Using Self Organizing Maps to Visualize Age Related Changes in Lumbar Vertebrae and Intervertebral Discs

Atif Ali Khan*[1], Daciana Iliescu[1], Evor Hines[1], Charles Hutchinson[2], Robert Sneath[3]

[1]School of Engineering, University of Warwick, [2]Warwick Medical School, University of Warwick
[3]University Hospital Coventry and Warwickshire, NHS Trust, Coventry, United Kingdom
{Atif.Khan; E.L.Hines; D.D.Iliescu; C.E.Hutchinson}@warwick.ac.uk, robert.sneath@uhcw.nhs.uk

Abstract

A human spine is a complicated structure of bones, joints, ligaments and muscles which all undergo a process of change with age. This paper describes the use of artificial intelligence in visualization and better understanding of the progressive and degenerative changes in human lumbar spine. Visualizing this pattern of change will be helpful in finding the correlations among the spinal features and understanding of how a change in one feature affects others. The self-organizing map (SOM) is an efficient tool for visualization of multidimensional numerical data. It is capable of projecting high-dimensional data onto a regular, usually 2-dimensional grid of neurons. In this paper, SOM is used to visualize the pattern of change in lumbar spine features with the varying age. The paper gives an idea of how the information can be acquired from SOM representations and how the SOM can be best utilized in exploratory data visualization. Data from the lumbar spine MRIs of 61 patients (both male and female) were used in this study. The age of patients ranged from 2 to 93 years. Information for vertebral height, disc height and disc signal intensities were recorded from the MRI scans. SOM then transformed the larger feature space to a smaller one for getting a more meaningful relation between the spinal features. Complexity is reduced and the data set is represented in the form of 2D map which is easier to understand and provides visual description.

I. Introduction

A human spine is a complicated and key component of human being. During the normal ageing process, spine undergoes progressive and regressive changes which presumably follow certain pattern. This research focuses explicitly on the study of progressive and degenerative changes occurring in the human lumbar spine with the normal ageing process. This research work concentrates on the identification and classification of age-related variations in "human spine" with the help of Magnetic Resonance Image (MRI). These scans of the lumbar spine area belong to patients of different age groups. Back pain is usually associated with the spine disorder. It is the second most common reason for visits to the doctor's clinic, outnumbered only by the upper-respiratory infections [1, 2, 3]. Back pain is one of the most common reasons for missed work too. One-half of all working Americans admit to having back pain symptoms each year [1, 4]. Before finding the specific cause of back pain, it is important to study the variation of spinal features with age first and their correlation with one another [5]. These variations and correlation among the features were studies using the data collected from University Hospital Coventry and Warwickshire, United Kingdom in the form of magnetic resonance images (MRI) of the lumbar spine. Scoring of features (feature selection and extraction) was done under the expertise of an orthopedic surgeon and radiologist. The model was designed and built using self-organizing maps.

A self-organizing map (SOM) is a special type of artificial neural network which is trained using unsupervised learning to produce a low-dimensional (typically two-dimensional), discretized representation of the input space of the training samples, called a map. Unlike to other artificial neural networks, self-organizing map uses a neighborhood function to preserve the topological properties of the input space [6]. One of the most interesting aspects of SOMs is that they learn to classify data without supervision. With this approach an input vector is presented to the network (typically a multilayer feedforward network) and the output is compared with the target vector. If they differ, the weights of the network are altered slightly to reduce the error in the output. This is repeated many times and with many sets of vector pairs until the network gives the desired output. Training a SOM requires no target vector and learns to classify the training data without any external supervision whatsoever [7]. To study the variations and correlation of the spinal features, a model was built which assigns patient a certain cluster to which he/she resembles the most on the basis of his/her spinal scores. This will help spine specialists to rank and categorize patients on the basis of their spinal scores. This research work will provide a better overview to the spine specialists and the patients about the abnormal behavior if any shown by their spine.

II. Lumbar Spine

A human spine consists of bones, joints, ligaments and muscles. There are a total of 33 vertebrae in the human spine: 7 in the neck (cervical region), 12 in the middle back (thoracic region), 5 in the lower back (lumbar region), 5 that are fused to form the sacrum and the 4 coccygeal bones that form the tailbone [8]. The anatomy of human spine is shown in the figure 1 below. The focus of this research work is to look at the age related changes in the lumbar spine area. This lumbar spine area consists of vertebrae L1, L2, L3, L4, L5 and intervertebral discs between these vertebrae.

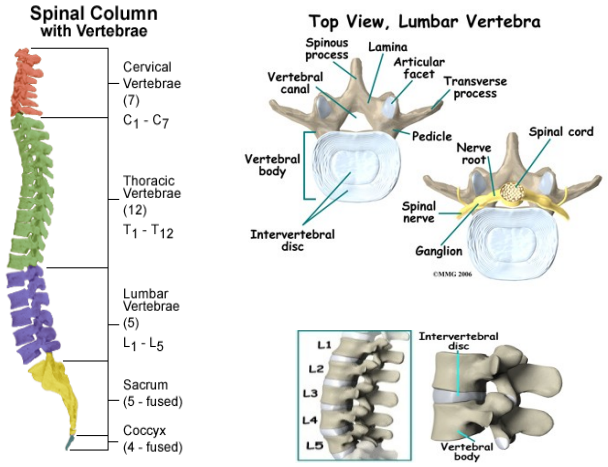

Figure 1: Anatomy of human spine

III. Data Set

The data set used in this research was taken from the University Hospital Coventry and Warwickshire (UHCW), United Kingdom. The raw data is in the form of Magnetic Resonance Images (MRI) specifically of the lumbar spine area. The format of data is Digital Imaging and Communications in Medicine (DICOM). These magnetic resonance images are the actual scans of the patients. Figure 2, shows the lumbar spine MRI.

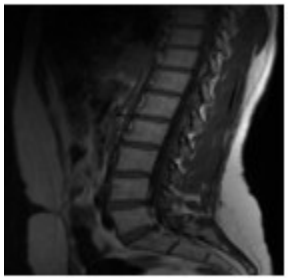

Fig 2: Sagittal and an axial view of the lumbar spine MRI

Information associated with each MRI scan is the age and gender of the patient which is used for SOM modeling. MRI scans of 61 patients were selected to develop an initial model. Age and gender distribution of patients are shown in the Table 1 below. Ten groups were formed on the basis of age decades as: G0: up to 10 years, G1: 11-20 years, G2: 21-30, years, G3: 31-40 years, G4: 41-50 years, G5: 51-60 years, G6: 61-70 years, G7: 71-80 years, G8: 81-90 years and G9: 91 and above years of age).

Table I. Age wise clustering of the samples

Age Group	Age (years)	Female	Male	Total
Group 0	10 and younger	3	1	4
Group 1	11 to 20	2	4	6
Group 2	21 to 30	4	2	6
Group 3	31 to 40	3	3	6
Group 4	41 to 50	3	2	5
Group 5	51 to 60	4	2	6
Group 6	61 to 70	6	2	8
Group 7	71 to 80	4	4	8
Group 8	81 to 90	4	3	7
Group 9	91 and over	2	3	5
	Total	35	25	61

There are lots of notable features which can be studied from a lumbar spine MRI scan. The scoring criteria were set to look initially on the vertebral height (L1, L2, L3, L4 and L5), disc height (T12-L1, L1-L2, L2-L3, L3-L4, L4-L5 and L5-S1) and disc signal (T12-L1, L1-L2, L2-L3, L3-L4, L4-L5 and L5-S1). These 17 spinal features were used as an input to build and test the initial model. These features were measured and recorded from the lumbar spine MRIs of 61 patients.

Table II. Extracted features of 5 samples from lumbar spine MRIs

		1	2	3	4	5
Gender	m/f	f	f	m	m	f
Age	Years	8	23	40	68	89
Vertebral height	L1	16.94	22.82	27.16	23.95	21.7
	L2	17.34	22.98	27.16	23.57	22.06
	L3	16.8	24.57	26.08	23.53	21.94
	L4	17.34	24.65	27.85	23.53	21.33
	L5	17.22	25.94	27.25	23.95	19.11
Disc height	T12 L1	5.95	7.51	9.33	9.48	4.45
	L1 L2	7.43	9.92	11.41	12.13	6.3
	L2 L3	7.75	10.22	13.05	13.27	5.35
	L3 L4	8.34	10.84	12.67	15.15	4.69
	L4 L5	8.0	9.06	11.83	15.41	7.15
	L5 S1	6.84	11.13	7.71	10.74	5.35
Disc Signal	T12 L1	272.4	132.5	189.4	138.8	61.9
	L1 L2	268.6	126.1	180.8	127.9	69.6
	L2 L3	255.1	123	185.2	120.2	43
	L3 L4	307.6	104.4	208.7	129.9	75.1
	L4 L5	263	95.3	138.4	137.6	67.2
	L5 S1	260	109.3	52.6	57.4	89.6

IV. Methodology

Self-organizing maps (SOMs) are a data visualization technique invented by Teuvo Kohonen which reduces the dimensions of data through the use of self-organizing neural networks. As the humans simply cannot visualize high dimensional data so this technique was created to help us understand high dimensional data. The way SOMs go about reducing dimensions is by producing a map of usually 1 or 2 dimensions which plot the similarities of the data by grouping similar data items together. So SOMs accomplish two things, they reduce dimensions and display similarities. The proposed model has a set of 17 input vectors arranged as columns in a matrix. SOM groups or ranks each sample (patient) on the basis of similarities in their 17 features and assigns certain location to each sample in the map. Figure 3 below; shows the step by step demonstration of the methodology used.

Figure 3. Steps involved in modeling

V. Self-Organizing Maps

Self-Organizing Map (SOM) is a data visualization technique which helps to understand high dimensional data by reducing data dimensions and displaying similarities among data. According to Teuvo Kohonen; the self-organizing map (SOM) is a new, effective software tool for the visualization of high dimensional data. It converts complex, nonlinear statistical relationships between high-dimensional data items into simple geometric relationships on a low-dimensional display. As it thereby compresses information while preserving the most important topological and metric relationships of the primary data items on the display, it may also be thought to produce some kind of abstractions [9].

SOM contains two processes: training and mapping. In training process, it constructs the map using input samples. After the training, it automatically classifiers a new input sample in the mapping process. The map consists of several neurons which associated with a weight vector that has the same dimension as the input sample and a position in the map. The neurons are arranged originally in physical positions according to a topology function, such as a grid, hexagonal, or random topology. The purpose of SOM is to detect regularities and correlations in their input, and also to recognize groups of similar input vectors [10, 11]. It can adapt their future responses to that input accordingly in such a way that neurons of competitive networks physically near each other in the neuron layer respond to similar input vectors.

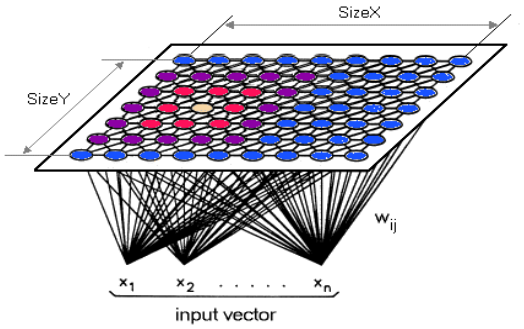

Figure 4, Structure of self-organizing map

With SOM, clustering is performed by having several units compete for the current object. Once the data have been entered into the system, the network of artificial neurons is trained by providing information about inputs. The weight vector of the unit is closest to the current object becomes the winning or active unit. During the training stage, the values for the input variables are gradually adjusted in an attempt to preserve neighborhood relationships that exist within the input data set. As it gets closer to the input object, the weights of the winning unit are adjusted as well as its neighbors [12, 13]. SOM training is shown below:

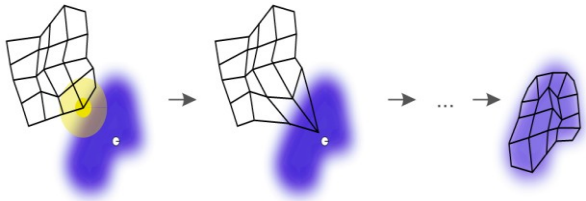

Figure 5. SOM training

The self-organization process involves four major components:

Initialization: All the connection weights are initialized with small random values.
Competition: For each input pattern, the neurons compute their respective values of a discriminant function which provides the basis for competition. The particular neuron with the smallest value of the discriminant function is declared the winner.
Cooperation: The winning neuron determines the spatial location of a topological neighbourhood of excited neurons, thereby providing the basis for cooperation among neighbouring neurons.
Adaptation: The excited neurons decrease their individual values of the discriminant function in relation to the input pattern through suitable adjustment of the associated connection weights, such that the response of the winning neuron to the subsequent application of a similar input pattern is enhanced.

SOM Algorithm:

Unlike other learning technique in neural networks, training a SOM requires no target vector. A SOM learns to classify the training data without any external supervision. Each node's weights are initialized. If the input space is D dimensional (i.e. there are D input units) we can write the input patterns as:

$$x = \{xi: i = 1, \ldots, D\}$$

And the connection weights between the input units i and the neurons j in the computation layer can be written as:

$$wj = \{wji : j = 1, \ldots, N; i = 1, \ldots, D\}$$

"N" is the total number of neurons. To determine the best matching unit, one method is to iterate through all the nodes and calculate the Euclidean distance between each node's weight vector and the current input vector. The node with a weight vector closest to the input vector is tagged as the BMU. The Euclidean distance is given as:

$$\text{Dist} = \sqrt{\sum_{i=1}^{n} [x_i - w(i,j)]^2}$$

Where x is the current input vector and w is the node's weight vector.

Network Architecture

In SOM, the network is created from a 2D lattice of 'nodes', each of which is fully connected to the input layer. Figure 6 shows a very small Kohonen network of 3 x 3 nodes connected to the input layer shown in dark blue.

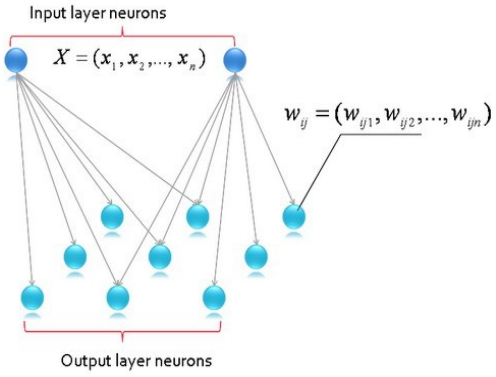

Figure 6, SOM network architecture

Each node has a specific topological position (an x, y coordinate in the lattice) and contains a vector of weights of the same dimension as the input vectors. That is to say, if the training data consists of vectors, X, of n dimensions: (x1, x2, x3...xn). Then each node will contain a corresponding weight vector W, of n dimensions: (w1, w2, w3...wn).

VI. Experimentation

The measurements taken from the lumbar MRI of 61 patients were used to model the SOM. Each patient has 17 features which were used as input to the model. These 17 input variables are vertebral heights (5 variables), disc height (6 variables) and disc signal (6 variables). So the variables 1-5 are the vertebral height (L1-L5), variables 6-11 are the disc heights from T12/L1--L5/S1 and variables 12-17 are the disc signals from T12/L1—L5/S1 respectively. The inputs vertebral heights, disc heights and disc signals have difference ranges. Initial model was built without normalization of the inputs. Figure 7, below shows the SOM model built on the basis of 17 input variables without normalization. In this mode, final quantization error was: 47.292 and final topographic error was: 0.00.

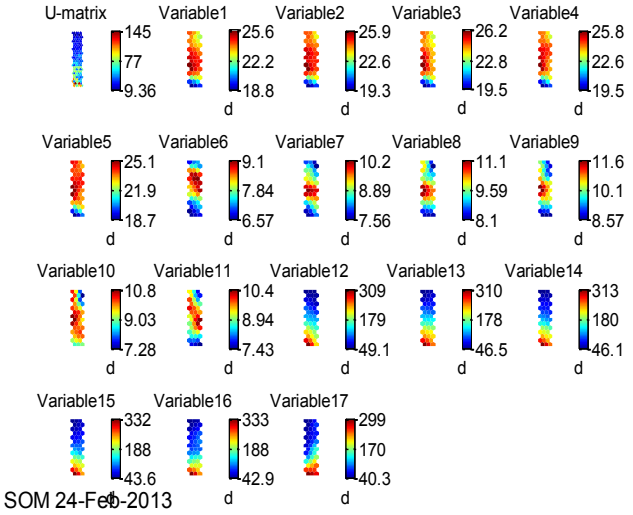

Figure 7, 17 variables SOM without normalization of Inputs

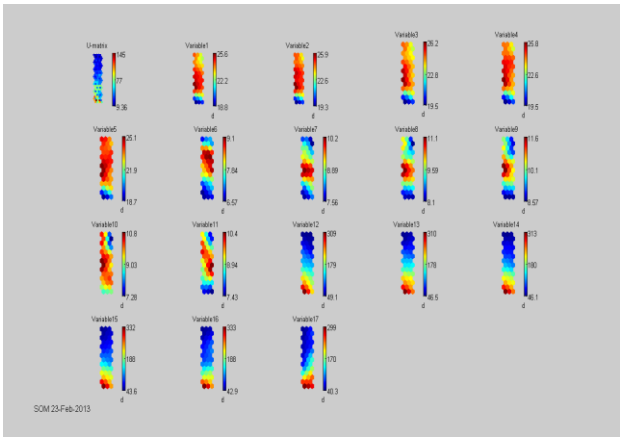

Figure 8, SOM and U-matrix without normalization of inputs

There are two separate parts of the SOM display. These include the unified matrix or U-matrix, and the component planes that are provided for individual variables [14, 15]. The U-matrix allows examination of the overall cluster patterns in the input data set after the model has been trained. [16, 17, 18] Each hexagonal cell represents individual neurons, which are the mathematical linkages between the input and output layers.

The neurons are drawn into distinct clusters during model training. Relative distances between neuron clusters are displayed by the intensity of the colors, with dark color representing greater distance [19, 20]. In the U-matrix generated here, a strong cluster is apparent, occurring in the top half (dark blue) and another one in the middle and lower middle half (light blue). This indicates that most of the input variables are covarying in one direction in n-dimensional space (where n is the number of input variables). A different trend is seen when SOM is modeled with normalized data. When the input variables are normalized, following trend was seen as shown in figure 9 below.

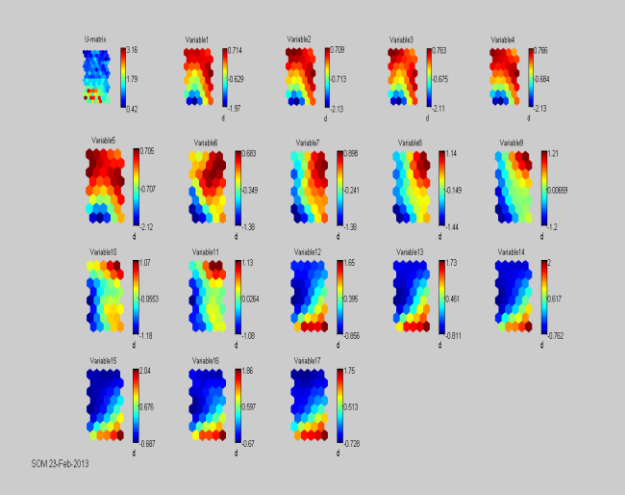

Figure 9, 17 variables SOM with normalized inputs

SOM model without input normalization showed final quantitation error of 47.292. However, by the normalization the inputs this quantization error is reduced to 1.989. The final quantization error was: 1.989 and the final topographic error was: 0.033. This shows that SOM analysis with normalized input variables provides far accurate and reliable results as compared to the results without normalization. The first map in the figure 10 below is the unified distance matrix or U-Matrix which represents overall behavior of the model. Variables 1 to 5 are the vertebral heights. Variable 6 to 11 are the disc heights and variable 12 to 17 are the disc signal intensities of all 61 patients. The color of the units (neurons) in the map shows the behavior of the specific neuron. Similar color shows that the neurons are located close to one another or similarity among the samples.

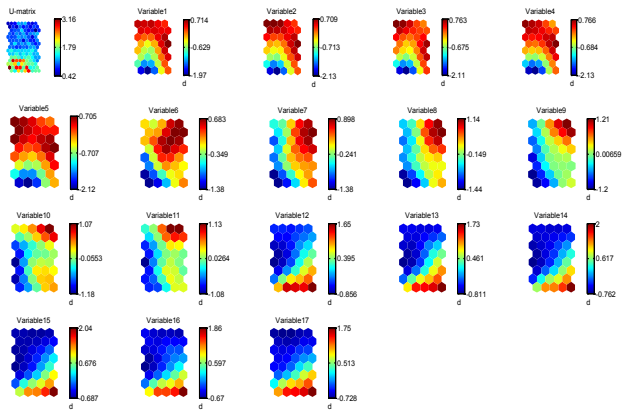

Figure 10, SOM and U-matrix with normalized inputs

VII. Results

In the component planes for individual variables, the color coding corresponds to actual numerical values for the input variables that are referenced in the scale bars adjacent to each plot. Blue colors show low values and red corresponds to high values. The relationships between each of the variables are visualized by comparing the color patterns for individual maps. In this manner, the relationships between all of the variables entered into the model can be examined simultaneously or in pair-wise combinations.

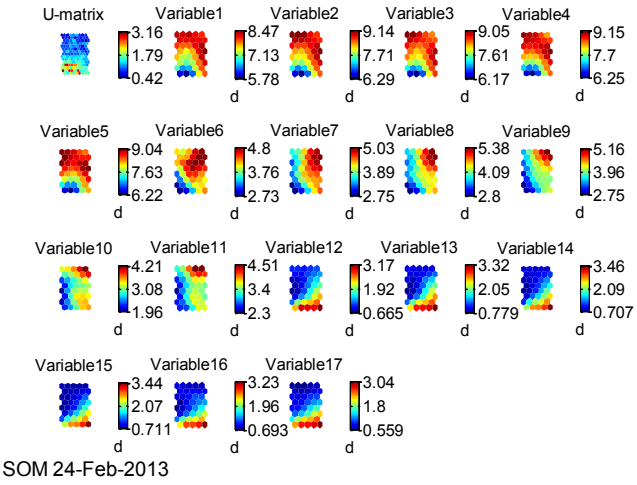

Figure 11, Visualization of SOM U-Matrix and variables

Here in figure 11 above, matching the color code of each variable with U-matrix it can be seen that vertebral heights L1, L2, L3, L4, L5 (corresponding to variables 1, 2, 3, 4, 5 respectively) do not correlate with the age (dissimilarity with U-matrix). Disc heights T12-L1, L1-L2, L2-L3, L3-L4, L4-L5 and L5-S1 (corresponding to variables 6, 7, 8, 9, 10, 11 respectively) show somewhat correlation with the age. However, disc signal T12-L1, L1-L2, L2-L3, L3-L4, L4-L5 and L5-S1 (corresponding to variables 12, 13, 14, 15, 16, 17 respectively) shows strong correlation with age.

VIII. Conclusion

The objective of the SOM analysis was to observe interrelationships that exist between 17 variables that were tested and thereby provides a basis for more advance analysis. The SOM does not replace existing statistical tools, but complements our ability to examine relationships between disparate types of variables in a visual presentation of the data. By visualizing the SOM results obtained by normalized dataset, it was concluded that lumbar spine vertebral height does not correlate with the age whereas disc height shows somewhat correlation with age. Disc signal intensities of lumbar spine show a strong correlation with the age. In future, other spinal features will be incorporated to study the spinal aging process in more depth.

Acknowledgments

This project was partially supported by Warwick Impact Fund, University of Warwick, United Kingdom. Authors would like to thank the University Hospital Coventry and Warwickshire (UHCW) NHS Trust, Coventry, United Kingdom for providing valuable data in the form of magnetic resonance images of the lumbar spine.

References

[1] Vallfors B. Acute, Subacute and Chronic Low Back Pain: Clinical Symptoms, Absenteeism and Working Environment. Scan J Rehab Med Suppl 1985; 11:1-98.

[2] Back Health at Work. HSE 2005.

[3] Back Pain Patient Outcomes Assessment Team (BOAT). In MEDTEP Update, Vol. 1 Issue 1, Agency for Health Care Policy and Research, Rockville.

[4] Department of Health Statistics Division. The prevalence of back pain in Great Britain in 1998. London: Government Statistical Service, 1999.

[5] Atif Khan, Daciana Iliescu, Evor Hines, Charles Hutchinson, and Robert Sneath, “Neural Network Based Spinal Age Estimation Using Lumbar Spine Magnetic Resonance Images (MRI)”, proceedings of 4th International Conference on Intelligent Systems, Modelling and Simulation (ISMS2013), 29-31 January, Bangkok Thailand, pp. 88-93.

[6] Kohonen, Teuvo (1982). "Self-Organized Formation of Topologically Correct Feature Maps". Biological Cybernetics 43 (1): 59–69.

[7] Kohonen, Teuvo, “Self-organizing maps (SOMs)”, Vol. 30. Springer Verlag, 2001.

[8] McKenzie, Robin A., and S. May. “The lumbar spine”, Spinal, 1981.

[9] Kohonen, Teuvo. "The self-organizing map" Neurocomputing 21.1 (1998): 1-6.

[10] Barlow, Horace B. "Unsupervised learning." Neural Computation 1, no. 3 (1989): 295-311.

[11] Carpenter, Gail A., and Stephen Grossberg, eds. “Pattern recognition by self-organizing neural networks”. MIT Press, 1991.

[12] Kohonen, Teuvo. "The self-organizing map." Proceedings of the IEEE 78, no. 9 (1990): 1464-1480.

[13] Vesanto, Juha, and Esa Alhoniemi. "Clustering of the self-organizing map." IEEE Transactions on Neural Networks, 11, no. 3 (2000): 586-600.

[14] Vesanto, Juha. "SOM-based data visualization methods." Intelligent data analysis 3, no. 2 (1999): 111-126.

[15] Vesanto, Juha, Johan Himberg, Esa Alhoniemi, and Juha Parhankangas. "Self-organizing map in Matlab: the SOM toolbox." In Proceedings of the Matlab DSP Conference, vol. 99, pp. 16-17. 1999.

[16] Ultsch, Alfred; Siemon, H. Peter (1990). "Kohonen's Self Organizing Feature Maps for Exploratory Data Analysis". Proceedings of the International Neural Network Conference (INNC-90), Paris, France, July 9–13, 1990. 1. Dordrecht, Netherlands: Kluwer. pp. 305–308. ISBN 978-0-7923-0831-7 (0-7923-0831-X).

[17] Ultsch, Alfred (2003); U*-Matrix: A tool to visualize clusters in high dimensional data, Department of Computer Science, University of Marburg, Technical Report Nr. 36:1-12.

[18] Ultsch, Alfred. "Maps for the visualization of high-dimensional data spaces." In Proc. Workshop on Self organizing Maps, pp. 225-230. 2003.

[19] Vesanto, Juha, Johan Himberg, Esa Alhoniemi, and Juha Parhankangas. SOM toolbox for Matlab 5. Helsinki, Finland: Helsinki University of Technology, 2000.

[20] Ong, Jason, and S. S. Raza. "Data mining using self-organizing kohonen maps: A technique for effective data clustering & visualization." In International Conference on Artificial Intelligence (IC-AI’99). 1999.