NeuroPipe - A Neural Network Based System for Ultrasonic Inspection

Robert Suna and Karsten Berns
Research Center for Computer Science
University of Karlsruhe
Haid-und-Neu-Str. 10-14,
76131 Karlsruhe
GERMANY

Abstract: With a probe for gas, oil and other pipelines a huge number of ultrasonic readings of the wall condition is collected. Based on the recorded wall thicknesses of this so-called pipe pig the Research Center for Computer Science (FZI) has developed an automatic inspection system called NeuroPipe. NeuroPipe has the task to detect defects like metal loss. The kernel of this inspection tool is a hybrid neural classifier which was trained using manually collected defect examples by the Pipetronix company. This paper focuses on some aspects of successful use of learning methods in an industrial application and on the difficulty of interpretation of sometimes faulty sensor measurements.

Keywords: Neural network, interactive learning process, hybrid pattern recognition

Robert Suna studied Computer Science at the University of Karlsruhe, where his main research was in the field of neural networks. Since 1993 he has been a member of the Research Center for Computer Science at the University of Karlsruhe, the Interactive Planning Technology (IPT) group. His main research topics are the industrial and commercial application of neural networks, the use of neural networks for robotic problems as well as the improvements of neural learning functions.

Dr. Karsten Berns studied Computer Science at the University of Kaiserslautern. Since 1989 he has been employed at the Research Center for Computer Science of the University of Karlsruhe. He is head of Department at Interactive Planning Technology (IPT) group. His main research area is the application of neural networks in the field of robotics. In his Ph.D dissertation (1994) he examined the use of neural networks for the control of a six-legged walking machine.

1 Introduction

Worldwide several hundred thousands kilometers of gas and oil pipelines exist which have to be checked at regular intervals as a precaution against possible environmental catastrophes. Beside other inspection tools Pipetronix GmbH (PTX) Karlsruhe has developed a special ultrasonic based probe called UltraScan (see Figure 1). This tool is launched into the pipeline and propelled through the medium flow with about 1m/s. During this passive movement the ultrasound sensors mounted at the head of the pig (see Figure 2), recorded the stand-off, i.e. the distance between the sensor and the inner pipeline wall, as well as the wall thickness (see Figure 3). These two readings, which are made every 3mm, are stored on data carriers for offline processing.

The data provided by the sensors are processed as 2D images, the so-called C-scan (see Figure 5). The C-scan shows the readings obtained from all sensors as tracks positioned next to each other. The deviations from the nominal wall thickness are colour-coded. Since two values exist for every reading, two C-scans (for stand-off distance and wall thickness, resp.) are necessary for a complete representation of the data.

In the past, the pipeline inspection was done manually based on the coloured C-scan. In the Pipetronix Interpretation Department, several people interpret these ultrasound images over many weeks for one pipe run. In order to facilitate interpretation work, the NeuroPipe program has been developed which classifies the possible defects into five basic classes described in Table 1. The program is designed in a way that the classification of defects in the pipeline is as reliable as possible and generally conservative.

In the next section the different parts of the system architecture of NeuroPipe are described. A detailed description can be found in [11]. The main part of this paper concentrates on the experience of the interactive learning process used to generate the hybrid neural defect classifier. At the end of this paper an improvement of this interactive learning concept is presented.

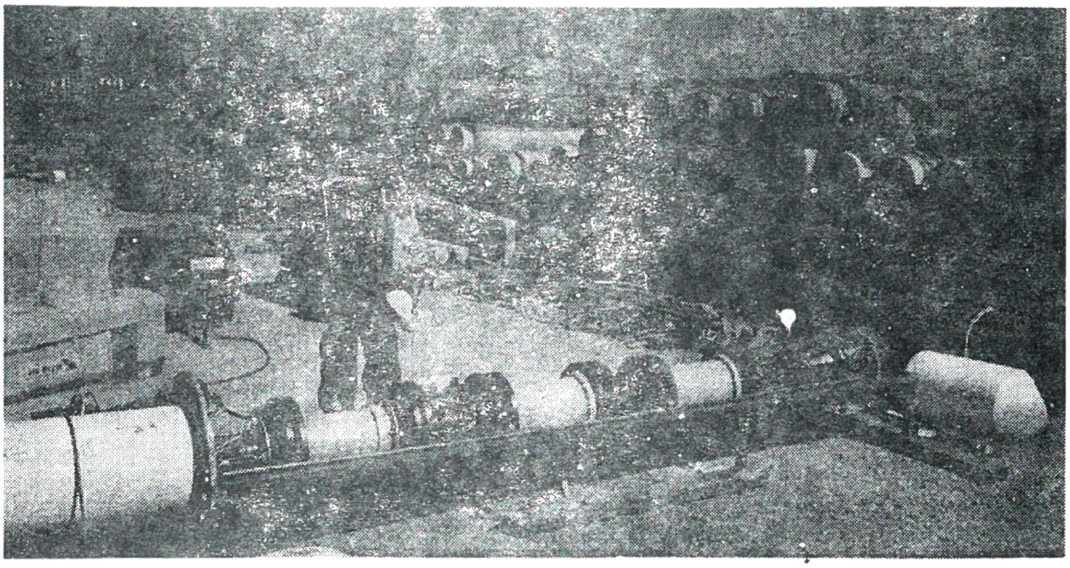

Figure 1: UltraScan Tool Before Inserting in A Pipeline

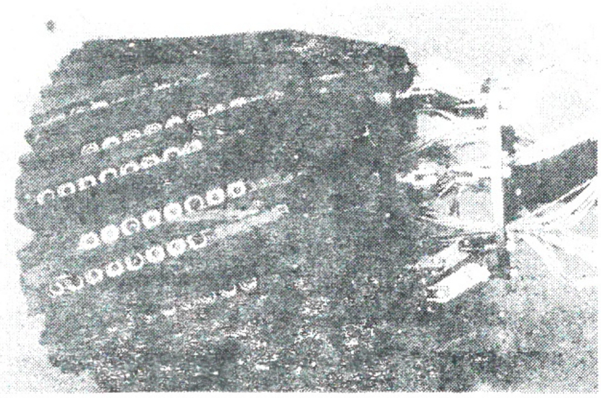

Figure 2: Ultrasound Sensor Carrier of the Pipe Pig. Based on the Diameter of the Pipe, There Are Up To 512 Mounted Ultrasound Sensors

Table 1: The Defect Classes Used in NeuroPipe

Defect type	Description
Metal loss	usually related to internal or external corrosion
Lamination	plane material separation mostly located in mid-wall position, manufacturing-related
Dent	geometric anomalies due to an external force applied to the pipe
Deposit	pollution, e.g. corrosion products, wax
No error	anomaly, but no defect

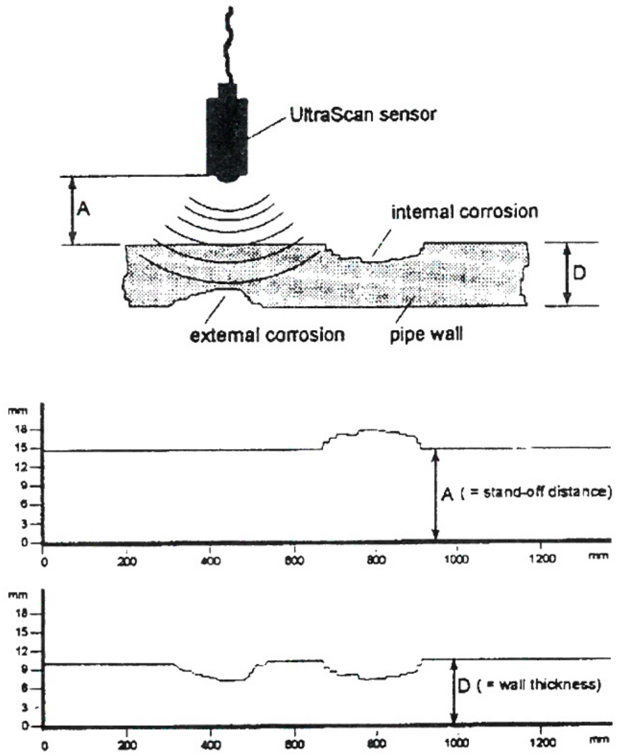

Figure 3: The Ultrasound measurement sensor. In the upper part of this Figure, the position of the sensor is outlined along with the corresponding anomalies in the wall. With the help of the B-scan, which shows the readings of a sensor over the distance (see middle and lower parts), the course of a defect can be tracked along the length of the pipeline. In the center of the illustration the corresponding stand-off distance is shown and in the lower part you see the appropriate wall thickness.

2 System Architecture

The system architecture can be divided into preprocessing, weld detection and extraction, anomaly search, defect classification and postprocessing.

The first step of the preprocessing is the decompression and the normalisation of the data measured by the pig. The main responsibility of this stage is the filtering of the sensor data. This process eliminates any small fluctuations as well as implausible or stray values. The implementation was done with the aid of standard image processing algorithms.

The task of the weld detection process is the localisation of any girth, longitudinal and spiral welds. The detection of defects inside of welds is not possible because of the physics of measurement. Therefore, these regions could not be analysed. They are located by different neural networks and excluded so that they are not treated by mistake as anomalies. Backpropagation networks are used to detect longitudinal welds and a modification of Kohonen networks for the determination of the shape of girth welds (see [14]).

All areas in which wall thickness or standoff distance sensor values deviate from specifications are detected as anomaly regions. After location, it is essential to separate closely bordering areas from one another. This makes sure that the subsequent classification of the anomalies get only one potential defect in a region. This guarantees that the decision between two defects, which lie very close to one another, by overlooking possibly more serious defect is incorrect. These areas – located and separated in such a way – are subsequently described as defect candidates.

Due to the high resolution of the utilised ultrasound technology, however, large numbers of additional anomalies are generated. They are caused by welds or internal sound reflectors (for example inclusions, slags). These indications are largely eliminated by the above- mentioned weld search and exclusion algorithms as well as by other preprocessing routines. Nevertheless, a number of these defect candidates remains for the subsequent classification.

The next module of the system architecture is the hybrid neural classification of the defect candidates. This part of the inspection system will be described in detail in the next section.

In the postprocessing step those anomalies which are classified as serious defects are stored in a database. For each of these defects, features like depth, remaining wall thickness, volume, etc. are calculated. These determined values are used for evaluation of the pipeline condition.

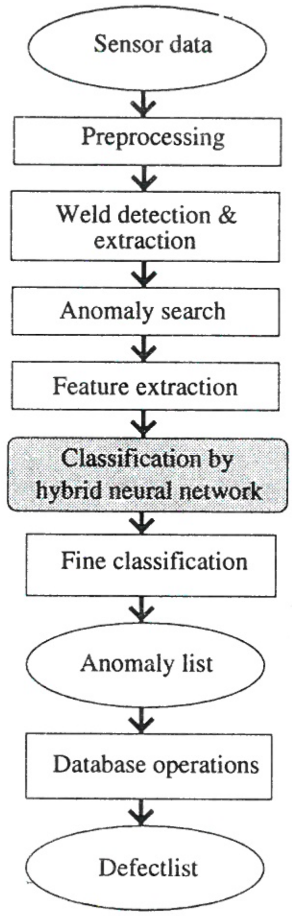

Figure 4: Data Flow of the NeuroPipe System

3 The Neural Classification Process

In a preliminary study on defect classification in pipelines (see [12, 13]), two approaches were examined and compared. The first approach was based on the raw data. For this purpose rectangular areas of a fixed size were selected from defect candidates and as a whole used for input into a neural network. The second approach established a constant number of 18 characteristic values from the defect candidate which were again classified by a neural network. Because of the varying geometric expansion of the defects size independent features must be applied as a basis for a classifying network.

The results of the study made it clear that only the approach regarding the classification of characteristic values led to useful results. The reason for this lies in the varying geometric expansion of the defects. The sizes of the rectangular regions of these defects range from a small number of readings up to several millions. Because of this large variation it is not possible for all the readings of a defect candidate to serve as an input for the neural network. Through the selection of a subregion essential information can be lost. Another way may be to interpret the regions as temporal sequences and use this representation for the classifier. This would increase the complexity of the problem too much.

In the NeuroPipe system therefore, a constant number of 86 characteristic values is determined by the defect candidate (e.g. wall thickness/stand-off data) which then serves as input to the neural networks. For example, the first and the second derivatives in the x,y direction, histograms representing the distribution of the readings on the wall, and statistical terms, related to the readings belonging to these raised values, are used.

The composition of the characteristic values was initially carried out arbitrarily since only vague expressions over the significance of individual characteristic values for classification were available. 1,400 correctly classified examples were available which were obtained through an extremely time-consuming, manual classification process (approx. 1/2 year). No particular representative set of data emerged since the "serious" characteristic values were not so frequent. Pipetronix, however, required a high quality classification.

The collection of a representative data set is very expendable. The supervisor must select learning examples which are currently not classified properly and of manifold type. He does not know how important different misclassified examples are and if it is possible to reduce the misclassification rate by them. Neural nets would be taught with all examples to achieve the best possible classification rate. If the new examples are not representative i.e. descriptive enough, this iteration process will be very time consuming.

In order to improve the generalisation of the network, the number of inputs was reduced according to the following procedure:

1. The neural network was trained.
2. Gradually each characteristic value which had not contributed to the classification was deleted from the input samples.
3. If there are deleted dimensions continue with step 1.

This way the input vector was reduced from 86 to 41 dimensions.

To get a better idea of the structuring of the input space, and as a basis for the hierarchical setup of the neural networks (see [2]), a two-dimensional Kohonen map was prepared from $30 * 30$ neurons (see [4, 9]). Figure 6 shows the 900 neurons. Each of them is assigned to all classes which activated it. Since it is possible that a sample can be a member of several categories, the necessity for multiple memberships of a neuron to different classes is obvious. This experiment shows that there are regions on the map which homogeneously belong to one class and complex regions with interwoven margins.

As shown in Figure 6, there are large homogeneous areas which belong to only one class. The idea of a hierarchical network comes from the fact that more than 50% of the defects can easily be separated by using a simple interim network from such areas with a more complex structure (border areas on the map). The hierarchy of the classification is set up in this case by using an RBF network and 5 feedforward networks, for each class of a separate net (see Figure 7).

First it is attempted to classify a possible defect by means of the modified RBF network (see [5, 6]). If the RBF network is not responsible for the defect classification (there exists no RBF neuron, which is active) the five backpropagation networks are called upon.

The advantage is that the five specialized feedforward networks only concentrate on the data that are located in more complex interwoven margins at the area boundaries. The advantage, in comparison to one network with 5 output classes is that each neural network should only have the necessary knowledge for the classification of one class, without reference to other classes. This idea proved to be correct due to the fact that the complexity of the network topology varied strongly, according to different classes.

For the modified RBF network, a rectangle function was used as an activation function. Binary weights were used between the RBF layer and the output layer (see [5, 6]). The neurons were positioned by a clustering algorithm and their radius chosen in such a way that the clash of having two neurons from different classes would never occur.

All of the feedforward networks which were used for these tasks learned with the help of the RPROP-algorithm (see [7, 8]), which as with normal backpropagation process (see [10]) was based on the decline in the weight gradients. The advantage of RPROP over the backpropagation process lies in rapid convergence through the use of the second derivative like [3]. Furthermore, the RPROP algorithm requires no learning parameter.

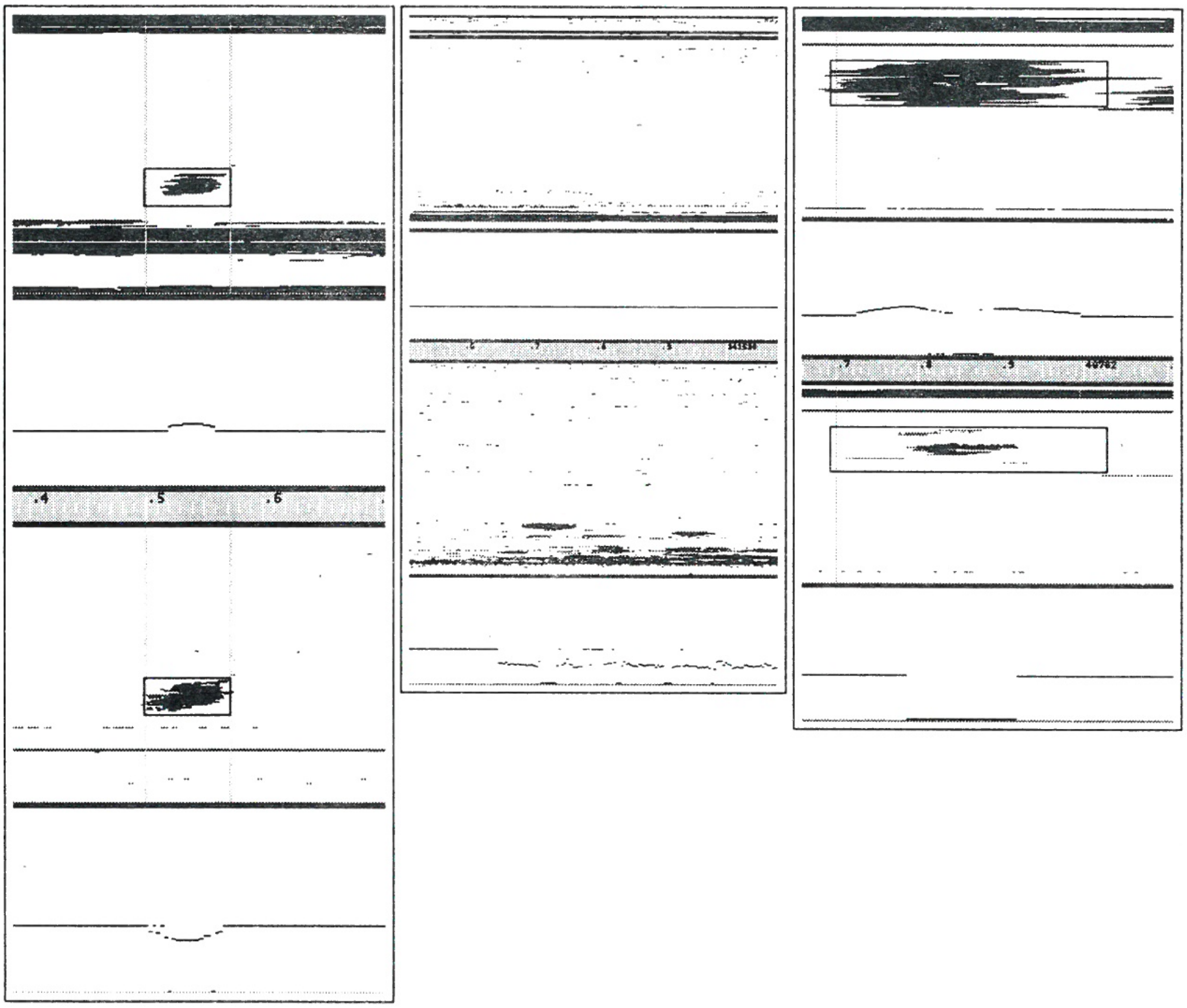

Figure 5: B-scan and C-scan show metal loss, lamination and dent. Each image is divided into two main areas. The upper half shows the stand-off distance, the lower half the wall thickness. Each area consists of two parts. The upper part shows the C-scan described above where white is used for the nominal values of the stand-off distance and the wall thickness respectively; the lower part contains the B-scan of one selected sensor.

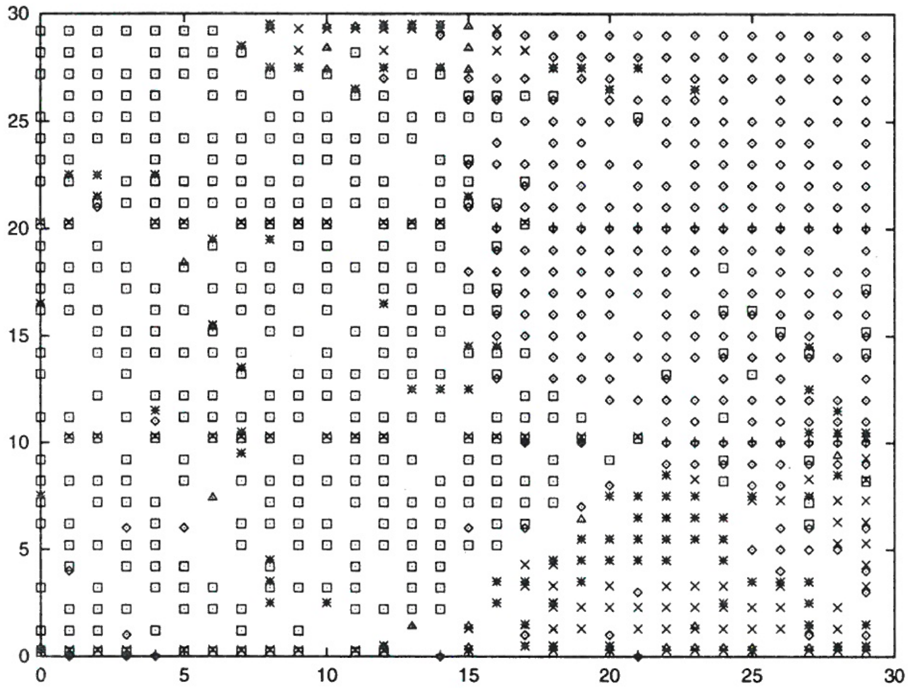

Figure 6: Two-dimensional Kohonen network trained by using 41-dimensional samples. The individual classes are described by different symbols. There where several symbols are activated on a neuron, the ambiguity of individual neurons becomes visible.

Using this hierarchical classifications method about 96% of all classifications were correct while the quality of mixed classifications[1] also showed a remarkable improvement.

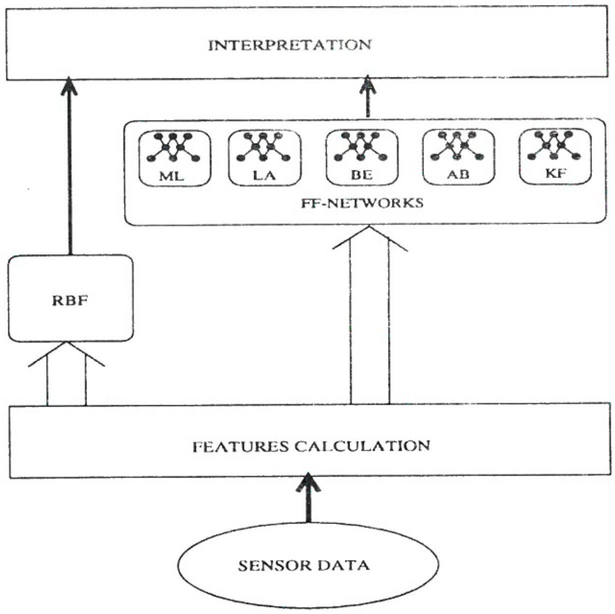

Figure 7: The Architecture of the Hierarchical Classification Network Using One RBF and Five FF Networks

4 Experience with the Iterative Learning Process

One of the major problems is the absence of known rules for the classification process. The description of the classes is on the mind of the analysts. It was not possible for them to describe the relevant class features. All descriptions we had were some high level features. For example "if the shape of the B-scan in the wall looks like this then it is a metal loss". But there were also some laminations with the same B-scan shape. The problem is that the image processing of the human brain is powerful enough to fade out all irrelevancies.

Because of this absence of rules and supported by unfixed experiences we had to develop our own features. So, these features were designed without any relation to the application and with no idea of what might actually happen to these features when they were to be classified. This heuristic procedure is the first step of uncertainty in the process because many different configurations are possible.

[1] Some areas of the pipeline wall consist of multiple defects. For such candidates a multiple class affiliation exists.

One of the effects of this procedure was that the features seemed to be a bit strange for the humans who generally worked on the recognition of the defects because the network used other features than the humans did. Another effect of this was that the controlling of the learning results of the network could not be done so easily. For example to detect which learning examples were needed for better training of the networks. Now one problem arises, namely that defects which are similar for a human being may not be similar for the classifier. This problem of measuring similarity influences the selection of learning examples.

Because of different points of view of the classifier it is impossible to predict which are the examples of defects necessary for improving the results of the system. To collect such learning examples high manual effort has to be made. As written above the connection of the features with the chosen learning examples depends on one another in an extremely high manner. Last but not least the neural network itself depends on the features and examples.

To improve the classification results it is always necessary to consider different points and adjust the parameters. But the major difficulty is to detect the right component for such an adjustment.

As one can easily recognise, the whole learning process acts like a reinforcement learning process. The classification rate of NeuroPipe acts as an estimation function. The developer decides which of the modules feature extraction, learning examples or neural networks should be changed. This iterative process is very time - consuming and highly application dependent.

In the case of NeuroPipe there have been many loops of such an iterative process. Given this it is necessary to automate this process for a general case.

5 Practical Application of NeuroPipe

The presented system has been used commercially since July 1994 by the Pipetronix company. One of the applications was a large project collecting some 90 GBytes of data from a pipeline several hundreds kilometers long. The following statistics obtained from a section of 10 km gives an impression of the amount of the collected data. However, this figure essentially depends on the diameter of the pipeline, the quality of steel and the state of the pipeline regarded corrosion. In addition, it depends on client's specifications such as reporting threshold or minimum interesting length of the individual defect types. Generally, a classification rate of much more than 90% is reached which significantly facilitates the work of the analyst.

The NeuroPipe program running on a Sparc 20 analyses about 20 times faster than a human can do. The tool never gets tired and makes no inadvertence mistakes. The quality of the classification is always the same. As NeuroPipe finds out all anomalies, classifies them and stores the pipeline position, the analyst can concentrate himself on more dangerous defects which are reviewed manually. Through the use of NeuroPipe the required time for the analysis of a whole pipe with review by the analysts was reduced by a factor of three and the accuracy of the defect lists is now higher than before.

Total number of entries	75,000
Number of entries after database postprocessing	4,000
Correct classification rate (incl. multiple significant indications of different neural networks)	appr. 95%
Features identified due to plausibility check	appr. 2-3%

6 Summary and Outlook

It has been successfully shown that it is possible to classify pipeline defects with the help of neural networks and that good results are obtainable by using hierarchical hybrid networks. This has also been confirmed by the actual deployment of the system. The used networks were taught using 1,400 example defects. This was carried out in three stages with increasing classification rates. Summing up, one can emphasise the fact that the defect classification success rate is much more than 90%. This was achieved by a skillful, hierarchical arrangement of different networks.

As described above, the definition of features for the classification task of an application is very difficult. It is highly related to the selection of learning examples. In the case of the NeuroPipe system, the recording of the examples was very time- consuming. It was not possible to decide which of the defects was learned well, because the internal representation through the calculated features was extremely different from the human point of view.

We are now concentrating on the development of a learning system (connectionist model simulator *CMS*) which also uses the neural networks for the feature extraction and learning examples selection tasks. The *CMS* should easily be integrated into different applications as an interactive learning component. Through the reinforcement learning a stepwise developing of a classifier from the learning of features up to the whole classifier will be possible. An improvement can be achieved by analysis and presentation of the system knowledge.

Acknowledgments

The work described here was carried out at *Research Center for Computer Science at the University of Karlsruhe*, in close co-operation with *Pipetronix GmbH*. We would like to specially give our thanks to Professor Rudiger Dillmann who supported this work as leader of the Interactive Planning Group, and to Klaus Germerdonk from Pipetronix for his co-operation in this project.

REFERENCES

[1] BERNS, K. and KOLB., T., **Neuronale Netze für technische Anwendungen**, SPRINGER -VERLAG, 1994.

[2] CHO, S.-B. and KIM, J.H., **Strategic Application of Feedforward Neural Networks to Large-Scale Classification**, COMPLEX SYSTEMS, No. 6, 1992, pp.363-389.

[3] FAHLMAN, S.E., **Faster - Learning Variations on Back - Propagation: An Empirical Study**, Proceedings of the 1988 Connectionist Models Summer School, MORGAN KAUFMANN, 1988.

[4] KOHONEN, T., **Self-organized Formation of Topologically Correct Feature Maps.** , BIOLOGICAL CYBERNETICS, No. 43, 1982, pp. 59-69.

[5] MOODY, J. and DARKEN., CH., **Fast Learning in Networks of Locally-tuned Processing Units.** , NEURAL COMPUTATION, Vol. 1,1989, pp. 281-294.

[6] REILLY, D., COOPER, L. and ELBAUM, C., **A Neural Model for Category Learning**, BIOLOGICAL CYBERNETICS, 45, 1982, pp. 35-41.

[7] RIEDMILLER, M., **Schnelle adaptive Lernverfahren für mehrschichtige Feedforeward Netzwerke -vergleich und weiterentwicklung-**. Master's Thesis, Universität Karlsruhe, February 1992.

[8] RIEDMILLER, M. and BRAUN, H. ,**A Direct Adaptive Method for Faster Back - propagation Learning**: The

RPROP algorithm, International Conference on Neural Networks, The Institute of Electrical and Electronics Engineers, San Francisco, USA, 1993, pp. 586-591.

[9] RITTER, H., MARTINETZ, T. and SCHULTEN, K., **Neuronale Netze - Eine Einführung in die Neuroinformatik selbstorganisierender Netzwerke**, ADDISON-WESLEY PUBLISHING COMPANY, 2nd edition, 1991.

[10] RUMELHART, D. E. , MCCLELLAND, J.L. and the PDP Research Group, **Parallel Distributed Processing**, Vol. 1 and Vol. 2 of Explorations in the Microstructure of Cognition, THE MIT PRESS, Cambridge, MA, 1986.

[11] SUNA, R., BERNS, K. and GERMERDONK, K., **Pipeline-diagnosis with Hybrid Neural Networks**, COMADEM, Vol. 8, Department of Mechanical Engineering, Queen's University, Kingston, Ontario, June 1995.

[12] SUNA, R., BERNS, K., GERMERDONK, K. and BARBIAN, A.O., **Pipeline Diagnosis Using Backpropagation Networks**, NEURO-NIMES, Nmes, October 1993, pp. 351-358.

[13] SUNA,R., **Neuronale Netze für die Klassifikation von Defekten in Pipelines**, COWAN, October 1994, pp. 84-94.

[14] SUNA, R. and EISENHAUER, P., **Schweissnahtverfolgung mit Kohonennetzen** in W.G. Kropatsch and H. Bischof (Eds.) Mustererkennung, No. 16 , DAGM Symposium, July 1994, pp. 601-608.