RESEARCH ARTICLE OPEN ACCESS

Object Detection in Medical Video Using Deep Learning

Madhankumar Y
Government Arts College (Autonomous)
Kumbakonam-Thanjavur

ABSTRACT
The development of live video surveillance is the most encouraging solution for surgery in hospital. As of late, a few contributions for video observation have been proposed. Moving objects detection from a video sequence is a fundamental and basic work in numerous computer-vision applications. After this procedure, in object detection, most essential features are separated from video frames and enhance our recognition execution enlivened improvement methods. At last the ideal features to deep learning approach to deal with distinguish the object from Laparoscopic Surgery Medical Video (LSMV) frames, the blood vessel and a few objects are recognized. Different frame based surveillance measurements are utilized to assess the proposed approach. Trial results and comparative examination obviously portray the effectiveness (95.56% precision) of the proposed approach.
Keywords: video surveillance, object detection, background subtraction, deep learning approach

I. INTRODUCTION

In these days, numerous surgeries are video-observed. The real-time video monitoring might be valuable to consequently convey data to the specialist inappropriate time [1-5]. Laparoscopic Surgery Video (LSV), is a kind of simple strategy in which different surgical instruments, besides, robot-helped surgery is less explained than that presented with hands [6-8]. Different explores have been directed on object detection in endoscopic or laparoscopic surgery images. Instrument detection is one focal point of the examination [9-12]. Video observation frameworks rely on the ability to distinguish moving objects in the video stream that is a suitable data extraction in a wide choice of computer vision applications. Manual video observation incorporates an examination of the video content by a human [13-16].

Extracting moving objects from image successions is fundamental issues in PC vision frameworks that are used in an assortment of uses, for example, video reconnaissance, human tracking, development examination and image synthesis [17-22]. Here background subtraction is helpful to detect the object. In object recognition, the initial step is recognized as moving object detection. The main aim of the detection model is that extracting the moving objects that are interested in Laparoscopic video sequence [23-29]. The algorithm, for the most part, utilizes either transient or spatial data in the image sequences to perform moving object detection, and the most normally utilized approach is pixel intensity [30-31].

Background subtraction is to subtract or contrast the present image from a reference background model. Like the third gathering of recognition strategies, the proposed framework does not expect that the surgical errands take after a predefined request and it requires segmented surgical assignments for training [32-34]. In the wake of training, the proposed framework can recognize, constantly, key LSV subsequences that regularly happen amid a given assignment, yet not among different task [35-38]. The task of action recognition is to bridge the gap between the numerical pixel level information and a high state unique action clarification. Then again, multi-view Laparoscopic video successions are repeatedly caught under contrasting brightening and lighting conditions and different cameras may have divergent and new parameters [39-45]. Defeating these issues and creating algorithms for the right detection of surgical instruments is a standout amongst the most critical research subjects for enhancing the exactness and safety of laparoscopic surgical robots and performing robotic surgery all the more astutely [46-50]. Deep learning innovation in light of artificial neural networks has been created in recent times with the appearance of huge data, the improvement of learning algorithms [51].

II. LITERATURE REVIEW

In 2016 Issam Elafia, *et al.* [52] have proposed the most object tracking techniques connected in the video reconnaissance field depend on the earlier pattern recognition of the moving items. These strategies are not sufficient for following a wide range of articles in the meantime on the grounds that the pattern of each moving item ought to be predefined. Accordingly, the creator acquaints another strategy with conquering this issue. In fact, another continuous approach is built up in light of the particle filter and background subtraction. This approach could recognize and track naturally, numerous moving items with no learning stage or earlier information about the size, nature or the underlying position.

In 2018 Wenqing Chu and Deng Cai [53] have proposed an object detection framework which joins the neighborhood appearance and the logical data. In particular, the logical data involves the connections among objects and the worldwide scene based relevant component created by a convolution neural system. The entire framework is planned as a completely associated Contingent Random Field (CRF) characterized by object recommendations. At that point the logical limitations among object proposals were displayed as edges normally. Moreover, a quick mean field estimation strategy was utilized to derive in this CRF model effectively.

In 2015 Binwen Fan and Xiaojiong Liu [54] have recommended the moving objects detection from a video arrangement is a central and basic work in numerous PC vision applications. To address the false classification, another background displaying technique was exhibited and another frontal area object detection show is proposed which use the proposed background comparability to intertwine shading and surface component. It can handle moderately complex situations with little movements and the enlightenment changes. The test comes about have demonstrated the viability of the strategy.

III. METHODOLOGY

The detection and tracking of a determined object in the video cuts have turned into an intriguing zone of research in the computer vision. Our proposed work is to build up the model for moving object detection LSMV clips utilizing deep learning procedure. At first, the predefined input video is separated into a number of frames and these frames are considered as input to the background subtraction process. Its generally utilized approach for distinguishing moving objects in videos from static cameras. Extracting frontal area objects from kept up background model. A frontal area object is any element that distinguished by creating a contrast of each frame of arrangement to background model with morphological task. Then some features like texture, wavelet and SIFT are considered for object recognition using classification approach. Video to Frame Conversion

3.1 Video to Frame Conversion

Detection of moving objects in the frame sequences is a challenging as well as the initial and foremost relevant information extraction video. Detection of moving objects in the frame successions is challenging and in addition the initial and leading data extraction video [55]. The estimations of most extreme and least pixel intensity along with mean of contrasts frame are utilized with temporal differencing approach to deal with acquired frames from video and changed over into images.

3.2 Background subtraction

The process of extracting moving closer view objects (input image) from stored background image (static image) or produced background outline from image series (LSMV) is called background subtraction. The method of reasoning in the approach is that of recognizing the moving objects from the distinction between the present frame and a reference frame, frequently called the "background image", or "background model". The intensity of the pixels relating to the static background remains to a great extent unaltered between two continuous frames.

$$|F_{c,i} - F_{p,i-1}| \quad \text{........................} (1)$$

Frame contrast is the least difficult type of background subtraction. The current frame f_c is just subtracted from the past frame f_p, and if the distinction in pixel esteems for a given pixel is more prominent than a threshold Th then the pixel is thought about piece of the foreground. In video scenes, undesirable regions exist in the closer view may cause a disgraceful tracking issue.

3.3 Morphological Operation

Consequently, to evacuate those undesirable segments, morphological task is performed. The opening and shutting operator is depicted to distinguish the foreground pixels, therefore binary mask images are acquired. The activities test evaluates whether the component fits inside the area. The coveted shape data is acquired after the morphological activity is performed. The midpoint pixel of the organizing component, called the root, perceives the pixel of enthusiasm being prepared [56].

3.4 Object Recognition

The chosen optimal features are given as input of object recognition stage, here we utilized DNN. This neural network design considering the multiple layers of hidden and output units are utilized, also its comprises of both pre training utilizing Deep Belief Network (DBN) and object recognizing stages in its parameter learning. To train the features of LSMV frames, to locate the

Pre Training Process

In the training stage, we utilize a DBN that is a deep design and a feed forward neural network, i.e. with various hidden layers and this model appeared in figure 2. The DBN model awards the system to deliver evident starts based on its hidden units' states, which depicts the system conviction. This layer includes an information layer that encases the information units, number of hidden layers; lastly, a output layer that has one unit for each class is pondered. The parameters of a DBN are the weights among the units of layers in addition to the bias of layer and to set up the parameters is one among the principle challenges for training DNN help of a confined Boltzmann machine (RBM) [57-59].

Restricted Boltzmann Machine (RBM)

A RBM is a two-layer repeated neural system in which stochastic binary inputs are connected to stochastic binary outputs by symmetrically weighted associations. Shown a training case, we disregard its class mark and we broaden it stochastically through the RBM in equation (11).The outputs of the hidden units take after the contingent dissemination is given [18]. This vector is coursed the other path through the RBM which impacts in a confabulation (redoing) of the exceptional data information.

$$R(q_i = 1/p) = \sigma(u_i + \sum W_{ji} p_j$$

…….. (2)

$$R(p_i = 1/q) = \sigma(v_i + \sum W_{ji} q_j$$

Where $u\ and\ v$ bias vectors for visible and hidden layers are, σ is logistic function the range as (0, 1) and $p\,and\ q$ inputs of first and hidden layers.

By adjusting the weights as well as biases to lower the energy of that data and to raise the energy of other data, the probability that the system appoints to training data can be raised primarily for those that have low energies and, henceforth, influence a major commitment to the partition function. For executing the stochastic steepest ascent in the log probability of the training data, this shows the course to an extraordinarily uncomplicated learning condition as

$$\Delta_{wij} \alpha \langle u_i y_i \rangle_{data} - \langle u_i y_i \rangle_{reconstruction}$$

……………….. (3)

If the RBM is trained, a unique RBM can be "stacked" over it to shape a multilayer network. Each time a unique RBM is stacked, the input visible layer is a prepared vector, and qualities for the units in the trained RBM layers are allotted by methods for the present weights and biases and the weight in each RBM to decline the multi-target work. The proficient deep network weights are occupied with preparing a detection stage.

IV. RESULT AND ANALYSIS

In this section we explain the experimental results of object recognition of LSMV is implemented in working platform MATLAB 2016a with i5 processors with 4GB RAM. The experimentation performance is analyzed for proposed and some existing recognition approaches and this process consider the two data base real time videos.

Table 1: Object Recognition Performance

Frames	Precision	PCC	PWC	FPR	TPR	Specificity
Frame 1	0.81	0.95	0.45	0.22	0.75	0.85
Frame 2	0.84	0.89	0.48	0.28	0.69	0.94
Frame 3	0.9	0.86	0.49	0.45	0.84	0.86
Frame 4	0.8	0.85	0.55	0.58	0.85	0.91
Frame 5	0.91	0.94	0.48	0.66	0.9	0.85
Frame 6	0.84	0.96	0.49	0.78	0.88	0.89

Table 1 describes the object recognition performance of dataset 1 i.e. laparoscopic surgery. In the dataset 1, six frames are analyzed and recognize the correct image from the Laparoscopic Surgery Video (LSV). By evaluating the performance measures such as precision, P Correctly Classified (PCC), P Wrongly Classified (PWC), False Positive Rate (FPR), True Positive Rate (TPR) and specificity, the recognition performance is examined. For frame 6, the recognition performance measures are 0.84 in terms of precision, 0.96 in PCC, 0.49 in PWC, 0.78 in FPR, 0.88 in TPR and 0.89 in specificity.

V. CONCLUSION

We have proposed a system to recognize the moving objects through an optimal features and deep learning approach. These maximum likelihood feature points are characterized into foreground pixels and remaining matching feature points are grouped into background based on the frame to frame contrast process. In any case, one can expand the work in future as to the issue of dealing with the drastic illumination changes and numerous moving objects in the scene, yet this strategy can extricate the moving object even under conditions with moderate illumination discrepancy. Some optimal features was (utilizing FP) extracted from video casings to detect the moving object. In future we intend to decrease the future measurements utilizing reduction approach to deal with enhance power, and constancy of object recognition process.

REFERENCE

[1] Chen, Y. T., Chen, C. H., Wu, S., & Lo, C. C. (2019). A two-step approach for classifying music genre on the strength of AHP weighted musical features. Mathematics, 7(1), 19.

[2] Elhoseny, M., Shankar, K., &Uthayakumar, J. (2019). Intelligent diagnostic prediction and classification system for chronic kidney disease. Scientific reports, 9(1), 1-14.

[3] Sivakumar, P., Velmurugan, S. P., & Sampson, J. (2020). Implementation of differential evolution algorithm to perform image fusion for identifying brain tumor.

[4] Khamparia, A., Gupta, D., Nguyen, N. G., Khanna, A., Pandey, B., & Tiwari, P. (2019). Sound classification using convolutional neural network and tensor deep stacking network. IEEE Access, 7, 7717-7727.

[5] Jansirani, A., Rajesh, R., Balasubramanian, R., & Eswaran, P. (2011). Hi-tech authentication for pslette images using digital signature and data hiding. Int. Arab J. Inf. Technol., 8(2), 117-123.

[6] Jain, R., Gupta, D., & Khanna, A. (2019). Usability feature optimization using MWOA. In International conference on innovative computing and communications (pp. 453-462). Springer, Singapore.

[7] Shankar, K., & Lakshmanaprabu, S. K. (2018). Optimal key based homomorphic encryption for color image security aid of ant lion optimization algorithm. International Journal of Engineering & Technology, 7(9), 22-27.

[8] Lyu, L., & Chen, C. H. (2020, July). Differentially Private Knowledge Distillation for Mobile Analytics. In Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval (pp. 1809-1812).

[9] Poonkuntran, S., Rajesh, R. S., & Eswaran, P. (2011). Analysis of difference expanding method for medical image watermarking. In International Symposium on Computing, Communication, and Control (ISCCC 2009) (Vol. 1, pp. 31-34).

[10] Sampson, J., & Velmurugan, S. P. (2020, March). Analysis of GAA SNTFT with Different Dielectric Materials. In 2020 5th International Conference on Devices, Circuits and Systems (ICDCS) (pp. 283-285). IEEE.

[11] Elhoseny, M., Bian, G. B., Lakshmanaprabu, S. K., Shankar, K., Singh, A. K., & Wu, W. (2019). Effective features to classify ovarian cancer data in internet of medical things. Computer Networks, 159, 147-156.

[12] Gochhayat, S. P., Kaliyar, P., Conti, M., Tiwari, P., Prasath, V. B. S., Gupta, D., & Khanna, A. (2019). LISA: Lightweight context-aware IoT service architecture. Journal of cleaner production, 212, 1345-1356.

[13] Dutta, A. K., Elhoseny, M., Dahiya, V., & Shankar, K. (2020). An efficient hierarchical clustering protocol for multihop Internet of vehicles communication. Transactions on Emerging Telecommunications Technologies, 31(5), e3690.

[14] Anand Nayyar, Vikram Puri, Nhu Gia Nguyen, BioSenHealth 1.0: A Novel Internet of Medical Things (IoMT) Based Patient Health Monitoring System, Lecture Notes in Networks and Systems. Springer, 2019

[15] Shankar, K., Lakshmanaprabu, S. K., Khanna, A., Tanwar, S., Rodrigues, J. J., & Roy, N. R. (2019). Alzheimer detection using Group Grey Wolf Optimization based features with convolutional classifier. Computers & Electrical Engineering, 77, 230-243.

[16] Paramathma, M. K., Pravin, A. C., Rajarajan, R., & Velmurugan, S. P. (2019, April). Development and Implementation of Efficient Water and Energy Management System for Indian Villages. In 2019 IEEE International Conference on Intelligent Techniques in Control, Optimization and Signal Processing (INCOS) (pp. 1-4). IEEE.

[17] Chen, C. H., Song, F., Hwang, F. J., & Wu, L. (2020). A probability density function generator based on neural networks. Physica A: Statistical Mechanics and its Applications, 541, 123344.

[18] Kathiresan, S., Sait, A. R. W., Gupta, D., Lakshmanaprabu, S. K., Khanna, A., & Pandey, H. M. (2020). Automated detection and classification of fundus diabetic retinopathy images using synergic deep learning model. Pattern Recognition Letters.

[19] Gupta, D., & Ahlawat, A. K. (2016). Usability determination using multistage fuzzy system. Procedia Comput Sci, 78, 263-270.

[20] Amira S. Ashour, Samsad Beagum, Nilanjan Dey, Ahmed S. Ashour, Dimitra Sifaki Pistolla, Gia Nhu Nguyen, Dac-Nhuong Le, Fuqian Shi (2018), Light Microscopy Image De-noising using Optimized LPA-ICI Filter, Neural Computing and Applications, Vol.29(12), pp 1517–1533, Springer, ISSN: 0941-0643. (SCIE IF 4.664, Q1)

[21] Pan, M., Liu, Y., Cao, J., Li, Y., Li, C., & Chen, C. H. (2020). Visual Recognition Based on Deep Learning for Navigation Mark Classification. IEEE Access, 8, 32767-32775.

[22] Chen, C. H., Hwang, F. J., & Kung, H. Y. (2019). Travel time prediction system based on data clustering for waste collection vehicles. IEICE TRANSACTIONS on Information and Systems, 102(7), 1374-1383.

[23] Shankar, K., & Eswaran, P. (2015). ECC based image encryption scheme with aid of optimization technique using differential evolution algorithm. Int J Appl Eng Res, 10(55), 1841-5.

[24] Anand Nayyar, Vikram Puri, Nhu Gia Nguyen, Dac Nhuong Le, Smart Surveillance Robot for the Real Time Monitoring and Control System in Environment and Industrial Applications, Advances in Intelligent System and Computing, pp 229-243, Springer

[25] Le Nguyen Bao, Dac-Nhuong Le, Gia Nhu Nguyen, Vikrant Bhateja, Suresh Chandra Satapathy (2017), Optimizing Feature Selection in Video-based Recognition using Max-Min Ant System for the Online Video Contextual Advertisement User-Oriented System, Journal of Computational Science, Elsevier ISSN: 1877-7503. Vol.21, pp.361-370. (SCIE IF 2.502, Q1)

[26] Chakchai So-In, Tri Gia Nguyen, Gia Nhu Nguyen: Barrier Coverage Deployment Algorithms for Mobile Sensor Networks. Journal of Internet Technology 12/2017; 18(7):1689-1699.

[27] Le, D.-N.a, Kumar, R.b, Nguyen, G.N., Chatterjee, J.M.d, Cloud Computing and Virtualization, DOI: 10.1002/9781119488149, Wiley.

[28] Bhateja, V., Gautam, A., Tiwari, A., Nhu, N.G., Le, D.-N, Haralick features-based classification of mammograms using SVM, Advances in Intelligent Systems and Computing, Volume 672, 2018, Pages 787-795.

[29] Khamparia, A., Saini, G., Gupta, D., Khanna, A., Tiwari, S., & de Albuquerque, V. H. C. (2020). Seasonal crops disease prediction and classification using deep convolutional encoder network. Circuits, Systems, and Signal Processing, 39(2), 818-836.

[30] Uthayakumar, J., Elhoseny, M., & Shankar, K. (2020). Highly Reliable and Low-Complexity Image Compression Scheme Using Neighborhood Correlation Sequence Algorithm in WSN. IEEE Transactions on Reliability.

[31] Huyen, D.T.T., Binh, N.T., Tuan, T.M., Nguyen, G.N, Dey, N., Son, L.H, Analyzing trends in hospital-cost payments of patients using ARIMA and GIS: Case study at the Hanoi Medical University Hospital, Vietnam, Journal of Medical Imaging and Health Informatics, 7(2), pp. 421-429.

[32] Van, V.N., Chi, L.M., Long, N.Q., Nguyen, G.N., Le, D.-N, A performance analysis of openstack open-source solution for IaaS cloud computing, Advances in Intelligent Systems and Computing, 380, pp. 141-150.

[33] Shankar, K., & Eswaran, P. (2016, January). A new k out of n secret image sharing scheme in visual cryptography. In 2016 10th International Conference on Intelligent Systems and Control (ISCO) (pp. 1-6). IEEE.

[34] Dey, N., Ashour, A.S., Chakraborty, S., Le, D.-N., Nguyen, G.N, Healthy and unhealthy rat hippocampus cells classification: A neural based automated system for Alzheimer disease classification, Journal of Advanced Microscopy Research, 11(1), pp. 1-10

[35] Velmurugan, S. P., & Rajasekaran, P. S. M. P. (2017). CLASSIFICATION OF BRAIN TUMOR USING MULTIMODAL FUSED IMAGES AND PNN. International Journal of Pure and Applied Mathematics, 115(6), 447-457.

[36] Shankar, K., Elhoseny, M., Perumal, E., Ilayaraja, M., & Kumar, K. S. (2019). An Efficient Image Encryption Scheme Based on Signcryption Technique with Adaptive Elephant Herding Optimization. In Cybersecurity and Secure Information Systems (pp. 31-42). Springer, Cham.

[37] Wu, L., Chen, C. H., & Zhang, Q. (2019). A mobile positioning method based on deep learning techniques. Electronics, 8(1), 59.

[38] Lydia, E. L., Kumar, P. K., Shankar, K., Lakshmanaprabu, S. K., Vidhyavathi, R. M., & Maseleno, A. (2020). Charismatic document clustering through novel K-Means non-negative matrix factorization (KNMF) algorithm using key phrase extraction. International Journal of Parallel Programming, 48(3), 496-514.

[39] Sujitha, B., Parvathy, V. S., Lydia, E. L., Rani, P., Polkowski, Z., & Shankar, K. (2020). Optimal deep learning based image compression technique for data transmission on industrial Internet of things applications. Transactions on Emerging Telecommunications Technologies, e3976.

[40] Lo, C. L., Chen, C. H., Hu, J. L., Lo, K. R., & Cho, H. J. (2019). A fuel-efficient route plan method

based on game theory. Journal of Internet Technology, 20(3), 925-932.

[41] Kung, H. Y., Chen, C. H., Lin, M. H., & Wu, T. Y. (2019). Design of Seamless Handoff Control Based on Vehicular Streaming Communications. Journal of Internet Technology, 20(7), 2083-2097.

[42] Elhoseny, M., & Shankar, K. (2019). Reliable data transmission model for mobile ad hoc network using signcryption technique. IEEE Transactions on Reliability.

[43] Shanmugam, P., Rajesh, R. S., & Perumal, E. (2008, May). A reversible watermarking with low warping: an application to digital fundus image. In 2008 International Conference on Computer and Communication Engineering (pp. 472-477). IEEE.

[44] Shankar, K., & Elhoseny, M. (2019). Trust Based Cluster Head Election of Secure Message Transmission in MANET Using Multi Secure Protocol with TDES. J. UCS, 25(10), 1221-1239.

[45] Parvathy, V. S., Pothiraj, S., & Sampson, J. (2020). Optimal Deep Neural Network model based multimodality fused medical image classification. Physical Communication, 101119.

[46] Subbiah Parvathy, V., Pothiraj, S., & Sampson, J. (2020). A novel approach in multimodality medical image fusion using optimal shearlet and deep learning. International Journal of Imaging Systems and Technology.

[47] Parvathy, V. S., & Pothiraj, S. (2019). Multi-modality medical image fusion using hybridization of binary crow search optimization. Health Care Management Science, 1-9.

[48] Velmurugan, S. P., Sivakumar, P., & Rajasekaran, M. P. (2018). Multimodality image fusion using centre-based genetic algorithm and fuzzy logic. International Journal of Biomedical Engineering and Technology, 28(4), 322-348.

[49] Chen, C. H. (2018). An arrival time prediction method for bus system. IEEE Internet of Things Journal, 5(5), 4231-4232.

[50] Shankar, K., Perumal, E., & Vidhyavathi, R. M. (2020). Deep neural network with moth search optimization algorithm based detection and classification of diabetic retinopathy images. SN Applied Sciences, 2(4), 1-10.

[51] Elafi, I., Jedra, M. and Zahid, N., 2016. Unsupervised detection and tracking of moving objects for video surveillance applications. Pattern Recognition Letters,pp. 84, pp.70-77.

[52] Chu, W. and Cai, D., 2018. Deep feature based contextual model for object detection. Neurocomputing, Vol.275, pFan, B. and Liu, X., 2015, August. Foreground detection based on color and texture features. In Fuzzy Systems and Knowledge Discovery (FSKD), 2015 12th International Conference on pp. 1940-1944. IEEE.

[53] Babaee, M., Dinh, D.T. and Rigoll, G., 2017. A deep convolutional neural network for background subtraction. arXiv preprint arXiv:1702.01731,pp.1-28.

[54] Mohanty, S. N., Ramya, K. C., Rani, S. S., Gupta, D., Shankar, K., Lakshmanaprabu, S. K., & Khanna, A. (2020). An efficient Lightweight integrated Blockchain (ELIB) model for IoT security and privacy. Future Generation Computer Systems, 102, 1027-1037.

[55] Elhoseny, M., & Shankar, K. (2020). Energy efficient optimal routing for communication in VANETs via clustering model. In Emerging Technologies for Connected Internet of Vehicles and Intelligent Transportation System Networks (pp. 1-14). Springer, Cham.

[56] Chen, C. H. (2020). A cell probe-based method for vehicle speed estimation. IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, 103(1), 265-267.

[57] Khamparia, A., Singh, A., Anand, D., Gupta, D., Khanna, A., Kumar, N. A., & Tan, J. (2018). A novel deep learning-based multi-model ensemble method for the prediction of neuromuscular disorders. Neural computing and applications, 1-13.

[58] Shankar, K., Zhang, Y., Liu, Y., Wu, L., & Chen, C. H. (2020). Hyperparameter tuning deep learning for diabetic retinopathy fundus image classification. IEEE Access, 8, 118164-118173.