International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Automatic Segmentation Using U-Net for Accurate Liver Segmentation of CT Images

S. V. Vanmore[1], S. S. Nikam[2], R. B. Dhumale[3], N. R. Dhumale[4], P. B. Mane[5], A. N. Sarwade[6]

Submitted: 08/05/2023 **Revised**: 15/07/2023 **Accepted**: 06/08/2023

Abstract*:* In this paper a high-resolution U-Net architecture for the segmentation of liver Computed Tomography (CT) scan images with high accuracy is presented. The contraction and expansion paths present in the proposed U-Net allow for reducing computational time and increase the resolution of information of the liver CT scan image respectively. The high-resolution U-Net architecture has been trained and validated using an IRCAD 2D liver CT image dataset. Performance of the U-Net architecture has been verified by calculating performance metrics of segmented 2D liver CT scan images. Results of performance metrics show that the proposed U-Net architecture achieves the best quality performance. The maximum dice coefficient of the liver segmentation during training phase is 95.79 % whereas during validation phase is 89.27 %. The high-resolution U-Net architecture has also compared with other researcher's work in this paper.

Keywords: *Liver CT images; Segmentation; Deep neural network; U-Net; Convolution Neural Network.*

1. Introduction

Liver size, texture abnormalities and liver abnormalities are important parameters for the diagnosis of disease, so segmentation of liver is significant factors in CT and are important diagnostic markers for disease progression in primary and secondary liver tumor diseases [1].Manual or semi manual techniques of segmentation have been using in clinical routines, but those techniques are subjective, operator controlled, and very time consuming. Computer-aided segmentation methods have been established to increase the efficiency of radiologists, however still some challenges in automated segmentation of the combined liver and lesions have left, such as segmentation of different sizes of the liver with varying amounts of the lesion [2][3]. Several segmentation techniques proposed that includes the intensity threshold, region growth, deformation models and multi threshold for accurate segmentation of the liver and lesion in CT images. The graph cut and level set techniques have been employed for liver and liver lesion segmentation [4]. In present years the convolution neural networks have been developed for semantic segmentation; however particular conventional artificial neural network and fully convolution neural networks have achieved extensive achievement [5]. The conventional artificial neural network has been used for liver segmentation, but this network is slow and giving imprecise results of segmentation. It requires a large set of data for training. A fully connected convolution neural network gives precise liver segmentation results. However, the network is slower, since it uses several kernel filters for retrieval of feature vectors using the sliding window technique; also it requires a large set of data for the training. To overcome the drawbacks of these networks the U-Net has proposed in this paper. The proposed U-Net network is fast, giving precise segmentation results and it required a limited set of data for training. The execution has been done on the Tens or Flow Processing Unit (TPU).

1Electronics Engineering, Department of Technology, Shivaji University, Kolhapur, 416004, Maharashtra, India
2,3,5 AISSMS Institute of Information Technology, Pune, India.
4,6 Sinhgad College of Engineering, Pune, India.
**Correspondence: shobha.nikam@aissmsioit.org; Tel.: (+91 9860709687)*

2. Data Set Information

The segmentation of the liver has carried out using CT scan images provided by the IRCAD data set. This data set is standard data set, which is used by the various researchers for their work on biomedical engineering. The IRCAD dataset of liver images consists of 3D CT scan images of ten women and ten men volunteers with age group 26 years to 75 years. The data set consists of 20 3D medical examination CT scan images. These images are NifTi format. Live resource images and mask images have been used from this data set. For image segmentation, CT scan images are read by using the in-babel python library. The 2D slices of CT scan images are extracted from 3D medical examination images. Each image size has a memory of 84 kB, with 512 rows and 512 columns. There are 2379 2D CT scan images from which 1155 2D data set for training of the proposed network have been selected. The same number of the 2D mask has been used as target

images during training. The trained network usage 289 CT scan images have been used for validation of the proposed network. However, the remaining 935 CT scan images for the testing of the proposed network. Liver CT scan images are down sampled to 256×256 resolution to reduce the computation cost. The training of the U-Net network has been performed by 1155 iterations for each epoch. There are 10 epochs, batch size=10, and learning rate equal to 0.001 in the training phase. For validation 231 livers CT images and for testing 935 liver CT scan images have been used. The overall accuracy of liver segmentation is dependent on the number of training images. As the number of training images is more the accuracy is more. The sampled training images and train masks are shown in Fig.1.

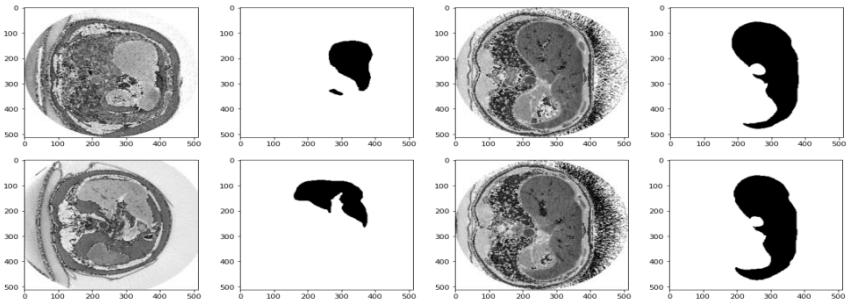


Fig.1: Sampled IRCAD liver CT images with its masks

3. Proposed U-Net Architecture

The U-Net is U shaped symmetric architecture; it consists of two parallel paths called a left path and a right path. The left path is the contraction path in which images have down sampled with the max-pooling layers replaced by up sampling layers. Up sampling layers have a large number of feature channels, which propagates information of unique features to the next layer of the proposed network. Hence the resolution of O/P segmentation results has been increased. The proposed U-Net consists of 9 layers from which 5 layers are in the contraction path and 5 layers in the expansion path. However, the bottom layer is common for both expansion and contraction paths.

Layer-wise information about U-Net Architecture has been given in Fig. 2. As discussed above, the left path of the U-net is a contraction path. The first layer of the contraction path is an input layer. Liver CT images available from the data set is of 512×512 sizes, but it down sampled by2tomakeasizeof 256×256 and applied to an I/P layer. The I/P of contraction path, there are two convolutions layers which consist of 32 filters with kernel filter size3×3using stride 1. The stride is applied horizontally and vertically with the same padding the method followed by the ReLu activation function. To maintain the size of the convolution layer as of I/P image size, the padding method has been applied to convolution layers. The purpose of adding two convolution layers is to extract the relation of information about different feature maps. At output of the 2nd convolution layer, the 2D max pooling of 2×2 filter size has been applied with stride one.

In the following layers of the contraction path kernel filter size activation function, padding methods and max-pooling technique are the same except for the number of filters in convolution layers. In the 2nd, 3rd, 4th, 5th layers number of filters inconvolutionlayers.64, 128, 256, and 512 respectively. The expansion path consists of five layers (layer No.5, layer No.6, layer No.7, layerNo.8, layer No.9). In layer number 6, the output of the 5th layer is up sampled by 2 and concatenated with O/P of the 4thlayer. The optimized channel information has been obtained from this layer. Due to up sampling, the number of feature channels has increased, and it allowed the network to propagate the inform action to high-resolution layers.

After upsampling,2 convolution layers of 256 filters of kernel filter size 3×3 with the ReLu activation function and the same padding, the method has been applied to maintain the size of consecutive layers.

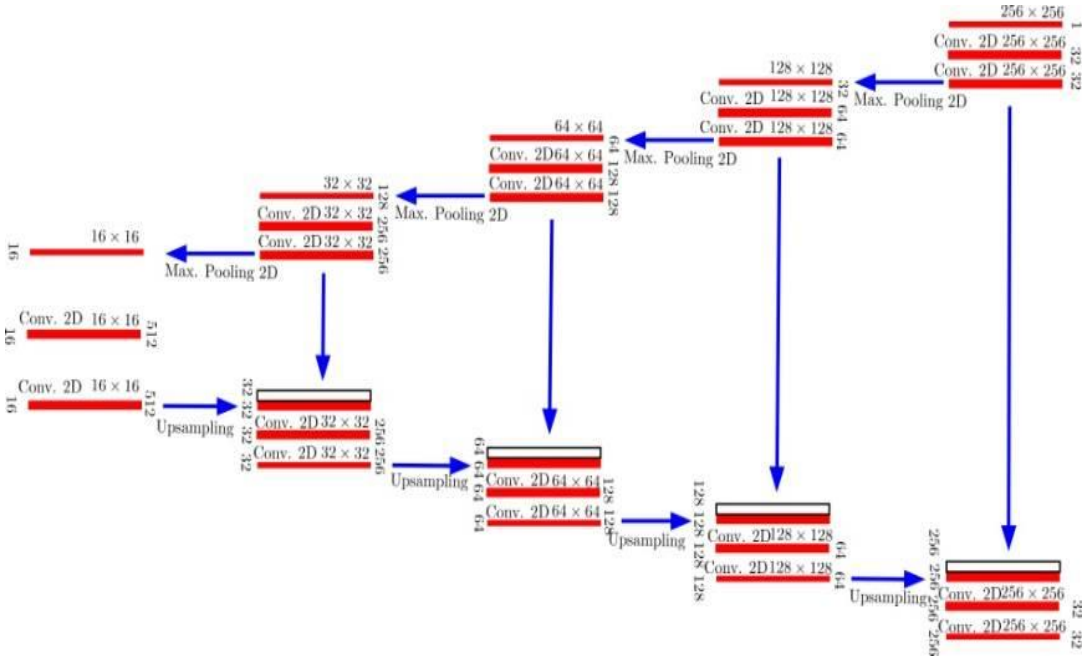


Fig. 2: The Proposed U-Net Architecture

In 11 the following layers of expansion path, kernel filter size, activation function, padding method are the same except for the number of filters. In the 6th, 7th, 8th, 9th layers the number of filters in convolution layers is 256, 128, and 64, and 32 respectively.

4. Performance Metrix of the U-Net Architecture for Liver Segmentation

Different evaluation indices of the U-net Architecture model have been obtained during the training and validation phase for the validity of Liver CT scan images. This is done for simultaneous truth and performance level estimation evaluation indices are as follows.

Dices Similarity Coefficient: This measure quantifies the match of two sets MS and GT. It is the ratio of the normalized value of the intersection of two sets over the average of two sets.

$$\text{Dice} = \frac{2|\text{MS}\cap\text{GT}|}{|\text{MS}|+|\text{GT}|} \quad (1)$$

Where MS is a machine segmented set of voxels, GT is a ground truth set of voxels [6][7].

Volume Overlap Error (VOE): It quantifies the error measure of two sets of MS and GT. VOE is defined as in [8][9].

$$VOE = 1 - \frac{2|MS\cap GT|}{|MS\cup GT|} \quad (2)$$

Relative Volume Difference (RVD): It measures the dissimilarity of two sets of MS and GT. It is defined as,

$$\text{VOE} = \pm\frac{2|\text{MS}|-|\text{GT}|}{|\text{GT}|} \quad (3)$$

Symmetric Volume Difference (SVD): It quantifies symmetric measures of the difference in two sets of MS and GT. The SVD is also used to estimate segmentation error. It is defined as,

$$\text{VD} = 1 - \text{Dice} = 1 - \frac{2|\text{MS}\cap\text{GT}|}{|\text{MS}|+|\text{GT}|} \quad (4)$$

Average Symmetric Surface Distance (ASD): It quantifies average distance from points on the boundary of MS set to the boundary of GT set. It is defined as,

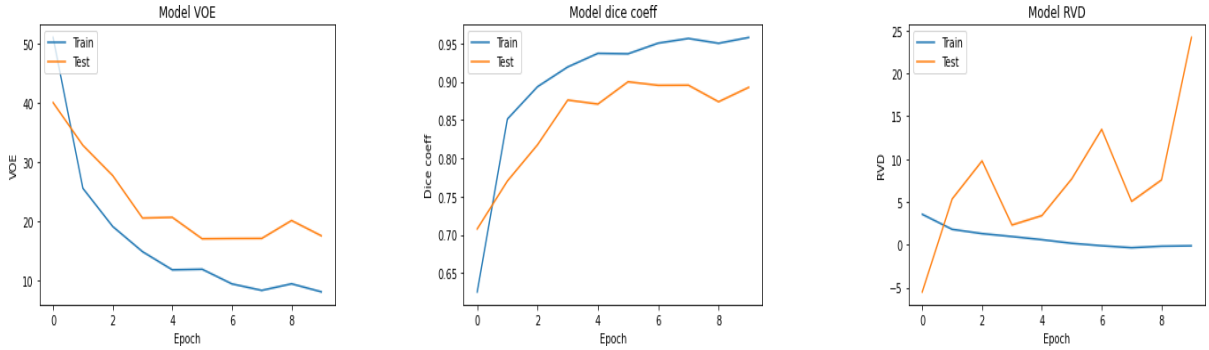
$$\text{ASD} = \frac{1}{|\text{B}_{\text{MS}}|+|\text{B}_{\text{GT}}|}\times \left(\sum_{\text{x}\in\text{B}_{\text{ms}}} \text{d}(\text{x},\text{B}_{\text{GT}}) + \sum_{\text{x}\in\text{B}_{\text{ms}}} \text{d}(\text{y},\text{B}_{\text{MS}})\right) \quad (5)$$

Where BMS is a boundary of MS and BGT is a Boundary of GT set. d(x;BGT) is the distance between the point on the boundary of MS set and the boundary of GT set. d(y;BMS) is the distance between a point on the boundary of GT set and the boundary of the MS set [10].

Jaccard: It quantifies the similarity coefficient between MS and GT is defined as [7].

$$Jaccard = \frac{|MS\cap GT|}{|MS\cup GT|} \quad (6)$$

the validation data set. In addition to DC, VOE, and RVD



(a)DSC versus Epoch (b)RVD versus Epoch (c)RVD versus Epoch

Fig.3: Graphs of segmentation performance metrics

5. Results and Discussion

Graphs of performance metrics of U-net architecture for liver CT scan segmentation during training and testing phases have shown in Fig. 3. All performance metrics have been calculated by using the nearest neighbor technique and set theory. The graph of Dice Coefficient or Dice coefficient (DC) versus epoch for the training and testing phase has been given in Fig. 3.a. As epoch and iterations increase, the dice similarity coefficient increases rapidly up to epoch 3, as shown in the figure. After epoch number 3, the dice similarly coefficient increases gradually. As shown in graphs dice similarity coefficient at the 10th epoch is 0.9579 of the training phase. However, the dice similarity coefficient at the 10th epoch of the validation phase is 0.8927. It has been observed that the dice coefficient of the training phase is greater than that of the validation phase because of the number of training liver CT scan images. The expected value of the dice coefficient has been obtained since U-net architecture has been specially designed for high-resolution segmentation.

As shown in Fig. 3.b, the graph of VOE decreases by increasing epochs. It has been shown in the graph of VOE of the training phase is less than that of the testing phase because of the number of training liver CT scan images. However, if the number of epochs and iterations increases VOE will further decrease. The graph of RVD versus epoch shown in Fig. 3.c. The RVD at the 10th epoch of training phase 0.1215, which indicates better agreement over the validation phase can be improved by increasing other performance metrics including ASD, SVD, Loss,

Jaccard have calculated during training and validation phases for 10 epochs with 1155 iterations per epoch, as shown in Table 1. Comparison of performance metrics of high-resolution U-net architecture for liver CT scan segmentation with other researcher's work has been shown in the Table 2. The performance of the proposed U-net network has compared with the graph cut method, active contour method, semi-automatic Laplacian meshes, and U-Net network as shown in Table3. However, the first three methods are the traditional method and semi-automatic.

The performance metrics of the traditional liver segmentation are better than the proposed U-Net network. The sample segmented liver CT images are shown in Fig. 4. However, the proposed U-Net network is automatic whereas traditional methods are semi-automatic, so the traditional methods require more time for liver segmentation. The proposed U-Net has also compared with the U-Net of O. Ronne bergeret al. [11]; both networks are automatic and have employed the same network. It has observed that the dice coefficient and ASD of the proposed U-Net are better than U-Net of O.Ronnebergeretal.

Table1: Performance metrics of the liver segmentation per epoch during training

Epoch	Dice Coefficient Loss	Dice Coefficient	VOE	RVD	SVD	Surface loss	Jaccard
1	-0.6243	0.6255	51.0976	3.555	0.3745	2.3887	0.489
2	-0.8512	0.8512	25.5796	1.8	0.1488	-0.4693	0.7442
3	-0.8936	0.8936	19.1265	1.302	0.1064	-0.5993	0.8087
4	-0.9192	0.9193	14.8674	0.96	0.0807	-0.6345	0.8513
5	-0.937	0.9371	11.7942	0.592	0.0629	-0.6741	0.8821
6	-0.9365	0.9365	11.8932	0.171	0.0635	-0.6602	0.8811
7	-0.9506	0.9504	9.4146	0.115	0.0496	-0.6826	0.9059
8	-0.9565	0.9565	8.3153	0.344	0.0435	-0.6897	0.9168
9	-0.9502	0.9502	9.4299	0.169	0.0498	-0.6685	0.9057
10	-0.958	0.9579	8.0707	0.122	0.042	-0.6889	0.9193

Table 2: Performance metrics of the liver segmentation per epoch during validation

Epoch	Dice Coefficient Loss	Dice Coefficient	VOE	RVD	SVD	Surface loss	Jaccard
1	-0.7019	0.7077	40.0719	-5.4949	0.2923	-0.3723	0.59953
2	-0.7714	0.7703	32.8543	5.3448	0.2297	-0.5165	0.6714
3	-0.8189	0.8177	27.7363	9.7886	0.1823	-0.521	0.7226
4	-0.8771	0.8761	20.5577	2.3074	0.1239	-0.5705	0.7944
5	-0.8713	0.8708	20.6713	3.4027	0.1292	-0.5827	0.7933
6	-0.9004	0.8999	17.0467	7.6986	0.1001	-0.6099	0.8295
7	-0.8953	0.8953	17.0897	13.471	0.1047	-0.614	0.8291
8	-0.8956	0.8955	17.1061	5.0581	0.1045	-0.5986	0.8289
9	-0.8738	0.8739	20.1336	7.5673	0.1261	-0.5873	0.7987
10	-0.8926	0.8927	17.5522	24.2	0.1073	-0.6185	0.8245

Table 3: Comparison of the liver segmentation performance metrics with other researchers work

References	Network/Method	Dice Coefficient [%]	VOE [%]	RVD [%]	ASD
Lietal.(liver-only)[12]	Graph cut		09.20	-11.20	01.60
Chartrandetal. (semi-automatic) [13]	-	60.80	01.70	01.60	-
Chan-Vese [11]	Active contour	94.70	07.12	09.11	02.49
Ronnebergeretal [14]	U-Net	72.90	39.00	87.00	19.40
Proposed	**High resolution U-Net**	**89.27**	**17.55**	**24.19**	**-00.61**

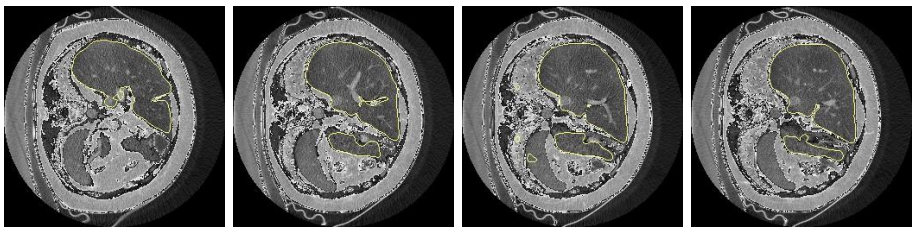


Fig.4: Sampled segmented liver CT images

6. Conclusion

This research paper has presented high-resolution U-net architecture for the segmentation of liver CT scan images. Experimental work has been done by employing IRCAD-data of liver CT scan images to the U-network. In the study, it has been observed that the conventional multilayer neural network is slow and not so accurate for the segmentation of liver CT scan images. The convolution neural network is accurate and required big data set, so it becomes sluggish. These problems have been overcome using a high resolution U-net network. Also, it has been observed that the high accuracy of the liver CT scan segmentation with moderate data set has been obtained. It is because of the up sampling of the information in the expansion path. In training phase DC=95.79%, SVD=0.0421, ASD=-0.6889, Jaccard=0.9193, RVD=0.1215 have been obtained at the 10^{th} epoch, which shows better results compare to other networks. Invalidation phase DC= 89.27%, SVD =0.1073, ASD=-0.6185, Jaccard=0.8245, RVD=24.2 have been obtained at 10^{th} epoch. However, segmentation performance metrics can be improved by increasing epochs. This network has been suitable for high precision liver lesion segmentation. The proposed system can be modified for automatic diagnosis of the liver lesion in medical applications by using a lesion mask from an IRCAD set.

References:

[1] M. Zerunian et al., "Updates on Quantitative MRI of Diffuse Liver Disease: A Narrative Review," Genet. Res. (Camb)., vol. 2022, 2022, doi: 10.1155/2022/1147111.

[2] M. Y. Ansari et al., "Practical utility of liver segmentation methods in clinical surgeries and interventions," BMC Med. Imaging, vol. 22, no. 1, pp. 1–17, 2022, doi: 10.1186/s12880-022-00825-2.

[3] H. Rahman, T. F. N. Bukht, A. Imran, J. Tariq, S. Tu, and A. Alzahrani, "A Deep Learning Approach for Liver and Tumor Segmentation in CT Images Using ResUNet," Bioengineering, vol. 9, no. 8, pp. 1–19, 2022, doi: 10.3390/bioengineering9080368.

[4] M. Ahmad et al., "A Lightweight Convolutional Neural Network Model for Liver Segmentation in Medical Diagnosis," Comput. Intell. Neurosci., vol. 2022, 2022, doi: 10.1155/2022/7954333.

[5] J. M. H. Noothout et al., "Knowledge distillation with ensembles of convolutional neural networks for medical image segmentation," J. Med. Imaging, vol. 9, no. 05, pp. 1–20, 2022, doi: 10.1117/1.jmi.9.5.052407.

[6] T. Heimann et al., "Comparison and evaluation of methods for liver segmentation from CT datasets," IEEE Trans. Med. Imaging, vol. 28, no. 8, pp. 1251–1265, 2009, doi: 10.1109/TMI.2009.2013851.

[7] W. R. Crum, O. Camara, and D. L. G. Hill, "Generalized overlap measures for evaluation and validation in medical image analysis," IEEE Trans. Med. Imaging,
vol. 25, no. 11, pp. 1451–1461, 2006, doi: 10.1109/TMI.2006.880587.

[8] X. Chen, J. K. Udupa, U. Bagci, Y. Zhuge, and J. Yao, "Medical image segmentation by combining graph cuts and oriented active appearance models," IEEE Trans. Image Process., vol. 21, no. 4, pp. 2035–2046, 2012, doi: 10.1109/TIP.2012.2186306.

[9] Y. Chen, Z. Wang, J. Hu, W. Zhao, and Q. Wu, "The domain knowledge based graph-cut model for liver CT segmentation," Biomed. Signal Process. Control, vol. 7, no. 6, pp. 591–598, 2012, doi: 10.1016/j.bspc.2012.04.005.

[10] S. Arya, D. M. Mount, N. S. Netanyahu, R. Silverman, and A. Y. Wu, "An optimal algorithm for approximate nearest neighbor searching in fixed dimensions," J. ACM, vol. 45, no. 6, pp. 891–923, 1998, doi: 10.1145/293347.293348.

[11] P. Getreuer, "Chan – Vese Segmentation Simplified Mumford – Shah Model Level Set Functions," vol. 2, pp. 1–11, 2012.

[12] G. Li, X. Chen, F. Shi, W. Zhu, J. Tian, and D. Xiang, "Automatic Liver Segmentation Based on Shape Constraints and Deformable Graph Cut in CT Images," IEEE Trans. Image Process., vol. 24, no. 12, pp. 5315–5329, 2015, doi: 10.1109/TIP.2015.2481326.

[13] T. C.-G. K. C. S. G. G. J. A. De Chav Ramnada Cresson, "Kidney Segmentation by Hierarchic Surface Deformation and Surface Anamorphing from CT-Scan or MRI Datasets and Prior Shape," pp. 641–644, 2012.

[14] W. Weng and X. Zhu, "INet: Convolutional Networks for Biomedical Image Segmentation," IEEE

Access, vol. 9, pp. 16591–16603, 2021, doi: 10.1109/ACCESS.2021.3053408.
[15] Goar, V. ., Yadav, N. S. ., & Yadav, P. S. . (2023). Conversational AI for Natural Language Processing: An Review of ChatGPT. International Journal on Recent and Innovation Trends in Computing and Communication, 11(3s), 109–117. https://doi.org/10.17762/ijritcc.v11i3s.6161
[16] Mark White, Kevin Hall, Ana Silva, Ana Rodriguez, Laura López. Predicting Educational Outcomes using Social Network Analysis and Machine Learning . Kuwait Journal of Machine Learning, 2(2). Retrieved from http://kuwaitjournals.com/index.php/kjml/article/view/182