ROBUST NONLINEAR ADAPTIVE NETWORK CLASSIFICATION OF ANAESTHESIA

Baumgart-Schmitt, R.; Walther, C.*; Backhaus, K.*; Reichenbach, R.*; Sturm, K.-P.**; Jaeger, U.****

* Lab. Neuroinformatics, Department of Electrical Engineering,
University of Applied Sciences Schmalkalden, Germany
** Department of Anaesthesia, Hospital of Schmalkalden, Germany
*** Medical practice of Zella-Mehlis, Germany

ABSTRACT

To control the administration of hypnotics during operations it is necessary to classify the depth of anaesthesia in a robust and efficient way. The frontal EEG was selected as a feature source. Different populations of topologically optimized trained neural networks solved the problem of robust classification. Robustness in this context means that the performance of the classifier is independent of the agent administration strategies used to induce different depth of anaesthesia. Training and optimization of the neural networks were supported by genetic programming and simulated evolution. The results are compared to the performance of the BIS XP monitor. For this purpose we applied this monitor to all patients in the cooperating hospital and measured the frontal EEG in a parallel way. The anaesthetist used several autonomic parameters like heart rate and blood pressure to recognize the depth of anaesthesia. The performance of both approaches using the frontal EEG has been measured by confusion matrices which represent the concordances and deviants between the scores of the anaesthetist and the results of the automatic procedure. Our approach led to higher degrees of concordances for all stages especially if the anaesthetic agent ketamine is included. The extension of the evaluated EEG frequency range improved the results for the difficult recognition of transitional stages.

Index Terms— Robustness, classification, anaesthesia, EEG, artificial neural networks, evolutionary and genetic algorithms, genetic programming, ketamine, BIS XP.

1. INTRODUCTION

In developing a flexible mobile device which should be able to classify and forecast the depth of consciousness during anaesthesia in a robust and economical way only one frontal EEG channel was evaluated. The classification was done by populations of optimized neural networks. The networks were trained by the parallel computable system SASCIA [1] [2] [3] [4] which is running on a cluster of PCs. The different parallel running components in the master slave architecture of SASCIA are synchronized by the LINDA tuple space, which is described below. The tuple space originally introduced by Gelernter [5] was implemented by means of Python and a corresponding C interface (PyLinda) by Wilkinson [6]. The distributed computing algorithms offered opportunities to evaluate the feature space in an automatic way. The quality of the approach is estimated by the concordance between the automatically classified stages of anaesthesia and the stages scored by the anaesthetist. According to our aim of developing adaptive and robust nonlinear classifiers the effects of

- different strategies of anaesthesia (agent or drugs effects),
- extended frequency range of the used features and
- different learn, test and validation sets

on the results are studied.

Robustness of the classifier in this context means that different mixtures of drugs or agents used during anaesthesia should not lead to a lower degree of concordances. For example it is well known that the effects of the agent ketamine on EEG differ significantly from the effects of other agents.

The ketamine effects on the quality of classification were considered. Taking into account that some monitors which are able to classify the depth of anaesthesia are offered on the market we selected the BIS XP monitor as a reference device. This monitor is extensively documented and tested, so we tried to compare the performance of this device with our approach.

Kreuer et. al. [7] compared the performance of this BIS XP Monitor with the Narcotrend monitor by measuring the effects of the agent isoflurane on the EEG during general aneasthesia. They did not found any remarkable differences. Hans et. al. [8] published the warning that Ketamine may confound the index of the BIS monitor. That could be lead to difficulties if this index is used to steer administration of anaesthesia. Neural networks were also used to predict the responses to incision during anaesthesia by Huang et. al. [9]. They distinguished between responder and non responder.

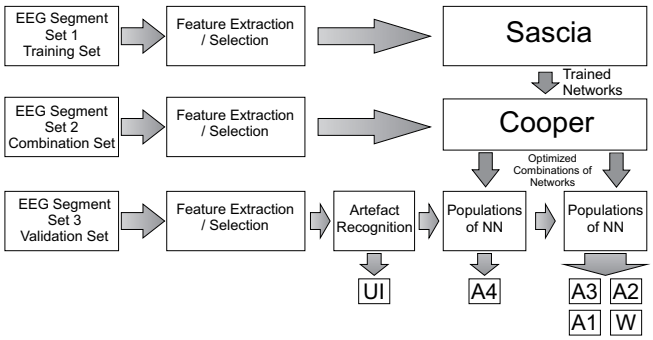


Fig. 1. Approach to train, optimize and apply populations of neural networks which serve as robust adaptive classifiers. Evolutionary algorithms, simulated genetic algorithms (SASCIA) and genetic programming (COOPER) are included in a multi-layer-approach for training purposes. The validation is performed by a chain of classifiers.

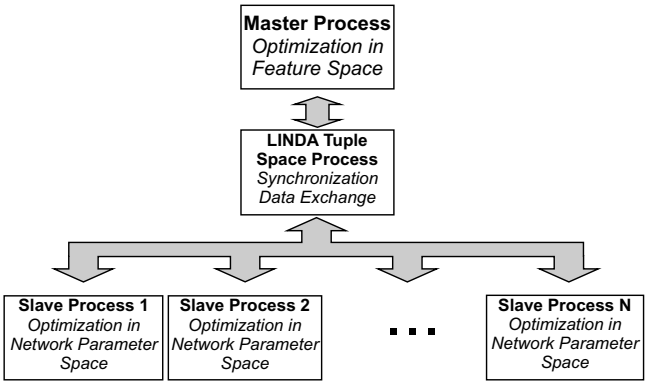


Fig. 2. Different parallel working processes of the training and optimization tool SASCIA for neural networks are synchronized by the LINDA concept. The master slave architecture of SASCIA is running on a PC cluster in parallel mode. The data exchange between master and slave processes is realized by using the LINDA tuple space.

2. METHODS

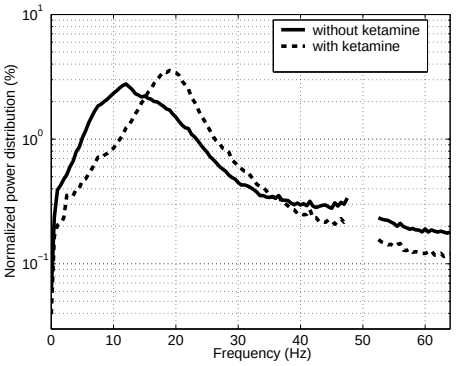
2.1. Data recording

The EEG and vegetative parameters of 63 patients were recorded during surgery in the cooperating hospitals of Schmalkalden and Zella Mehlis (Germany). The complete data pool of all patients was divided into three independent parts: n EEG epochs with 10 second length and n= 14028 of 22 patients were included in the training data set, m epochs with m= 11884 of 22 patients were used to optimize the cooperation of the trained networks and the data of 19 patients with v epochs, v= 8889 served as validation set. Anaesthesia was induced either by a mixture of ketamine and propofol, or rapifen and propofol. The hypnotic stage was maintained by the administration of sevofluran. The EEG was measured bipolar near Fp1 and Fp2 and sampled by the rate of 505 values per second by our mobile recorder system Quisi mini. This system uses a 16 Bit Sigma-Delta analog-to-digital converter (16 Bit ADC) and a flash-RAM-card to store the data in two byte mode. The electrodes of the BIS XP monitor were fixed on the forehead of the 19 patients selected for the validation set according to the manual of the company. Therefore the EEG could be sampled by both devices in a synchronous way.

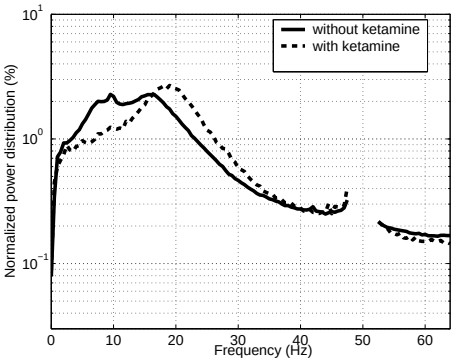
2.2. Nonlinear Classification

Fig. 1 shows the sequence of the main procedures which were used for EEG feature selection, training of neural networks, optimizing the topologies of networks and the classification of the validation set.

At least four stages and awake should be distinguished by the nonlinear classifier to guide the anaesthesia: The label A1 marks the transitional stage of sedation or light anaesthesia, A2 the stage of moderately deep anaesthesia or the stage of surgical tolerance, A3 the stage of deep anaesthesia and A4 the stage of very deep anaesthesia or stage with the occurrence of so called burst suppression patterns. These staging are in accordance with the well known definitions of Kugler [10]. The developed and continuously improved system SASCIA is the most important part of our approach. Features of EEG and manually scored segments are included to train and optimize neural networks. SASCIA can be regarded as a flexible multi layer system controlled by external files. The core of the system consists of a training algorithm based on error back propagation. The middle layer serves the topological optimization by simulated evolution. The optimal selection of the features is realized in the outer layer by genetic algorithms. Selection, recombination, mutation and inversion were used as genetic operators. From the complete set of trained networks a subset of networks is tested by the tool COOPER [11] to find the optimal subset of networks by genetic programming. That means each neural network is represented by one bit in a bit string and the bits are switched by genetic operators. If one bit is switched on the networks is activated in the selected subset. According to the included neural networks the length of the string varied from 150 to 188. Optimality aims at excellent cooperation of the networks in their subsets. According to Fig. 1 COOPER uses the second part of EEG segments which is called combination set. The third set of EEG segments in Fig. 1 serves the validation of the population of networks. The first population of neural networks which is able to distinguish between stage A4 and the set of other stages consists of 8 networks. Eleven neural networks serve in the second population to recognize the three stages A1, A2, A3 and Wake. The third population contains 55 neural networks. The topologies of the networks differ in all populations according to the selected feature subsets by genetic algorithms and the optimization. The mean



(a) A2

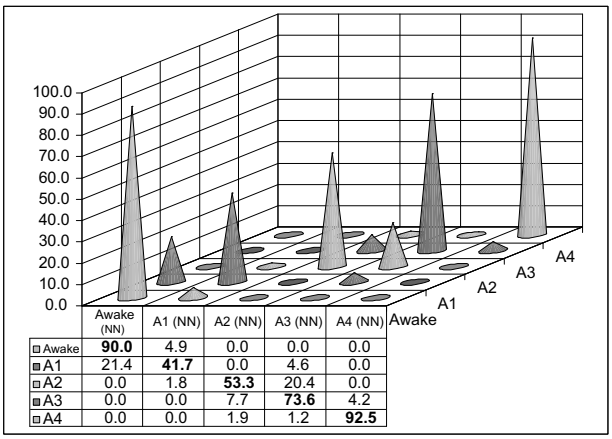


(b) A3

Fig. 3. First differences of the EEG time series were transformed into the frequency range. The means and standard deviations of EEG power distributions in dependence of stage A2 (top) and A3 (bottom) show remarkable differences induced by the agent ketamine (gray coloured). All logarithmic values of the graph are normalized to the complete accumulated power until 64 Hz. The 50 Hz band was eliminated.

of the number of inputs, the used learning rates of the batch training and the momentum were 18, 0.661 and 0.333 for all members of the three populations. Generally 2 hidden layers with different numbers of units were used.

During recording of EEG several sources of artefacts should be taken into consideration: active or passive body movements of the patients, eye movements, EEG deflection by an electro surgery system or other electrical devices. Artefacts were eliminated in a separate way by a multistage procedure and labelled by UI which means that it is an unidentified segment. The following classification is performed by a chain of two classifiers. In a first chainlink the burst suppression (A4) epochs are separated from the rest. The EEG during stage A4 can be characterized by more or less short burst and suppression lines. Than the rest is partitioned into the different stages of anaesthesia A1, A2, A3 and wake. The membership of 10 seconds EEG epochs to classes of depth of anaesthesia is simultaneously calculated by sets of trained networks. The



	Awake (NN)	A1 (NN)	A2 (NN)	A3 (NN)	A4 (NN)
Awake	**90.0**	4.9	0.0	0.0	0.0
A1	21.4	**41.7**	0.0	4.6	0.0
A2	0.0	1.8	**53.3**	20.4	0.0
A3	0.0	0.0	7.7	**73.6**	4.2
A4	0.0	0.0	1.9	1.2	**92.5**

Fig. 4. The performance of the nonlinear adaptive classifier concerning the evaluation of EEG epochs is demonstrated by confusion matrices. The degrees of concordances represent the medians on all subjects. That means the values in each rows do not accumulate to 100%. The sum of the single values for one operation is equal to 100.

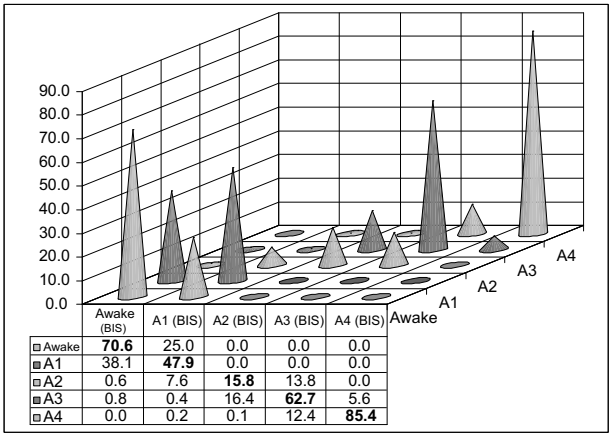
networks contribute to final decision by the calculation of the median.

2.3. Parallel distributed training

The system SASCIA can run on a single computer or a distributed platform. To save computing time SASCIA was implemented on PC Cluster. The PC Cluster consists of 10 PCs which are connected by a local area network. The master PC controls 9 slave PCs. The master slave architecture of the distributed implementation can be seen in Fig. 2. Up to now N is equal to 9. The design of the architecture is not restricted to a certain number of slaves. In accordance with the existing resources the control files of SASCIA have to be adapted to guaranty the efficiency of the computing process. The parallel running processes are coordinated by the LINDA concept. This concept serves for synchronization of the different processes which have access to the common storage with a data container. One process is created which is responsible for that data container called tuple space. The master process and the N slave processes are able to access that container in a bidirectional mode. The tuples correspond to ordered parameter sets which control the concerted action of master and slaves.

3. RESULTS

The first results serve the comparison of our approach to the BIS XP monitor. The main topic of our contribution should be dedicated to the robustness of classification. The overwhelming majority of the 53 features were extracted in the frequency range. Therefore the differences in frequency range due to different agents should be explained by the Fig. 3.



	Awake (BIS)	A1 (BIS)	A2 (BIS)	A3 (BIS)	A4 (BIS)
Awake	**70.6**	25.0	0.0	0.0	0.0
A1	38.1	**47.9**	0.0	0.0	0.0
A2	0.6	7.6	**15.8**	13.8	0.0
A3	0.8	0.4	16.4	**62.7**	5.6
A4	0.0	0.2	0.1	12.4	**85.4**

Fig. 5. Confusion matrix calculated on the basis of the BIS XP monitor results.

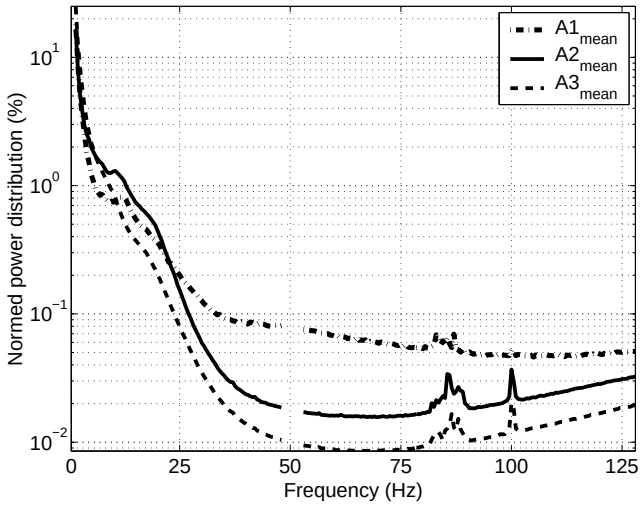


Fig. 6. Means of EEG power distributions are dependent of stage A1 (top), A2 (middle) and A3 (bottom) above 40 Hz especially. All logarithmic values of the graph are normalized to the complete accumulated power until 128 Hz. The 50 Hz band was eliminated. The sample rate was 505.3 Hz. The means were calculated on all 63 subjects. The classification was done manually by the anaesthetist who used the complete sources of information.

First differences of the EEG time series were transformed in the frequency range. The means and standard deviations of EEG power distributions in dependence of stage A2 (top) and A3 (bottom) show remarkable differences induced by the agent ketamine. All logarithmic values of the graph are normalized to the complete accumulated power. According to the 128 Hz down sampled EEG time series the maximum frequency is 64 Hz. The 50 Hz band was eliminated. The dotted lines of EEG influenced by ketamine show a shift of the peak power values to higher frequencies in each class A2 and A3. The black lines serve as references because in these cases no ketamine were applied. The means of each line were calculated on 9 different subjects. The classification was done manually by the anaesthetist who used the complete sources of information. The results of classification are demonstrated by confusion matrices. The matrix in Fig. 4 show the degrees of concordance in percentage between the manually scored 10 seconds EEG epochs and the results received by the populations of neural networks. The complete validation data base consisting of 8889 records from 19 patients was used. The features were selected in the frequency range up to 64 Hz. The first step of classification was performed by 8 networks, which separated stage A4 (burst suppression) from the pool of stages. The four class problem was solved by 11 networks simultaneously.

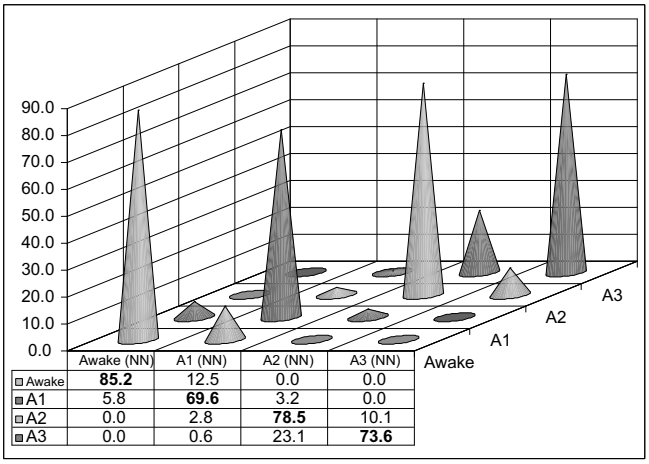
The results of the BIS XP monitor in Fig. 5 differ remarkably. The values of concordance measured in percentage show that our approach is more robust against drug influences especially ketamine than the algorithm of the BIS XP monitor. The relatively low degree of concordance for stage A1 was the cause to improve the performance of the neural networks. Regarding to the power distributions in Fig. 6 the upper frequency range show differences between the line which corresponds to the stage A1 and the A2, A3 lines. The number of extracted features was extended to 62. Nine features represent the frequency range above 64 Hz. By means of SASCIA 55 networks were topologically optimized. All 55 networks contribute to the results represented in Fig. 7.

In comparison with the confusion matrix of Fig. 4 the A1 concordance could be raised from 41.7% to 69.6%. If the matrix will be completed by the results A4 versus rest of Fig. 4 the matrix of Fig. 7 represents the lowest degree of discordance level. To reduce the number of classifying neural networks and to improve the results further the optimization tool COOPER was applied. The confusion matrix of the optimized subset of nine neural networks can be seen in Fig. 8 and compared to the matrix of Fig. 7. Two concordance values could be improved at the expense of the other two values. The results of Fig. 7 and 8 were received by the inclusion of the same validation data base consisting of 14038 records from 22 patients.

4. INTERPRETATION

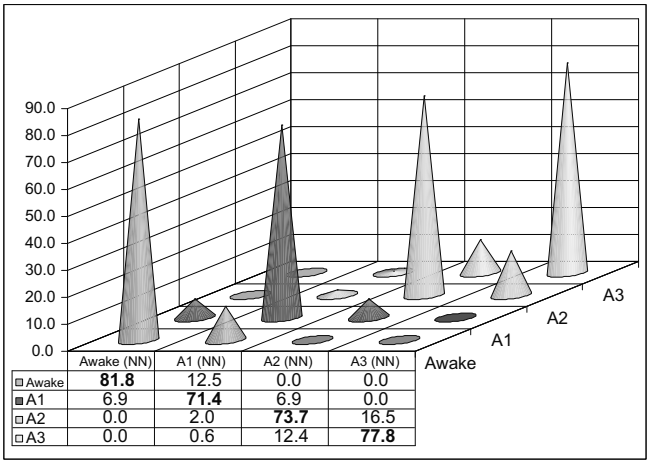
Because the character of the EEG is more transient in A4 segments than in other segments the frequency range offers no advantages to the time range. Therefore features of time range were only used to separate A4 segments from the rest. This motivates the processing in a chain of classifiers. The advantages of this approach could be verified by the improved results.

Comparing Fig. 7 with Fig. 4 the extension of the frequency range from 64 until 128 Hz for feature extraction led to remarkable effects in A1 classification. Both the concor-



	Awake (NN)	A1 (NN)	A2 (NN)	A3 (NN)
Awake	**85.2**	12.5	0.0	0.0
A1	5.8	**69.6**	3.2	0.0
A2	0.0	2.8	**78.5**	10.1
A3	0.0	0.6	23.1	**73.6**

Fig. 7. In comparison with the confusion matrix of Fig. 4 the A1 concordance could be raised from 41.7% to 69.6%. The EEG frequency range which served as the primary basis of feature extraction was extended from 64 Hz to 128 Hz. The validation data base consisting of 14038 records from 22 patients was used. The classification was done by 55 networks. All 55 networks are topologically optimized by the evolutionary layer of SASCIA and use different subspaces of the complete features space.



	Awake (NN)	A1 (NN)	A2 (NN)	A3 (NN)
Awake	**81.8**	12.5	0.0	0.0
A1	6.9	**71.4**	6.9	0.0
A2	0.0	2.0	**73.7**	16.5
A3	0.0	0.6	12.4	**77.8**

Fig. 8. Results of an optimally selected set of nine cooperating neural networks are presented. The set of nine networks are a subset of the set of 55 networks. The subset was selected by the tool COOPER of Fig. 1. The data base used by cooper consists of 11909 feature vectors with associate labels from 22 patients. The validation data base consisting of 14038 records from 22 patients as in Fig. 7 was implemented.

dance between the opinion of the expert and the results of the classifier and the type of misclassification could be improved. The stage A1 can be regarded as a transitional stage between the initial administration of medications and the loss consciousness. There are muscle activities which can be measured by EMG (electromyography). The influence of muscle activities on EEG can not be neglected above 64 Hz especially (Fig. 6). Therefore the reasons for the improved results are in concordance with the physiology.

The set of neural networks could be reduced from 55 to 9 networks by means of the optimization tool COOPER without loss of performance. That is important for the real time implementation of our mobile system. Training and optimizing populations of neural networks consume much computing time. According to the distributed computing implementation of SASCIA a lot of computing time can be saved. The chance that the computing time drops down will be increased by further inclusion of slaves. But the exact saving of computation time in dependence of the number of slaves is unpredictable. There are several reasons for that: The global aim of SASCIA consists in finding the optimal neural networks. That means the parameter of the networks, the number of features, the number of hidden units, the number of iteration for instance have to be varied. A comprehensive search in the parameter space will lead to great variability of the network architectures. A great variance of computation times follows. According to the LINDA concept the slaves have to wait until the master generates the next task tuple. The optimization procedures need the complete set of goal values to make the next steps. The different waiting times of the slaves make it impossible to predict the saving of computing times. Furthermore all slave processes should be supported by platforms with equal performance to minimize the loss of efficiency.

5. CONCLUSION AND FUTURE WORK

The results of concordances base on the evaluation of the differences between the manually scored EEG and the automatic classification by means of four criteria. Because of the unbalanced distributions of class specific segments more than one criterion is necessary to control the optimization procedure adequately. Up to now the four criteria were parametrically aggregated to one criterion which was used to control the simulated evolution. The first steps in including the multi-criteria optimization were done and we expect a further improvement in performance. In our approach are contained two problems which can be solved by a two class solution. The separation of unidentified segments and of A4 segments from the rest respectively was successfully performed by supported vector machines. This procedure seems to be an alternative way to increase the validation. The successful multi-criteria prediction of anaesthetic stages is a necessary step before the implementation of the algorithms can be accepted by the anaesthetist.

6. REFERENCES

[1] Baumgart-Schmitt, R., “Praktische Anwendungen Neu-

ronaler Netze," *LIT Verlag Hamburg*, 1996.

[2] Baumgart-Schmitt, R. ; Herrmann, W. M. ; Eilers, R. ; Bes, F., "On the use of neural network techniques to analyse sleep eeg data, first communication: Application of evolutionary and genetic algorithms to reduce the feature space and to develop classification rules," *Pharmacoelectroencephalography*, vol. 36, pp. 194–210, 1997.

[3] Baumgart-Schmitt, R. ; Herrmann, W. M. ; Eilers, R., "On the use of neural network techniques to analyze sleep eeg data, third communication: Robustification of the classificator by applying an algorithm obtained from 9 different networks," *Pharmacoelectroencephalography*, vol. 37, pp. 49–58, 1998.

[4] Baumgart-Schmitt, R. ; Wenzel, A., "The application of multiobjective genetic programming for the adaptive sleep eeg-analysis by neural networks," *Hybrid and Adaptive Computational Intelligence in Medicine and Bio- Informatics. Aegean : Univ. of Aegean*, pp. 64–77, 2003.

[5] Gelernter, D., "Generative communication in Linda," *ACM Transactions on Programming Languages and Systems*, vol. 7(1), pp. 80–112, January 1985.

[6] Wilkinson, A., "PyLinda," *www-users.cs.york.ac.uk/aw/pylinda*, 2005.

[7] Kreuer, S.; Bruhn, J.; Larsen, R.; Grundmann, U.; Shafer, S. L.; Wilhelm, W., "Application of bispectral index and narcotrend index to the measurement of the electroencephalographic effects of isoflurane with and without burst suppression," *Anesthesiology*, vol. 101, pp. 847–854, 2004.

[8] Hans, P.; Dewandre, P.-Y.; Brichant, J. F.; Bonhomme, V., "Comparative effects of ketamine on bispectral index and spectral entropy of the electroencephalogram under sevoflurane anaesthesia," *British Journal of Anaesthesia*, vol. 94 (3), pp. 336–340, 2005.

[9] Huang, L.; Wang, W.; Singare, S., "Using back propagation feedback neural networks and recurrence quantification analysis of eegs predict responses to incision during anesthesia," *Lecture Notes in Computer Science, Springer-Verlag Berlin Heidelberg*, vol. 4222/2006, pp. 364–373, 2006.

[10] Kugler, J., "Elektroencephalographie in Klinik und Praxis," *Thieme Verlag, Stuttgart, New York*, 1981.

[11] Wenzel, A., "Robuste Klassifikation von EEG-Daten durch Neuronale Netze," *Shaker Verlag Aachen*, 2005.