pattern Classification and Scene Analysis*. Wiley, New York, 1973.

[6] M. Eisen, P. Spellman, and D. Botstein. Cluster analysis and display of genome-wide expression patterns. *Proceedings of the National Academy of Sciences*, 95:14863–14868, 1998.

[7] H. Fischer and J. Hennig. Neural-network based analysis of MR time series. *Magn. Reson. Med.*, 41(1):124–131, 1999.

[8] E. Forgy. Cluster analysis of multivariate data: Efficiency vs. interpretability of classifications. *Biometrics*, 21:768, 1965.

[9] M. Galicki, U. Möller, and H. Witte. Neural clustering networks based on global optimisation of prototypes in metric spaces. *Neural Computing and Applications*, 5:2–13, 1997.

[10] F. Girosi and T. Poggio. Networks and the best approximation property. *Biological Cybernetics*, 63:169–176, 1990.

[11] T. Greapel and K. Obermayer. A stochastic self-organizing map for proximity data. *Neural Computation*, 11(1):139–155, 1999.

[12] T. Hofmann and J. Buhmann. Pairwise data clustering by deterministic annealing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(1):1–14, 1997.

[13] J. Huhse and A. Zell. Evolution strategy with neighborhood attraction - A robust evolution strategy. In L. Spector, E. D. Goodman, A. Wu, W. B. Langdon, H.-M. Voigt, M. Gen, S. Sen, M. Dorigo, S. Pezeshk, M. H. Garzon, and E. Burke, editors, *Proceedings of the Genetic and Evolutionary Computation Conference (GECCO-2001)*, pages 1026–1033, San Francisco, California, USA, 7-11 2001. San Francisco, CA 94104, USA.

[14] T. Kohonen. The self-organizing map. *Proceedings of the IEEE*, 78(9):1464–1480, 1990.

[15] Y. Linde, A. Buzo, and R. Gray. An algorithm for vector quantizer design. *IEEE Transactions on Communications*, 28:84–95, 1980.

[16] T. Martinetz and K. Schulten. A 'neural gas' network learns topologies. In *Proceedings of the International Conference on Artificial Neural Networks ICANN*, pages 397–402, Amsterdam, 1991. Elsevier Science Publishers.

[17] Z. Michalewicz. *Genetic Algorithms + Data Structures = Evolution Programs*. Springer–Verlag Berlin Heidelberg New York, third, revised and extended edition, 1996.

[18] J. Moody and C. Darken. Fast learning in networks of locally-tuned processing units. *Neural Computation*, 1:281–294, 1989.

[19] H. Ritter, T. Martinetz, and K. Schulten. *Neural Networks*. Addison Wesley, New York, 1991.

[20] K. Rose, E. Gurewitz, and G. Fox. Vector quantization by deterministic annealing. *IEEE Transactions on Information Theory*, 38(4):1249–1257, 1992.

[21] M. Su and H. Chang. A new model of self-organizing neural networks and its application in data projection. *IEEE Transactions on Neural Networks*, 12(1):153–158, 2001.

[22] P. Tamayo, D. Slonim, J. Mesirov, Q. Zhu, S. Kitareewan, E. Dmitrovsky, E. Lander, and T. Golub. Interpreting patterns of gene expression with self-organizing maps: Methods and application to hematopoietic differentiation. *Proceedings of the National Academy of Science USA*, 96:2907–2912, March 1999.

[23] A. Wismüller, J. Behrends, O. Lange, D. Dersch, G. 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 Suppl.*, 221(P), 2001.

[24] A. Wismüller, J. Behrends, O. Lange, M. Jukic, K. Hahn, and D. Auer. Flexible machine learning image analysis for high-precision computer-assisted segmentation of multispectral MRI data sets in patients with multiple sclerosis. *Radiology Suppl.*, 221(P), 2001.

[25] A. Wismüller and D. Dersch, editors. *Symposion über biologische Informationsverarbeitung und Neuronale Netze – SINN '95, Konferenzband*. Hanns-Seidel-Stiftung, München, 1996.

[26] A. Wismüller and D. Dersch. Neural network computation in biomedical research: chances for conceptual cross-fertilization. *Theory in Biosciences*, 116(3):229–240, 1997.

[27] A. Wismüller, D. Dersch, B. Lipinski, K. Hahn, and D. Auer. A neural network approach to functional MRI pattern analysis – clustering of time-series by hierarchical vector quantization. In L. Niklasson, M. Bodèn, and T. Ziemke, editors, *ICANN'98 — Proceedings of the 8th International Conference on Artificial Neural Networks, Skövde, Sweden. Perspectives in Neural Computing*, volume 2, pages 123–128, London, Berlin, New York, 1998. Springer-Verlag.

[28] A. Wismüller, D. Dersch, F. Vietze, G. Leinsinger, K. Hahn, and D. Auer. Cluster analysis of biomedical image time-series. *International Journal of Computer Vision*, 46(2):103–128, 2002.

[29] A. Wismüller, F. Vietze, and D. Dersch. Segmentation with neural networks. In I. Bankman, R. Rangayyan, A. Evans, R. Woods, E. Fishman, and H. Huang, editors, *Handbook of Medical Imaging*, Johns Hopkins University, Baltimore, 2000. Academic Press. ISBN 0120777908.

[30] A. Wismüller, F. Vietze, D. Dersch, J. Behrends, K. Hahn, and H. Ritter. The deformable feature map – a novel neurocomputing algorithm for adaptive plasticity in pattern analysis. *Neurocomputing*, 2001. Accepted for publication.

[31] A. Wismüller, F. Vietze, D. Dersch, K. Hahn, and H. Ritter. The deformable feature map — adaptive plasticity in function approximation. In L. Niklasson, M. Bodèn, and T. Ziemke, editors, *ICANN'98 — Proceedings of the 8th International Conference on Artificial Neural Networks, Skövde, Sweden. Perspectives in Neural Computing*, volume 1, pages 222–227, London, Berlin, New York, 1998. Springer-Verlag.

[32] A. Wismüller, F. Vietze, D. Dersch, G. Leinsinger, H. Ritter, and K. Hahn. Adaptive self-organized template matching of the gray-level feature space for automatic segmentation of multispectral MRI data of the human brain. *Radiology Suppl.*, 213(P), 1999.

[33] C. Albani, T. Villmann, B. Villmann, A. Körner, M. Geyer, D. Pokorny, G. Blaser, and H. Kächele. Kritik und erste Reformulierung der kategorialen Strukturen der Methode des Zentralen Beziehungs–Konflikt–Themas (ZBKT). *Psychotherapie, Psychosomatik und Medizinische Psychologie*, 49(11):408–421, 1999.

[34] J. Barber, P. Crits-Christoph, and L. Luborsky. A guide to the CCRT Standard Categories and their classification. In L. Luborsky and P. Crits-Chrostoph, editors, *Understanding Transference*, pages 37–50. Basic Books New York, 1990.

[35] H.-U. Bauer, R. Der, and M. Herrmann. Controlling the magnification factor of self–organizing feature maps. *Neural Computation*, 8(4):757–771, 1996.

[36] T. Graepel, M. Burger, and K. Obermayer. Self-organizing maps: generalizations and new optimization techniques. *Neurocomputing*, 21(1–3):173–90, 1998.

[37] M. M. V. Hulle. *Faithful Representations and Topographic Maps*. Wiley Series and Adaptive Learning Systems for Signal Processing, Communications, and Control. Wiley & Sons, New York, 2000.

[38] E. C. Ifeachor, A. Sperduti, and A. Starita. *Neural Networks and Expert Systems in Medicine and Healthcare*. World Scientific, Singapore, 1998.

[39] S. Kaski. SOM-based exploratory analysis of gene expression data. In N. Allinson, H. Yin, L. Allinson, and J. Slack, editors, *Advances in Self-Organizing Maps*, pages 124–131. Springer, London, 2001.

[40] S. Kaski, J. Sinkkonen, and J. Nikkilä. Clustering gene expression data by mutual information with gene function. In G. Dorffner, H. Bischof, and K. Hornik, editors, *Artificial Neural Networks—ICANN 2001*, pages 81–86. Springer, Berlin, 2001.

[41] T. Kohonen. *Self-Organizing Maps*, volume 30 of *Springer Series in Information Sciences*. Springer, Berlin, Heidelberg, 1995. (Second Extended Edition 1997).

[42] J. Lampinen and T. Kostiainen. Generative probability density model in the self-organizing map. In U. Seiffert and L. Jain, editors, *Self-Organizing Neural Networks*, Studies in Fuzziness and Soft Computing, pages 75–92. Physica-Verlag, Heidelberg, New York, 2001.

[43] L. Luborsky. The core conflictual relationship scheme. In N. Freedman and S. Grand, editors, *Communicative Structure and Psychic Structures*. Plenum Press New York, 1977.

[44] H. Malmgren, M. Borga, and L. Niklasson. *Artificial Neural Networks in Medicine and Biology*. Springer-Verlag London, 2000.

[45] L. Sachs. *Angewandte Statistik*. Springer Verlag, 7-th edition, 1992.

[46] T. Saito. Presenting symptoms and natural history of wilson disease. *European Journal of Pediatrican*, 146:261–265, 1987.

[47] J. Sinkkonen and S. Kaski. Clustering based on conditional distributions in an auxiliary space. *Neural Computation*, 14:217–239, 2002.

[48] J. Vesanto and E. Alhoniemi. Clustering of the self-organizing map. *IEEE Transaction on Neural Networks*, 11(3):586–600, 2000.

[49] T. Villmann. Neural networks approaches in medicine – a review of actual developments. In *Proc. Of European Symposium on Artificial Neural Networks (ESANN'2000)*, pages 165–176, Brussels, Belgium, 2000. D facto publications.

[50] T. Villmann. Evolutionary algorithms with subpopulations using a neural network like migration scheme and its application to real world problems. *Integrated Computer- Aided Engineering*, 9(1):25–36, 2002.

[51] W. Hermann, H. Barthel, S. Hesse, F. Grahmann, H.-J. Kühn, A. Wagner, Th. Villmann. Comparison of clinical types of Wilson's disease and glucose metabolism in extrapyramidal motor brain regions. *Journal of Neurology*, to appear, 2002.

[52] T. Villmann, R. Der, M. Herrmann, and T. Martinetz. Topology Preservation in Self–Organizing Feature Maps: Exact Definition and Measurement. *IEEE Transactions on Neural Networks*, 8(2):256–266, 1997.

[53] T. Villmann, W. Hermann, and M. Geyer. Variants of self-organizing maps for data mining and data visualization in medicine. *Neural Network World*, 10(4):751–762, 2000.

[54] D. Wang, H. Ressom, M. Musavi, and C. Dommisoru. Double self-organizing maps to cluster gene expression data. In M. Verleysen, editor, *Proc. of European Symposium on Artificial Neural Networks (ESANN'2002)*, page in this Proc., Brussels, Belgium, 2002. D facto publications.

[55] P. L. Zador. Asymptotic quantization error of continuous signals and the quantization dimension. *IEEE Transaction on Information Theory*, (28):149–159, 1982.