Analysis of Tumours in Urinary Bladder using Machine Learning

G. Sudha[1], M.M.Hussani[2],M Lavanya[1], M Sabeena[1], A Abinaya[1]
1- Department Of Biomedical Engineering, Muthayammal Engineering College
2- Department Of EEE, Government College Of Engineering, Erode.

ABSTRACT - Using biopsy, tumors classification is performed, which is not normally conducted before definitive surgery. The machine learning technology helps radiologists to diagnose tumor without invasive measures. Convolution neural network (CNN) is the machine-learning algorithm which achieved substantial results in image classification and segmentation. Some of the most notable primary kidney disease types are tumor, stone and cyst (cystoscopy). Urinary bladder cancer is one of the most common malignancies of the urinary system. The root of cancer lays in the mutation of the bladder mucous cells. The main characteristics of urinary bladder cancer are high re-occurrence rate and high metastatic potential. This paper develops a new CNN architecture to classify disease types and checks whether the image is having tumor, stone, cyst data or not. With i) good generalization capability and ii) good execution speed, newly developed CNN architecture is being used as an effective decision-support tool for radiologists in diagnostics. Python is used for development of the project.

Keywords-:Urinary, Bladder-Cancer, Cystoscopy, Convolution Neural Network.

INTRODUCTION

Urinary bladder cancer is one of the most common malignancies of the urinary system. The root of cancer lays in the mutation of the bladder mucosa cells. The main characteristics of urinary bladder cancer are high re-occurrence rate and high metastatic potential.

For these reasons, it can be concluded that an accurate and fast diagnostic procedure is of vital importance. With aim of developing faster and more accurate diagnostic procedures, artificial intelligence (AI) and machine learning (ML) are introduced [3-5]. The AI utilization for urinary bladder cancer diagnostics was introduced here. The presented diagnostic procedures are based on two different diagnostic procedures: a cystoscopy and computerized tomography.

Cystoscopy

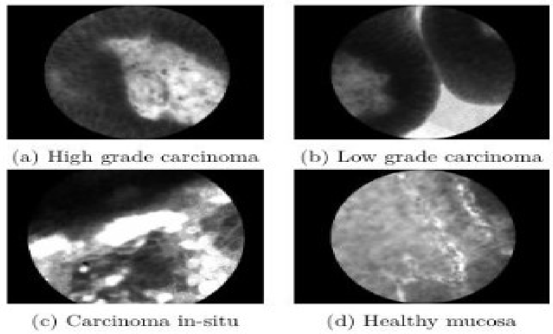
The first used data set is the data set collected during optical cystoscopy and examination of urinary bladder mucosa with a confocal laser endomicroscope. The images of urinary bladder mucosa are divided into four different classes:

• High-grade carcinoma,

• Low-grade carcinoma,

• Carcinoma in-situ, and

• Healthy mucosa



An example of each class is given in Figure 1.

Figure 1: Example of each class from cystoscopy data set

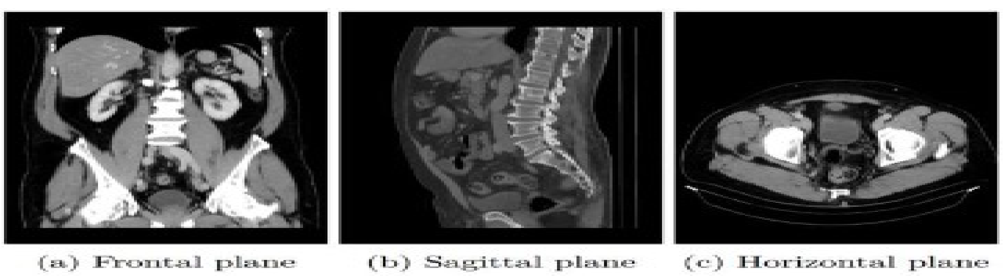
The presented data set was applied to recognize the grade of urinary bladder cancer.

Computerized Tomography

The other data set used in this paper is the data set collected using Computerized Tomography technique.

The images collected with CT were captured in three planes (a. frontal, b. horizontal, and c. sagittal) and were divided into six classes (images with healthy bladder, without a bladder, unilateral bladder wall thickening, circular bladder wall thickening, exophytic formation, and also invasion outside the contour of bladder).

An example for each plane used in this research is shown in Figure



2.

Figure 2: Example of each plane from CT data set

This data set, aim is to develop the system for semantic segmentation for urinary bladder cancer masses in CT images. A brief process of the semantic segmentation is shown in Figure 3.

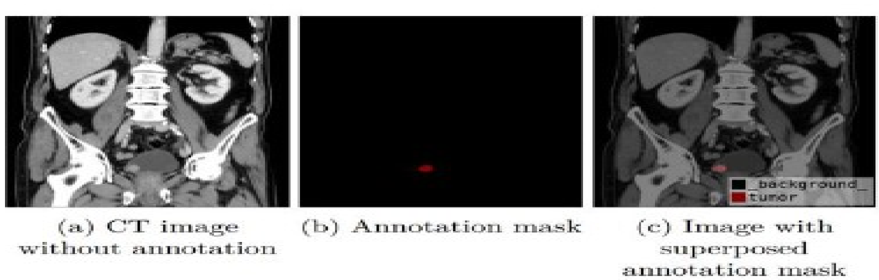


Figure 3: Semantic segmentation process

RELATED WORKS

In this paper[1], a fully connected. Convolution neural network is used to segmented and classification of blood cell microscope WBC images for healthy and unhealthy conditions. The performance of the classifier was analyzed. The accuracy sensitivity specificity and pression are 96.84%, 96.26%,97.35% and 96.39% respectively.

[2] In recent times, so numerous Computers backed opinion (CAD) systems are designed for opinion of several conditions. Lung cancer discovery at early stage has come veritably important and also veritably easy with image processing and deep literacy ways. In this design lung case Computer Tomography (CT) checkup images are used to descry and classify the lung nodes and to descry the malice position of that node. The CT checkup images are segmented using U-Net armature. We employed deep transfer literacy to handle the failure of available data and designed a Convolution Neural Network (CNN) model along with the Machine literacy styles Random Forest (RF), Support Vector Machines (SVM), and Decision Tree (DT). The proposed approach was estimated on intimately available Lung CT checkup dataset. .

[3] In order to identify the affected region by the help of segmentation technique. An effective way for detecting the breast cancer affected regions in mammograms by image segmentation method. It will give better diagnosis for tumor affected regions. The proposed model helps to improve the diagnostic performance of breast cancer disease. In order to detect the tumor affected regions by segmentation technique. Further those segmented regions features are extracted and it is trained completely, finally trained images are classified by the efficient classifier of different classes in mammogram. Our process is to compare the deep learning model(CNN image features).

In this paper [4] the authors discussed that Non muscle-invasive bladder cancer had a relatively high-post operative recurrence rate despite of the implementation of normal/conventional treatment methods. Cystoscopy is must for diagnosing as well as monitoring bladder cancer, but lesions were overlooked while utilizing white-light imaging. By using cystoscopy, tumors with the small diameter; flat tumors, like carcinoma in situ; and extent for flat lesions associated with elevated lesions were difficult to identify.

Materials and Methods: The total of 2103 cystoscopic images, consisting of 1672 images of normal tissues and 432 images of tumor lesions, was used to creating a dataset with an 8:2 ratio of training/ test images. They constructed a tumor classifier based upon the CNN convolutional neural network. The performance of trained classifier were evaluated using the test data. True-positive rate and false-positive rate are plotted when threshold were changed as receiver operating characteristic (ROC) curve. Results: In the test data (tumor image: 88, normal image: 336), 79 images were true positive, 315 true negative, 21 false positive, and 10 false negative.

The area under ROC curve were 0.98, with the maximum Youden index of 0.836, sensitivity for 89.6%, and specificity of 93.9%. Conclusion: By objectively evaluating cystoscopic image with CNN, it was possible to classify the images, including tumor lesions and normality. The objective evaluation of the cystoscopic images using AI was expected to contribute for improvement in accuracy of the diagnosis and treatment of bladder cancers.

In this paper [5] the authors stated that one of the major challenges in Machine Learning in application medicine was data collection. Either ethical concerns or patient's lack, data might be scarce. Here, Deep Convolution Generative Adversarial Networks (DCGAN) had been applied for data augmentation purpose. bladder mucosa images were used in order to generate novel images using DCGANs.

Then, original and generated images combination were used to train AlexNet as well as VGG16 architectures. The results showed that the improvements in AUC score in some cases/ equal scores with apparent lowering of the standard-deviation across data folds in cross-validation; indicating networks trained with addition of generated data is having the lower sensitivity across hyperparameter range.

Urinary bladder cancer is one of the most common urinary tract cancers. Standard diagnosisprocedure can be invasive and time-consuming. For the above reasons, procedure named optical-biopsyis was introduced.

This procedure allowed in-vivo evaluation of the bladder mucosa without biopsy need. Although faster and invasive, accuracy is often lower. For the reason, machine learning (ML) algorithms were being used to increase the accuracy. The issue with the ML algorithms is the sensitivity to input data amount. In medicine, collection is time-consuming becuase of potentially low count of patients. Authors demonstrate the use of Deep Convolutional Generative Adversarial Networks (DCGAN) for the image generation .

Urinary bladder cancer the most common urinary tract malignant diseases , which is a consequence of mutation in bladder’s mucosa cells which causes the uncontrolled growth. Such growth will show a high tendency to spread to rest of mucosa, and also to various other parts of human body. Above the stated facts, urinary bladder cancer shows elevated recurrence rates which are ranging from 61% in first year, 78% in first five years. The recurrence rates are among highest when compared to other malignant diseases.

So, diagnosis, treatment and follow-up are most extremely challenging. There is the multiplicity of urinary bladder cancer kinds, among which most common are:

a) urothelial carcinom

b) squamous cell carcinoma

c) adenocarcinoma

d) small cell carcinoma and sarcoma .

Most urinary bladder cancers are a) urothelial carcinomas, b) transitional cell carcinomas (TCC). These carcinomas are, in the papillary form, denoted by low-grade meta-static potential. High-grade cancer as well as carcinoma in-situ (CIS) are being characterized by high metastatic potential. Unlike TCC, the CIS lesion of the urinary bladder mucosa is most dominantly flat, and making harder to distinguish from the benign growths. The dominant medical examination in diagnosis of the urinary bladder cancer is the cystoscopy, which is an endoscopic method in which a probe known as cystoscope is inserted in urinary bladder via the ureter. Modern cystoscopes are also equipped with confocal-laser-endomicroscopes (CLE) which allow in-vivo evaluation of mucosa, without need for biopsy as well as patho-histological examination. This method which is called the optical biopsy, will show high results from point of view in detecting papillary lesions. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions

In this paper [6] authors stated that in the last decade, deep learning (DL) computing methodologyhas been deemed Gold Standard in machine learning (ML) communities. In addition, it gradually become the most widely adapted computational approach in ML field , and achieving the outstanding results on various complex cognitive tasks, matching/ beating those provided by the human performance. One of the DL benefits is the capacity to learn data with huge amounts . The DL field is growing fast in last few years and it is extensively used successfully which address the wide range of traditional applications.

Despite it has been contributed in several works reviewing State-of-the-Art on DL, all of them tackled only one aspect of the DL, which leads to the overall lack of knowledge about it.

Therefore, in this contribution here, they were proposed the more holistic approach to provide a more suitable starting point from which for developing a full understanding of DL.

Specifically, the review attempts in providing a more comprehensive survey of most important DL aspects and includes those enhancements recently added to field. In particular, this study outlines the DL importance, presenting the types of DL techniques and networks.

PROPOSED METHODOLOGY

In existing system, the image is being visualized manually and tumor/stone presence is checked by doctors. Normally images are processed with median filter so that the noise pixels are eliminated and image clarity is improved. Only through manual checking, tumor content presence could be identified. Some studies classified the tumor types using augmented tumor region of interest, image dilatation, and ring-form partition. SVM and KNN based classification is carried. It classifies the image is having tumor content or not only. Disease type such as tumor/stone/cyst could not be found out. Drawbacks are:

- Median filter process required.
- Not effective when tumor data is very low in size.
- Accurate image processing for tumor/stone/cyst presence is not possible.
- Manual verification/checking are required.
- SVM/KNN accuracy will be reduced if data-set size grows large.
- Multi class disease type such as tumor/stone/cyst is not possible.

In proposed system, the gray scale image is taken. The RGB image if taken, is converted into gray scale image first. Then all the images are re-sized into same size. Then the training data set image with tumor class factor ‘yes’ are taken along with ‘no’. For each image, all the pixels’ grayscale value are found out and written in a row. So the total number of rows is equal to total number of images. The number of columns is the number of pixels in the image. These data is saved in comma separated file. For testing data, all these operations are carried out and saved in another comma separated value file. Then Convolutional neural network is applied to train the model with test data. The accuracy is found out and displayed. From the given test image, it can be found out as tumor present or not by checking with training data images. Disease type such as tumor/stone/cyst could be found out. The following modules are present in the project.

1) Image File Selection

2) CSV File Preparation

3) Image Classification

4) SVM/KNN Classification

5) CNN Based Model Prediction

1) IMAGE FILE SELECTION

In this module, the image file selection is carried out. The data set is taken from ' https://www.kaggle.com/datasets/nazmul0087/ct-kidney-datasets-normal-cyst-tumor-and-stone' which contains training data with ‘tumor’, 'stone', 'cyst' and ‘normal' set of images. Moreover, it contains testing data with same category of images. During CNN processing these data are to be given as input and accuracy percent will be found out.

2) CSV FILE PREPARATION

The dataset folders images are being taken and all the images are sequentially processed. The CSV file with name ‘kidney_training.csv’ is created with header row as Label, Pixel1, Pixel2 upto Pixel10000. Then, for each image, resize operation for width and height to 100 is made, the grayscale values are found out and appended in ‘csv’ file such that each row contains gray scale values of each image. The label column is filled with ‘1’ for kidney tumor ‘yes’ image and ‘0’ for kidney tumor ‘no’ image.

3) IMAGE CLASSIFICATION

In this module, a file name is given as input which is the name of the image file name to classify and it should be present in root folder of the project. The grayscale values of the image are found out after resize to 100x100. Then, for each row in the ‘csv’ file, the pixels are checked for similarity. First it is checked against all ‘1’ label rows and then if not matched, checked with ‘0’ labels rows. If matched in label ‘1’ , the given image is identified as tumor present. If matched in label ‘0’ , then the image is grouped as tumor not present. If not matched in both, then the image cannot be classified as kidney tumor found or not.

4) SVM/KNN CLASSIFICATION

In this module, 75% of records in given data set are taken as the training data and 25% of records are taken as the test data. The pixel values are taken as numerical values. Then the model is trained using data in training group and predicted thereafter with test data. Of which, records are classified as disease present or not.

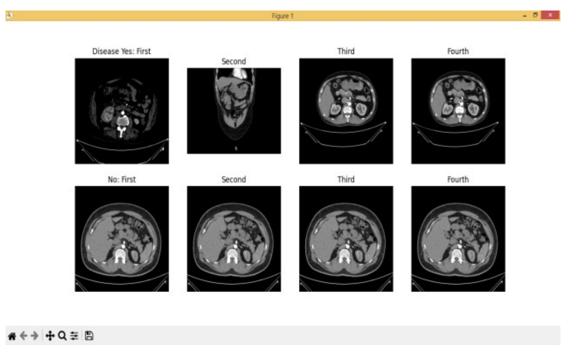
5) CNN BASED MODEL PREDICTION

Here image dataset is taken first. It can be seen that the image data is saved in form of pixel values. But it cannot be feeded to the CNN model in this format. So, it is again converted into numpy array. It is must to convert categorical data into one hot vector encoding. Then, reshaping the data and cast it into *float32* type so it will be used conveniently. Preprocessing data is finished by carrying the process called normalizing. Normalizing the image data will map the entire pixel values in all images to values between ‘0’ to ‘1’. This helps to minimize inconsistencies in the data. Before normalizing, image data will have large variations in pixel values that lead to some unusual behavior during training process.

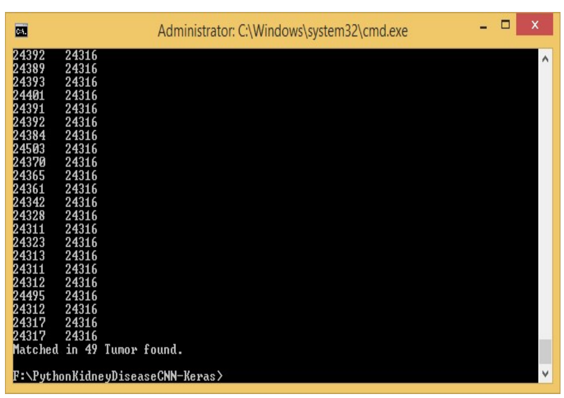
Once the model is created, it will be imported and compiled using the ‘model.compile’. The model is trained to just 5 epochs but it can be increased the number of epochs. After training process completed, we make predictions on test set. The accuracy value is displayed during the iterations. Multi class image labelling is also possible here.

EXPERIMENTAL RESULTS

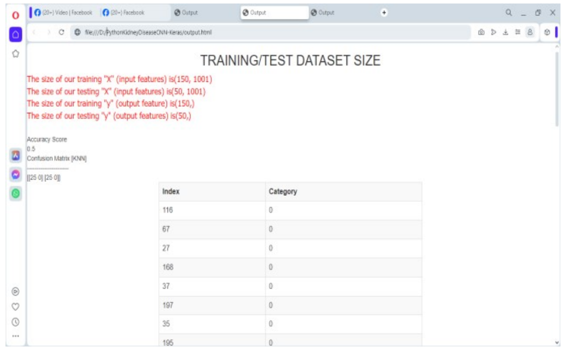
- There is no need of Median filter process .
- Even tumor data is very low in size, it is effective.
- Accurate image processing for checking tumor presence is available.
- Manual checking/verification is not required.
- CNN algorithm increased the accuracy score.
- Multi class disease types like tumor/stone/cyst is possible.



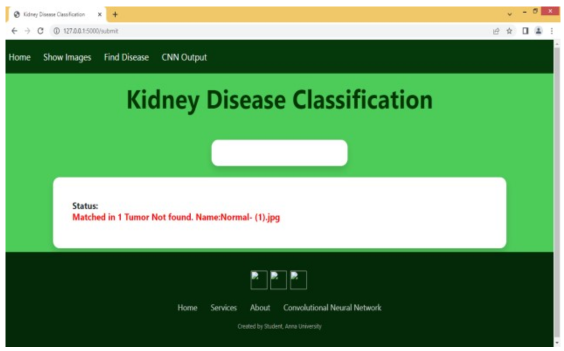
4.1 DISPLAYING IMAGES



4.2 FINDING TUMOR DATA



4.3 KNN OUTPUT



4.4 TUMOR CLASSIFICATION

CONCLUSION AND FUTURE ENHANCEMENT

This project introduced a new CNN architecture for classify people with / without kidney tumor and checks if image is having tumor data or not. With a) good capability of generalization and b) fast execution speed, the developed CNN architecture is used as an efficient decision-support tool for admins in detecting the presence of tumor . Python is used as coding language for project development. If this application is developed as web service, it could be integrated in various network applications. The application is designed such that the said enhancements could be integrated with current application.

REFERENCES

[1] Classification of WBC cell classification using fully connected convolution neural network K Gokul Kannan1, T R Ganesh Babu2, R Praveena2, P Sukumar3, G Sudha4 and M Birunda4 Published under licence by IOP Publishing Ltd Citation K Gokul Kannan et al 2023 J. Phys.: Conf. Ser. 2466 012033 DOI 10.1088/1742-6596/2466/1/01203

[2] G.Sudha,M.Birunda, R. Khowshalya, S. Preneshvar, S. Jayanthiswari, V. Sathya, N. Shneka. Lung Cancer Image Segmentation and Detection Using Deep Learning Algorithms. Cardiometry; Issue No. 26; February 2023; p.556-562;DOI: 10.18137/cardiometry.2023.26.556562;

[3] G. Sudha, C. Selvi, N. Kaliammal, M. Varsha, R. Sujithra, E. Theerthammal, S. Kamali. Comparision Study Analysis Using Deep Learning Algorithm for Breast Cancer. Cardiometry; IssueNo. 26; February 2023; p. 563-568; DOI:10.18137/cardiometry.2023.26.563568;

[4] Atsushi Ikeda et.al.,,Support System of Cystoscopic Diagnosis for Bladder Cancer Based on Artificial Intelligence. Journal of Endourology.Mar 2020.352- 358.http://doi.org/10.1089/end.2019.0509.

[5] Lorencin I, Baressi Šegota S, Anđelić N, Mrzljak V, Ćabov T, Španjol J, Car Z. On Urinary Bladder Cancer Diagnosis: Utilization of Deep Convolutional Generative Adversarial Networks for Data Augmentation. Biology (Basel). 2021 Feb 26;10(3):175. doi: 10.3390/biology10030175. PMID: 33652727; PMCID: PMC7996800.

[6] Alzubaidi L, Zhang J, Humaidi AJ, Al-Dujaili A, Duan Y, Al-Shamma O, Santamaría J, Fadhel MA, Al-Amidie M, Farhan L. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. J Big Data. 2021;8(1):53. doi: 10.1186/s40537-021-00444-8. Epub 2021 Mar 31. PMID: 33816053; PMCID: PMC8010506.

[7] Lorencin, Ivan, et al. "Using multi-layer perceptron with Laplacian edge detector for bladder cancer diagnosis." Artificial Intelligence in Medicine 102 (2020): 101746.

[8] Zlatev, D.V.; Altobelli, E.; Liao, J.C. Advances in imaging technologies in the evaluation of high-grade bladder cancer. Urol. Clin. 2015, 42, 147–15.

[9] Cancer Genome Atlas Research Network. Comprehensive molecular characterization of urothelial bladder carcinoma. Nature 2014, 507, 315–322.

[10] Takahashi,K.; Kimura, G.; Endo, Y.; Akatsuka,J.; Hayashi, T.; Toyama,Y.; Hamasaki, T.; Kondo, Y. Urothelial carcinoma of the bladder, lipid cell variant: A case report and literature review. J. Nippon. Med Sch. 2019._86-602.

[11] Dotson, A.; May, A.; Davaro, F.; Raza, S.J.; Siddiqui, S.; Hamilton, Z. Squamous cell carcinoma of the bladder: Poor response to neoadjuvant chemotherapy. Int. J. Clin. Oncol. 2019, 24, 706–711.

[12] Celis, J.E.; Wolf, H.; Østergaard, M. Bladder squamous cell carcinoma biomarkers derived from proteomics. Electrophor. Int. J. 2000, 21, 2115–2121.

[13] Dadhania, V.; Czerniak, B.; Guo, C.C. Adenocarcinoma of the urinary bladder. Am. J. Clin. Exp. Urol. 2015, 3, 51.

[14] Tsai IJ, Shen WC, Lee CL, Wang HD, Lin CY. Machine Learning in Prediction of Bladder Cancer on Clinical Laboratory Data. Diagnostics (Basel). 2022 Jan 14;12(1):203.

[15] JD Cohen, et al., Detection and localization of surgically resectable cancers with a multi-analyte blood test. Science 359, 926–930 (2018).

[16] C.Nagarajan and M.Madheswaran - 'Experimental verification and stability state space analysis of CLL-T Series Parallel Resonant Converter' - Journal of ELECTRICAL ENGINEERING, Vol.63 (6), pp.365-372, Dec.2012.

[17] C.Nagarajan and M.Madheswaran - 'Performance Analysis of LCL-T Resonant Converter with Fuzzy/PID Using State Space Analysis'- Springer, Electrical Engineering, Vol.93 (3), pp.167-178, September 2011.

[18] C.Nagarajan and M.Madheswaran - 'Stability Analysis of Series Parallel Resonant Converter with Fuzzy Logic Controller Using State Space Techniques'- Taylor & Francis, Electric Power Components and Systems, Vol.39 (8), pp.780-793, May 2011.

[19] C.Nagarajan and M.Madheswaran - 'Experimental Study and steady state stability analysis of CLL-T Series Parallel Resonant Converter with Fuzzy controller using State Space Analysis'- Iranian Journal of Electrical & Electronic Engineering, Vol.8 (3), pp.259-267, September 2012.

[20] Nagarajan C., Neelakrishnan G., Akila P., Fathima U., Sneha S. "Performance Analysis and Implementation of 89C51 Controller Based Solar Tracking System with Boost Converter" Journal of VLSI Design Tools & Technology. 2022; 12(2): 34–41p.

[21] C. Nagarajan, G.Neelakrishnan, R. Janani, S.Maithili, G. Ramya "Investigation on Fault Analysis for Power Transformers Using Adaptive Differential Relay" Asian Journal of Electrical Science, Vol.11 No.1, pp: 1-8, 2022.

[22] G.Neelakrishnan, K.Anandhakumar, A.Prathap, S.Prakash "Performance Estimation of cascaded h-bridge MLI for HEV using SVPWM" Suraj Punj Journal for Multidisciplinary Research, 2021, Volume 11, Issue 4, pp:750-756

[23] G.Neelakrishnan, S.N.Pruthika, P.T.Shalini, S.Soniya, "Perfromance Investigation of T-Source Inverter fed with Solar Cell" Suraj Punj Journal for Multidisciplinary Research, 2021, Volume 11, Issue 4, pp:744-749

[24] C.Nagarajan and M.Madheswaran, "Analysis and Simulation of LCL Series Resonant Full Bridge Converter Using PWM Technique with Load Independent Operation" has been presented in ICTES'08, a IEEE / IET International Conference organized by M.G.R.University, Chennai.Vol.no.1, pp.190-195, Dec.2007

[25] M Suganthi, N Ramesh, "Treatment of water using natural zeolite as membrane filter", Journal of Environmental Protection and Ecology, Volume 23, Issue 2, pp: 520-530,2022

[26] M Suganthi, N Ramesh, CT Sivakumar, K Vidhya, "Physiochemical Analysis of Ground Water used for Domestic needs in the Area of Perundurai in Erode District", International Research Journal of Multidisciplinary Technovation, pp: 630-635, 20