ISSN No:-2456-2165

Estimation of Survival Time for the Post Liver Transplantation using Hidden Markov Model

Jishnu Nair[1]
KJCOEMR, Pune

Prananjay Bhujbal[2]
KJCOEMR, Pune

Karan Nalawade[3]
KJCOEMR, Pune

Sushma Akhade[4]
KJCOEMR, pune

Abstract:- Due to increase of Technology in Medical science, there is a significant growth is seen in organ transplantation. Most of the time organ receivers life span can be predicted based on his /her health conditions. But this may not be true in case of Liver transplantation, this is mainly due to complicated working structure of the liver and its unexplainable importance in the functioning of the human body. More often the predicted lifespan of the liver transplant receiver end quite early. So machine learning can be a boon to predict accurate life span of the receiver. Some methodologies are existed in machine learning to achieve the same using deep learning techniques, where it needs heavy data for learning purpose. It is very hard to provide the such large number of data for the learning process every time. So to provide the life span expectation of liver transplanted patient, even with a moderate size of data, this research article uses the Hidden Markov model along with the Dumpster–Shafer reasoning to enhance the process of life spa expectation.

Keywords:- *Hidden Markov Model, Dumster Shafer Reaoning,K means clustering, Fuzzy Classification.*

I. INTRODUCTION

The liver is the largest internal organ in the human body. It performs a plethora of activities and it absolutely necessary for the optimum functioning of the Human body. The Liver has a lot of tasks, primarily it is concerned with the production of bile juice that aids in the digestion, it is also the organ that monitors and regulates the blood glucose levels and also maintains the glycogen and fat levels in the body. It is truly a magnificent organ. The versatility means, most of the metabolic functions are managed by the liver.

This is impacted greatly when the liver contracts a disease, it reduces its efficiency and all the functions take a hit. This is problematic as the liver manages most activities that would cease if the liver is not responding. The liver could be under attack from various diseases, such as hepatitis B and C, autoimmune disease, alcohol, etc. This leads to the poor functioning of the liver. Liver cirrhosis is one of the leading causes of liver failure among adults. It is a condition in which the liver deteriorates rapidly. The liver could also be subject to some tumors, that can be benign or cancerous that also damage the liver irreparably.

Cirrhosis is caused by excessive alcohol consumption as most of the alcohol is broken down in the liver, it leads to the parts of the liver dying gradually and over time, the scar tissue takes its place, further deteriorating health by the formation of scar tissue that blocks the blood flow to the other parts of the liver, making the condition much worse. Abstaining from alcohol and remedial measures can help ameliorate the effects and live a longer life. But not all cases are able to revert the damage and lead a healthy lifestyle.

When all of the remedies are exhausted, even after abstinence, the only viable alternative is transplantation. It is one of the most critical procedures to be performed and has a lot of parameters that need to be taken care of before proceeding with the transplant. There are countless tests that need to performed and the state of the person is taken into consideration. transplanted organs are very prone to failure and rejection, which would lead to a lot more complications and can also be fatal for the patient. This is a very critical job and it needs to be done with utmost precision.

Hidden Markov Model is one of the variants of the Markov Model. It has been used extensively in applications ranging from data compression, molecular biology and speech recognition. As the Markov Models themselves belong to a subset of Bayesian Networks, the most basic form of the Markov Model. Markov Models are quite efficient in their applications and are widely used as a statistical modeling technique for pattern recognition.

Hidden Markov Model is a subclass of the Markov Models and is primarily used to model a time series data. Due to the fact that the Hidden Markov Model is a form of Markov Model, it is a statistically generated model that is capable of predicting response. Therefore. It has been used extensively for the purpose of Artificial intelligence and more importantly Patterns Recognition.

Hidden Markov Models perform the pattern recognition and other tasks due to the Bayesian Network that it is based on, which extracts the probabilities and segregates them across a selection of observations. The observations are reminiscent of a time series data as they are discrete and are sampled at regular time intervals. Which is very characteristic of a Markov Model, from it has been derived from.

The main difference between the Markov Models and the Hidden Markov Models is due to its implementations. The Hidden Markov Model effectively hides the observations of a particular state from the observer's view, which is not the case with Markov Models. Furthermore, in Hidden Markov Models, it is assumed that the hidden states conform the conditions of being a Markov Model. Hence, the is the difference between the Hidden Markov Model and the Markov model.

Classification techniques are usually utilized on complex data to reduce its dimensionality as it groups similar items together into a set. Fuzzy classification is a classification technique that closely follows the fuzzy method, which dictates that the data elements can belong to various classes at the same time due to their fuzzy inclination to some degree.

This is highly useful and different from the crisp classification where every class has a crisp parameter which has a value of 1 or 0, or true or false. And all the data elements have a straightforward class and does not belong to more than one class at the same time. This is highly useful for measurements and other values that do not conform to a hard classification technique. This leads to a lot more accurate representation of data by the Fuzzy classification in comparison with the Crisp classification technique.

This research paper dedicates section 2 for analysis of past work as literature survey, section 3 deeply elaborates the proposed technique and whereas section 4 evaluates the performance of the system and finally section 5 concludes the paper with traces of future enhancement.

II. LITERATURE SURVEY

R.Ribeiro[1] argues thatchronic liver disease in one of the most progressive diseases in recent years due to the fact that this disease does not have any symptoms so we cannot detect this disease until it's too late. The authors have introduced a hierarchical procedure to detect t stage CLD by using the ultrasound images, laboratory tests and clinical records. When any ill health condition is found the methodology defines it in three pathologies such as 1). chronic hepatitis; 2) compensated cirrhosis 3) decompensated cirrhosis. This CBC classification scheme outperformed the nonhierarchical one, achieving an overall accuracy of 98.67%.

Xiaozhou Liintroduces a human serum Raman spectra yo analysis the of liver cancer. This spectrum is used to measure the size of the liver between normal people and liver cancer patients. There are only two types of sharp peaks such as the Raman peak, and a second one is the relative intensity of Raman peaks[2].For the Raman spectrum peaks, the patient suffering from liver cancer may detect the normal peak or the very weak low peak. Thus, the result of this methodology playsan important role to explore the method of laser spectrum diagnosis

M.Saleem states that an orthotopic liver transplantation is one of the successful and the most effective cure in the macro steatotic livers, macrosteatotic livers has become never-ending form in our nation due to continued growth in the obesity,but there is a solution, for the orthotopic liver it is essential to build or to recondition livers into lean fully functioning form for transplant. This experiment will provide the screening tools to optimize defatting protocols and the reagents to build the protocols formacrosteatosis in marginal donor organs.[3]

Y.B.Kang expresses that there has been exponential progress in the last two decades in the liver cell viability and maintenance function in vitro.the proposed the methodology has developed a novel microfluidic device using poly-dimethylsiloxane for a very long time such as a co-cultured Petri dish of primary rat hepatocytes (PRH) and endothelial cells (EC) to mimic the liver sinusoid[4].Thus, the novel liver models that closely mimic human cells have been proven to work effectively for a very long-time in a layered co-culture of PRH and EC.

E.Vorontsov[5] utilizes convolutional neural networks to connect in tandem and train the algorithm for the purpose of liver lesion classification. The results have come closer towards 2017 MICCAL live tumor Segmentation challenge attaining competitive liver and liver lesion detection and segmentation towards a wide range of metrics. The proposed method depictsa very simple way in which a single stage model is trained end to end without using any data externally according to the challenge. This method matches the top segmentation performances and achieves the highest precision.

D.Hara explains thatLiver transplantations involve a long and complicated procedure which produces significant physiological changes. This include changes in temperature, hepatic blood flow hemodynamics and acid-base. Patients coming for liver transplantation with some critical liver dysfunction also encounter protein and coagulation factor abnormalities[6]. Thus, the plasma sampling and two-compartment modeling is the response for the weighting factor for blood loss allowed and degree of instability for improved design of future controllers.

J. S. Baba elaborates on the Optimizing wavelength selection as one of the important factors to be matched or to be focused on during the liver transplantation as it requires the deep characterization of liver optical wavelengths. To complete this the authors have developed a single integrating sphere-based technique using the spatially resolved diffuse reflectance system for multispectral optical[7]. They report the result in early estimation of the monochromatic source to measure properties of a well-characterized tissue phantoms made from polystyrene spheres and Trypan blue.

H. S. Ahn expresses that liver transplantation is one of the major components,which consists of liver resection and liver transplant, as the liver may be damaged by ischemia-reperfusion injury.Till now there has not been any method

designed to measure or to predict the level of the injury during the operation.This issue of real-time monitoring makes quantification injury in the liver.Thus, this paper clears the relationship between bioimpedance and ischemia-reperfusion injury.[8]

D.Katz[9] states that Orthotopic liver transplantation is one of the definitive treatments which is the last and the end stage of the liver transplantation. The liver transplantation specifies when cirrhosis, drug-induced acute hepatic failure, malignancy, or metabolic diseases are present.While the main aim is of saving a life but there is very limited level of organs, there is a need for large numbers of organs.

Y. Mokuno introduces the concept of Liver transplantation and that it is not easy as there are many things to be improved in the liver transplantation process as it is very sensitive. Thus, the fatty livers are very sensitive to ischemia-reperfusion and cold preservation-related injuries thus this type of liver cannot be acceptable for the liver transplantation[10]. Later it was found that it is possible to eliminate excess fat storage from fatty livers by short-term perfusion of fatty livers ex vivo.These results support the notion that liver perfusion can be used to reform the fatty livers that makes them suitable for transplantation.

D.G. Melvin explains that in recent years there has been vast technical growth in the field of medical science and in the neurological field. Thus,the authors in this paper research and develop a module of neurological technology with domain as liver transplant monitoring. There is a comparison between a neural network with a static using the biomedical information acquired from the sets of liver transplant patients. The non-linear technique is depicted to be better than the linear technique. The results of the proposed paper may decide the future use of temporal information.[11]

S.Matis[12] argues that Liver transplantation is an orthodox therapeutic option for the liver transplantation patient or for an end-stage liver of the liver disease. Most probably the 20% of transplanted liver fail and half of these liver transplanting fail. If there is early prediction in liver failure or liver disease it may improve the condition during the liver transplantation that is done before the patient's condition becomes critical. By using the neural network, the accuracy of liver transplantation is improved by 5.98% of the graft survivors in the test set were correctly predicted while 88% of graft failures in the test set were also correctly predicted.

K.Suzukiexplains that Computerized liver Transplanting has been extensively researched due to the fact that current gold standards of manual transplanting is personalized and it is very time-consuming. Thus, the traditional method for the observation of the liver is done by the CT or by the MRI machine[13]. There are numbers of researchers that are working on the computerized segmentation of the liver in CT. but there are very few researchers who are working on MRI. The purpose of the proposed methodology is to develop a general framework of the liver transplantation with both MRI and CT. The computerized liver segmentation framework gives an efficient and accurate way of measuring liver in both CT and MRI automatically.

III. PROPOSED METHODOLOGY

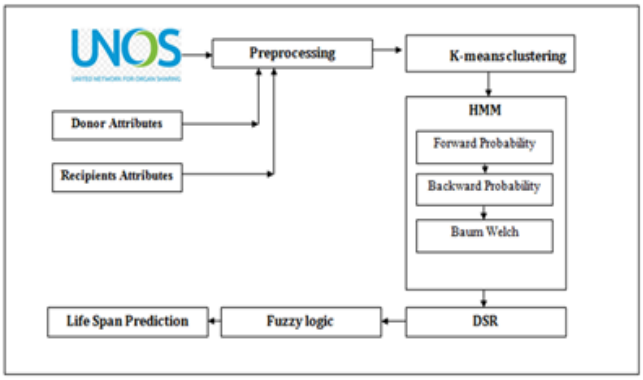

Fig 1:- Proposed Model Overview Diagram

The proposed methodology of life expectation prediction of liver transplantation patient can be seen in Figure1 and it is explained deeply with the below mentioned steps.

A. *Step 1: Data Collection and Preprocessing*

This is the initial step of the proposed model where the data are being collected according to the protocols of UNOS (United Network for Organ Sharing) for both the Donor and Receiver. The proposed model uses some Donor and Recipient attributes for the experiment as mentioned in the below Table 1 and Table 2.

DONOR ATTRIBUTES	DESCRIPTION
age_don	Age
clin_infect_don	Had Clinical Infection (Y /N)
create_don	Terminal Lab Creatinine
diabetes_don	History of Diabetes (Y/N)
don_ty	Type (Deceased / Dead)
gender_don	Gender
non_hrt_don	Non Heart Beating Donor
sgot_don	Terminal SGOT
sgpt_don	Terminal SGPT
tbili_don	Total Bilirubin

Table 1:- Donar Attribute Description

RECIPIENT ATTRIBUTES	DESCRIPTION
bmi_tcr	Body Mass Index
encph_tcr	Encephalopathy
exe_hcc	Hepatoblastoma
final_albumin	Recent Albumin level
final_ascites	Recent Ascites
final_bilirubin	Recent Bilirubin
final_inr	Recent INR
final_meld_or_peld	MELD / PELD
final_meld_peld_lab_score	MELD / PELD
final_serum_create	Serum Creatine
final_serum_sodium	Serum Sodium
gender	Gender
int_age	Age in Years
malig_tcr	Any Malignancies
gstatus	Graft Status

Table 2:- Donar Attribute Description

B. Step2: Preprocessing

Here in this step the most important factors which eventually plays a vital role in survival factor of the patient's post transplantation are selected in a list. For Donor some attributes are selected like CREATE_DON,SGOT_DON, SGPT_DON and TBILI_DON. And for Recipient some attributes are selected like bmi_tcr, final_albumin and final_meld_or_peld.

C. Step 3: K Means Clustering

Here this step receives the preprocessed list from the past step for both the recipient and Donor data to form the clusters based on the K means clustering technique. According to K means clustering process this step follows the following techniques to form clusters as described below.

➢ *Distance Evaluation*

Here each row of the pre processed list is subject to evaluate the Euclidean distances with all other remaining rows. This Euclidean distance will be appended at the end of the row and it is Reffered as the Row distance R_D. The Average of all this R_D yields the Distance of the complete preprocessed list referred as P_D.

➢ *Sorting*

Once the preprocessed list rows are appended with RD then this list is subject to sorting process using the bubble sort technique. This makes the list to come closer based on the distances, thereby it makes the data more qualitative.

➢ *Data Point Evaluation*

Some random points are allocated with respect to the total number of preprocessed list rows by normalizing the values.

➢ *Centroid Selection*

Based on the data points, respective rows and their R_D are selected to refer them as centroids.

➢ *Boundary Evaluation*

Here in this process boundary for all the centroids are evaluated based on the given equation 1.

$$B_D = (C_{RDi} - P_D) \ TO \ ((C_{RDi} + P_D) \quad (1)$$

Where,
C_{RDi} - *Entroid Row Distance*
P_D - Preprocessed List Distance
B_D - *Boundary*

➢ *Cluster Formation*

Based on the formed boundaries clusters are created with respect to the row distances. These clusters are formed both for the Donor and Recipient data.

D. Step 4: Hidden Markov Model

Here formed clusters are used to identify the worst possible factor for the given input by evaluating the Euclidean distance. By doing this proposed model put the strict protocols to evaluate the life expectancy in all worst conditions.

To do this a distance is evaluated with respect to the given input of the four attributes for both Donor and recipient data for each of the clusters. The cluster with the biggest distance is selected out of this, which is used to estimates the Forward and Backward Probability.

➢ *Forward and Backward probability Estimation*

Here the cluster with the biggest distance is used to evaluate the forward and backward probability. For this process mean and Standard Deviation of the Row distance R_D is evaluated for the complete selected cluster. Then the Forward Probability and Backward probability list are evaluated based on the following equations.

$$\mu = \frac{(\sum_{i=1}^{n} RDi)}{n} \quad (2)$$

$$\delta = \sqrt{\frac{1}{N}} \sum_{i=1}^{n} (RDi - \mu)^2 \quad (3)$$

$$F_{PB} = \mu - \delta \rightarrow \mu + \delta \quad (4)$$

$$B_{PB} = < \mu \rightarrow > \mu \quad (5)$$

Where
δ - Standard Deviation
μ- Mean
N- Number of Rows in cluster
F_{PB} - Forward Probability
B_{PB} - Backward Probability

Then the list with the biggest size is selected as the best probability list from both Forward and backward probability. This yields the two probability list each one from the Forward and backward probability.

➢ *Baum-welch*

Here the selected probability list is used to evaluate the some protocols by using the matrix evaluation process. Where each of the four selected attributes are used create the optimized protocols to select the best rows which eventually provides the correlated rows between the Donor and Recipient. This can be shown in the below shown algorithm1.

Algorithm 1: Baum-Welch

// Input : IP_{SET} ={ A_{T1} , $A_{T2,}$ $A_{T3,}$ A_{T4}} Input Protocol Set
// P_L [Probability List]
//Output : Probability Matrix List PM_{LIST}
Function : probabilityList($IP_{SET,}$ P_L)
Step 0: Start
Step 1: PM_{LIST}=Ø
Step 2: ***for*** i=0 **to** size of IP_{SET}
Step 3: MIN_{SCORE} =IP_{SETi} - K
Step 4: MAX_{SCORE} =IP_{SETi} + K
Step 5: ***for*** j=0 **to** size of P_L
Step 6: ***If*** (P_{Lj} >= MIN_{SCORE} **AND** P_{Lj} <= MAX_{SCORE}),**THEN**
Step 7: $PM_{LIST=}$ PM_{LIST}+ P_{Lj}
Step 8: ***END If***
Step 9: **End *for***
Step 10: **End *for***
Step 11: return PM_{LIST}
Step 12: Stop

E. Step 5: Dumpster Shaffer Reasoning and Fuzzy Classification

Here in this theory a reasonable evaluation is given for the Probability matrix list for the input attribute values given by the user. Here if the attribute values are equal, then they are assigned a count for this. This eventually helps the system to predict the life expectancy of the liver transplanted patient.

Here a probability value is evaluated based on the mid probability theory of DSR for the counted probability. Then, based on this three fuzzy crisp values are created like LOW ,MEDIUM and HIGH.

Then, based on this Fuzzy Crisp values and the DSR count, the maximum life span of 180 months for the patient is set and then they are classified by the IF THEN rules of Fuzzy Classification. Final output in terms of months is being shown to the user which was thoroughly evaluated using many bottleneck conditions.

IV. RESULT AND DISCUSSIONS

The proposed system of prediction of life expectancy in term of months for post liver transplanted patient is developed using the Windows based machine of Core i5 Processor with 6GB RAM. The system is developed using the Java Programming language using Netbeans as IDE and Mysql as the database server. To implement the system the proposed model uses the sample dataset of size 200 including for both Donors and Recipient. This dataset is created based on the protocol set by the UNOS. To measure the efficiency of the model, some experiments are conducted as described below.

The predictiveness can be measured using many parameters like RMSE, MAE ,MRR and many more. Here the percentage of error that out proposed model can commit can be measured by using the parameter Mean absolute Error (MAE) .

MAE measures the percentage of the error that our system may commit for the given input data. As this measuring parameter provides a better way to estimate the performance of the proposed model. MAE is a measure of difference between the two continues entities which in turn represents the same phenomenon.

Here in our experiment percentage of error is measured in between the Predicted life span and actual life span for the post transplantation. MAE can be represented by the given equation 6.

$$MAE = \frac{(\sum_{i=1}^{n} | xi - yi |)}{n} \quad (6)$$

Where,
xi - Correct prediction of Life Span
yi - Incorrect prediction of Life Span
n- Number of Trails

No of Input Claims	Correct prediction of Life Span	xi-yi
10	9	1
20	17	3
30	24	6
40	33	7
50	41	9
	MAE	**5.2**

Table 3:- MAE Evaluation

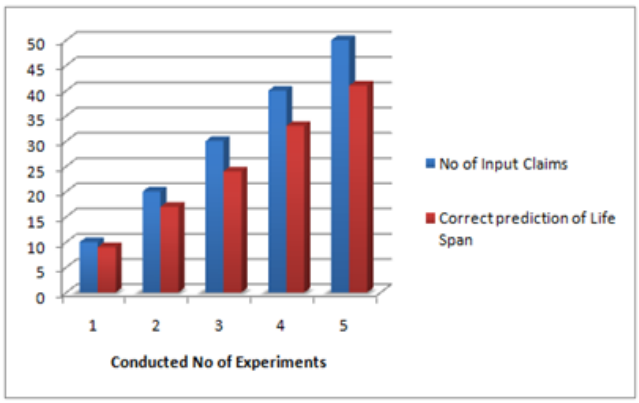

Fig 2:- MAE Evaluation Based on the Correct Prediction Values

The experiments conducted by using MAE yield average value of 5.2 and this can be seen in Figure 2 and Table 3 . This is actually a good result for such a small size of data.

V. CONCLUSION AND FUTURE SCOPE

The proposed model for prediction of life expectancy after liver transplantation is deployed using the machine learning process. Proposed model uses the moderate amount of 200 data , which are created using all the novelty rule followed by the UNOS. Proposed system uses the K-means clustering for efficient clustering of the likely data in between the Donor data and Recipient data. Once they are clustered then they are fed to Hidden Markov Model to evaluate the Probability matrix data. This is used by the Dumpster Shaffer reasoning to evaluate the proper value count, which is being used by the abstract Fuzzy classification to evaluate the Proper correct prediction of life span.

The proposed model can be improved in the future by implementing the model by considering the real time data from the hospitals with more attributes related to the medical paradigm.

REFERENCES

[1]. Ricardo T. Ribeiro∗, Rui Tato Marinho, and J. Miguel Sanches, Senior Member," Classification and Staging of Chronic Liver Disease from Multimodal Data"IEEE Transactions on Biomedical Engineering, VOL. 60, NO. 5, MAY 2013

[2]. Xiaozhou Li, Junxiu Lin, Jianhua Ding, Shu Wang, Qiang Liu1, Shumei Qing," Raman Spectroscopy and Fluorescence for the detection of Liver Cancer and Abnormal Liver Ti" Proceedings of the 26th Annual International Conference of the IEEE EMBS, San Francisco, CA, USA • September 1-5, 2004

[3]. Mubasher Saleem, Timothy J. Maguire," Combined Computational Fluid Dynamic and Metabolic Model of Liver Defatting Under Flow" 978-1-4799-3728-8/14/,IEEE,2014.

[4]. Y.B. Kang, T.R. Sodunke, J. Cirillo, M.J. Bouchard, and H. Noh," Liver on A Chip: Engineering the Liver Sinusoid" 978-1-4673-5983-2/13/, IEEE,2013.

[5]. Eugene Vorontsov, A Tang, Chris Pal, Samuel Kadoury," Liver Lesion Segmentation Informed by Joint Liver Segmentation" 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018) April 4-7, Washington, D.C., USA, 2018.

[6]. Dorene A. O'Hara, John G. Hexem, Garrett J. Derbyshire, Frank J. Overdyk, Bo Chen," The Use of a PID Controller to Model Vecuronium Pharmacokinetics and Pharmacodynamics During Liver Transplantation" 0018–9294/97,IEEE, 1997.

[7]. J. S. Baba, Member, B. S. Letzen, M. N. Ericson, G. L. Coté," Development of a Multispectral Tissue Characterization System for Optimization of an Implantable Perfusion Status Monitor for Transplanted Liver" 978-1-4244-3296-7/09/,IEEE,2009.

[8]. H. S. Ahn 1, S. S. Yun, H. J. Shin, J. Y. Kim, J. H. Choi, "Measurement of bioimpedance and cell viability during ischemia-reperfusion in the rat liver", Proceedings of the 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference Shanghai, China, September 1-4, 2005

[9]. Daniel Katz," Serious Gaming for Orthotopic Liver Transplantation" ,IEEE 28th International Symposium on Computer-Based Medical Systems.2015.

[10]. Y. Mokuno, F. Berthiaume, M. L. Yannush, "Reconditioning of Fatty Livers For Transplantation Increased Survival Of Recipients In A Rat Model" 0-7803-7612-9/02,IEEE, 2002.

[11]. D.G. Melvin, M. Niranjan, R.W. Prager, A.K. Trul1, V.F. Hughes," NEUROC0MPUTING AP P LI CAT10NS IN MONITORING. POST-OPERATIVE LIVER TRANSPLANT" 0-7803-4256-9/97/IEEE,1997.

[12]. Sherri Matis, Howard Doyle, Ignazio Marino, Richard Mural and Edward Uberbacher" Use of Neural Networks for Prediction of Graft Failure following Liver Transplantation" Eighth IEEE Symposium on Computer-Based Medical Systems.1995.

[13]. Kenji Suzuki, HieuTrung Huynh, Yipeng Liu, Dominic," Computerized Segmentation of Liver in Hepatic CT and MRI by Means of Level-Set Geodesic Active Contouring" 35th Annual International Conference of the IEEE EMBS Osaka, Japan, 3 - 7 July 2013.