4th
ICAII4.0
INTERNATIONAL CONFERENCE on ARTIFICIAL INTELLIGENCE towards INDUSTRY 4.0
11-12 November 2021
İskenderun / Hatay

icaii4.iste.edu.tr

ICAII 4.0

November 11-12. 2021
İskenderun, HATAY / TURKEY

PROCEEDINGS BOOK

Editors:

Assoc. Prof. Dr. Yakup Kutlu
Assoc. Prof. Dr. Sertan Alkan
Assist. Prof. Dr. İpek Abasıkeleş Turgut
Assist. Prof. Dr. Ahmet Gökçen
Assist. Prof. Dr. Gökhan Altan
Assist. Prof. Dr. Mehmet Sarıgül
Assist. Prof. Dr. Levent Karacan

ISBN: 978-605-73802-0-3

NATURE and SCIENCE
SOCIETY
www.dogavebilim.com

icaii4.iste.edu.tr

Conference Proceedings Book

Editors:

Assoc. Prof. Dr. Yakup Kutlu

Assoc. Prof. Dr. Sertan Alkan

Assist. Prof. Dr. İpek Abasıkeleş Turgut

Assist. Prof. Dr. Ahmet Gökçen

Assist. Prof. Dr. Gökhan Altan

Assist. Prof. Dr. Mehmet Sarıgül

Assist. Prof. Dr. Levent Karacan

Technical Editors:

Merve Nilay Aydın

Handan Gürsoy Demir

Kadir Tohma

Halil Okur

ISBN: 978-605-73802-0-3 **Publication No:1**

International Conference on Artificial Intelligence towards Industry 4.0 (ICAII4.0) has been organized in Iskenderun, Hatay/Turkey on 11-12 November 2021.

The main objective of ICAII4.0 is to present the latest research and results of scientists related to Artificial Intelligence, Industry 4.0 and all sub-disciplines of Computer Engineering. This conference provides opportunities for the different areas delegates to exchange new ideas and application experiences face to face, to establish business or research relations and to find global partners for future collaboration.

We hope that the conference results provide significant contribution to the knowledge in this scientific field. The organizing committee of conference is pleased to invite prospective authors to submit their original manuscripts to ICAII4.0.

All paper submissions will be double-blind and peer-reviewed and evaluated based on originality, technical and/or research content/depth, correctness, relevance to conference, contributions, and readability. Selected papers presented in the conference that match with the topics of the journals will be published in the following journals:

- Natural and Engineering Sciences
- Journal of Intelligent Systems with Applications

In particular we would like to thank Prof. Dr. Tolga Depci, Rector of Iskenderun Technical University; Natural and Engineering Sciences, Academic Publisher; Journal of Intelligent Systems with Applications. They have made a crucial contribution towards the success of this conference. Our thanks also go to the colleagues in our conference office.

Looking forward to see you in next Conference.

Conference Organizing Committee

SUPPORTERS

Iskenderun Technical University
www.iste.edu.tr

Natural and Engineering Sciences
www.nesciences.com

Journal of Intelligent Systems with Applications

asd.islerya.com

Nature and Science Society
www.dogavebilim.com

COMMITTEES

HONOUR COMMITTEE

Prof. Dr. Tolga DEPCİ *Rector*	*Iskenderun Technical University*

ORGANISATION COMMITTEE

Assoc. Prof. Dr. Yakup KUTLU	*Iskenderun Technical University*
Assoc. Prof. Dr. Sertan ALKAN	*Iskenderun Technical University*
Dr. Ahmet GÖKÇEN	*Iskenderun Technical University*
Dr. İpek ABASIKELEŞ TURGUT	*Iskenderun Technical University*
Dr. Gökhan ALTAN	*Iskenderun Technical University*
Dr. Mehmet SARIGÜL	*Iskenderun Technical University*
Dr. Levent KARACAN	*Iskenderun Technical University*

SCIENTIFIC COMMITTEE

Dr. Ahmet GÖKÇEN	*Iskenderun Technical University*
Dr. Gökhan ALTAN	*Iskenderun Technical University*
Dr. İpek ABASIKELEŞ TURGUT	*Iskenderun Technical University*
Dr. Levent KARACAN	*Iskenderun Technical University*
Dr. Mehmet SARIGÜL	*Iskenderun Technical University*
Dr. Mustafa YENİAD	*Ankara Yildirim Beyazit University*
Dr. Sertan ALKAN	*Iskenderun Technical University*
Dr. Yakup KUTLU	*Iskenderun Technical University*
Dr. Yalçın İŞLER	*Izmir Katip Celebi University*

INVITED SPEAKERS

Cem PEHLİVAN
ASELSAN

Hüsnü Levent PANDÜL
Altinok Consulting Engineering Inc.

CONTENT

Detecting Tagged People in Camera Images

Muhammed Telçeken[1*], Yakup Kutlu[1]

[1]Bilgisayar Mühendisliği Anabilim Dalı, İskenderun Teknik Üniversitesi, İskenderun
Email: mtelceken@gmail.com, yakup.kutlu@iste.edu.tr

Abstract

With the development of technology, cameras are used more widely. It is possible to evaluate the widespread use of cameras in various subjects in daily life. Especially face recognition systems are one of the most important areas of use of cameras. Facial recognition systems can be used in many areas such as cyber security, entertainment, security applications of daily used devices, and faster and easier transactions in financial areas. Although a lot of progress has been made in this regard, face recognition systems are still used widely enough because it is thought that they have weaknesses in terms of security. Many scientists are working on facial recognition. In this study, it is aimed to detect the faces of people determined from videos or live camera images in the best and safest way. Yolov4 object detection algorithm, a ready-made algorithm, was used for the detection of human faces on images. The faces of the people in the images were detected by training the data set we created in the Yolov4 algorithm. An accuracy of 99.0% has been achieved for detecting people's faces on images. The data set we created with pictures of certain people is trained in the CNN algorithm. The faces of the people detected on the images were classified on the model trained with the CNN algorithm for the identification of the people, and the accuracy value was examined for the detection of the identified people on the video recordings or live images from the cameras.

Keyword(s): *Face Recognition, Yolov4, CNN Algorith*

Diagnosis of Ovarian Cancer using Conventional Machine Learning Methods

Bartu Yesilkaya[1], Saadet Sena Egeli[1], Yalcin Isler[1,2]

[1] Department of Biomedical Engineering, Izmir Katip Celebi University, Turkey
[2]Islerya Medical and Information Technologies, Turkey
Emails: bartu.yesilkaya@yandex.com, egelisena@gmail.com, islerya@yahoo.com

Abstract

Ovarian cancer is one of the common types of cancer among women that can result in death unless it is not diagnosed at earlier stages [1]. This study utilized nine widely-used machine learning methods to diagnose ovarian cancer. An open-source database [2], which can be accessed free of charge on the Internet via Mendeley, was investigated via these algorithms. We evaluated the classifiers' performances with the 10-fold cross-validation method [3,4]. As a result, the classifier performances obtained 100% with Logistic Regression, 98% with Support Vector Machines, Stochastic Gradient Descent, Random Forest, and Multi-Layer Perceptron, 97% with AdaBoost and Decision Tree, 93% with k-Nearest Neighbors, and 85% with Naive Bayes. By comparing to the literature, the achieved results seem satisfactory. The outcomes of this study need to be verified by other studies using different databases before using in the clinic.

Keyword(s): Ovarian cancer, Pattern recognition, Machine learning, Cross-validation

Reference(s):

1. A. Sahin, N. Ozcan, G. Nur, Ovarian cancer prediction using PCA,K-PCA, ICA and random forest, Journal of Intelligent Systems with Applications 4 (2021) 103-108.
2. P. Yasodha, N. Ananthanarayanan, Detecting the ovarian cancer using big data analysis with effective model, Biomedical Research (0970-938X) (2018).
3. M.B. Selek, B. Yesilkaya, S.S. Egeli, Y. Isler. The effect of principal component analysis in the diagnosis of congestive heart failure via heart rate variability analysis, Proceedings of the IMechE Part H: Journal of Engineering in Medicine, ONLINE FIRST (2021).
4. Y. Isler, A. Narin, M. Ozer. Comparison of the effects of cross-validation methods on determining performances of classifiers used in diagnosing congestive heart failure, Measurement Science Review 15(4), (2015) 196-201.

EEG based Spatial Attention Shifts Detection using Time-Frequency features on Empirical Wavelet Transform

Gülçin İnat[1], Gokhan Altan[2]

[1,2] Department of Computer Engineering, Iskenderun Technical University, Turkey
Emails: gulcininat@gmail.com, gokhan.altan@iste.edu.tr

Abstract

The human nervous system has over 100b nerve cells, of which the majority are located in the brain. Electrical alterations, Electroencephalogram (EEG), occur through the interaction of the nerves. EEG is utilized to evaluate event-related potentials, imaginary motor tasks, neurological disorders, spatial attention shifts, and more. In this study, we experimented with 29-channel EEG recordings from 18 healthy individuals. Each EEG was sampled at 250 Hz with a standard reference electrode to the right mastoid. Each recording was decomposed using Empirical Wavelet Transform (EWT), a time-frequency domain analysis technique at the feature extraction stage. Afterwards, the statistical features which were calculated from each modulation are fed into conventional machine learning algorithms. The proposal achieved the best spatial attention shifts detection accuracy using the Decision Tree algorithm with a rate of 89.243%.

Keyword(s): Electroencephalogram (EEG), Empirical Wavelet Transform, EWT, k- NN, SVM, MLP

References

1. Altan, G., Kutlu, Y., & Allahverdi, N. (2016). Deep belief networks based brain activity classification using EEG from slow cortical potentials in stroke. International Journal of Applied Mathematics Electronics and Computers, (Special Issue-1), 205-210.
2. Reichert, C., Tellez Ceja, I. F., Sweeney-Reed, C. M., Heinze, H. J., Hinrichs, H., & Dürschmid, S. (2020). Impact of stimulus features on the performance of a gaze-independent brain-computer interface based on covert spatial attention shifts. Frontiers in Neuroscience, 14, 1250.
3. Hu, Y., Li, F., Li, H., & Liu, C. (2017). An enhanced empirical wavelet transform for noisy and non-stationary signal processing. Digital Signal Processing, 60, 220-229.
4. Altan, G., Yayık, A., & Kutlu, Y. (2021). Deep Learning with ConvNet Predicts Imagery Tasks Through EEG. Neural Processing Letters, 1-16.

High performance detection and locking application for Combat Unmanned Aerial Vehicle (UAV)

Kadir Tohma[1], Yakup Kutlu[2], Samet Koçdaş[3], Samet Tunay[4],
Osman Ünal[5], Osman Emer[6]

[1, 2, 3,4,5,6] Department of Computer Engineering, Iskenderun Technical University, Turkey
Emails: kadir.tohma@iste.edu.tr, yakup.kutlu@iste.edu.tr

Abstract

An Unmanned Aerial Vehicle suitable for the Combat Unmanned Aerial Vehicle category is being designed to be used in competitions within the body of İskenderun Technical University. This UAV has many tasks to perform. The first of these is autonomous locking. The designed UAV should be able to track the rival UAV elements autonomously by processing them on the flight card by means of the camera on the nose part, and be locked to the rival UAV elements for at least 4 seconds. In this study, we performed the detection and locking process by processing the images coming from the camera integrated in the front of the aircraft with the embedded card. According to the results obtained in the simulation environment, the application that we will use in our aircraft detected the competing UAVs at a rate of 90%, and locked to the rival UAV elements in an average of 3.12 seconds.

Keyword(s): Combat Unmanned Aerial Vehicle, UAV

Investigation of Classifier Performances with respect to the Difference of Flickering Frequencies of User Commands in Brain-Computer Interfaces (Oral)

Ebru Sayilgan[1], Yilmaz Kemal Yuce[2], Yalcin Isler[3]

[1] Department of Mechatronics Engineering, Izmir University of Economics, Turkey
[2] Department of Computer Engineering, Alanya Alaaddin Keykubat University, Turkey
[3]Department of Biomedical Engineering, Izmir Katip Celebi University, Turkey
Emails: ebru.sayilgan@ieu.edu.tr, yilmazkemalyuce@gmail.com, islerya@yahoo.com

Abstract

Recently, brain computer interfaces based on the steady-state visually evoked potentials becomes popular [1, 2, 3]. These use flashing user commands on the screen with different rates. Each user command is dedicated to a unique flickering frequency. This study investigates whether the increase in the difference between two dedicated frequencies enhances the classifier performances or not by utilizing the wavelet features. These features are the energy, entropy, and variance of each wavelet decomposition level. Six different mother wavelet functions are tested for this purpose. Seven commonly-used classifiers are investigated in the scope of this study. Mother wavelets of Bior3.5, Haar, and Rbio2.8 gives better mean classifier accuracies, while mother wavelets of Coif1 and Sym4 produces worse results as the frequency difference between user commands increases. The Db4 mother wavelet shows no effect of the frequency difference on the mean classifier accuracies.

Keyword(s): Steady-state visually-evoked potentials, Brain-computer interface, Mother wavelet selection, Machine learning algorithms

References

1. E. Sayilgan, Y.K. Yuce, and Y. Isler, Evaluation of wavelet features selected via statistical evidence from steady-state visually-evoked potentials to predict the stimulating frequency, Journal of the Faculty of Engineering and Architecture of Gazi University, vol. 36(2), 593-605, 2021.
2. E. Sayilgan, Y.K. Yuce, and Y. Isler, Investigating the Effect of Flickering Frequency in Steady-State Visually-Evoked Potentials on Dichotomic Brain-Computer Interfaces, Innovation and Research in BioMedical Engineering, UNDER REVIEW, 2020.
3. E. Sayilgan, Y.K. Yuce, and Y. Isler, Evaluation of Mother Wavelets on Steady-State Visually-Evoked Potentials for Triple-Command Brain-Computer Interfaces, Turkish Journal of Electrical Engineering & Computer Sciences, ONLINE FIRST, 2021.

Modification of posterior probability variable with frequency factor according to Bayes theorem

Mehmet Sait Vural[1], Muhammed Telçeken[2]

[1]Gaziantep Islam Science and Technology University, Turkey
[2]Department of Computer Engineering, Iskenderun Technical University, Turkey
Emails: mehmetsait.vural@gibtu.edu.tr, mtelceken@gmail.com

Abstract

Abstract Probability theory is a branch of science that statistically analyzes random events. Thanks to this branch of science, machine learning techniques are used inferences for the prediction or recommendation system. One of the statistical methods at the forefront of these techniques is Bayesian theory. Bayes is a simple mathematical formula used to calculate conditional probabilities and obtain the best estimates. The two most important parts of the formula are the concepts of a priori probability and posterior/conditional probability. In a priori probability, the most rational assessment of the probability of an outcome is made based on the available data, while in posterior probability, the probability of the event occurring is calculated after considering all evidence or data.

In this study, a new mathematical model is presented to calculate the posterior probability variable of Bayesian theory more precisely. According to this new mathematical model, equal priority probabilities of some variables should be recalculated according to frequency. Calculations are applied to two nodes. The first of these two nodes is the node consisting of the existing data, and the second is the queried node. The positive frequency value will be applied when the variables consisting of existing data and having the same a priori probabilities are found at the questioned node, and negative frequency value will be applied for the other variables. Thus, while calculating a standard probability value according to Bayesian Theory, frequency-based values are taken into account with the help of the newly created mathematical model. With the help of these frequencies, the modification of the system reveals more precise results according to these two basic principles. The results obtained were tested with the cross validation method and high accuracy rates were determined.

Keyword(s): Bayesian Theory, Machine Learning, Mathematical Model

Prediction of Alzheimer's Disease from MRI Images Using Deep Learning Techniques

Ceren Kaya[1], Mertcan Karakoç[2]

[1] Department of Biomedical Engineering, Zonguldak Bulent Ecevit University, Zonguldak, Turkey
[2]Department of Computer Engineering, Zonguldak Bulent Ecevit University, Zonguldak, Turkey
Emails: ceren.kaya@beun.edu.tr, mertcankarakoc92099@gmail.com

Abstract

Alzheimer's disease is a progressive mental deterioration and neurological disorder that can occur in middle or older age and mostly cause dementia because of degeneration of the brain. Early diagnosis of Alzheimer's disease is clinically vital due to irreversible nature of disease progression. In this study, a four-class dataset, obtained by Magnetic Resonance Imaging method and labeled according to Alzheimer's disease subtypes (Non, Moderate, Mild, Very Mild), was used. Three different deep learning models were used for the prediction of the disease. Two of these are models named “Model Basic” and “Model Branched”, which are not based on any model built from scratch. The other one is the ResNet50 based model, which is frequently used in deep learning research. The dataset is divided into three separate parts (60% train, 20% test and 20% validation). Batch size was set to 16 and 50 epochs were carried out in training of all models. In the light of the findings, “Model Branched” model achieved higher classification performance compared to “Model Basic” (97.75%) and ResNet50 (95.12%) models with an accuracy of 98.14%. The application of automatic deep learning tools, which can be highly sensitive for predicting Alzheimer's disease, will help diagnose the disease in its early stages and therefore achieve a better clinical outcome.

Keyword(s): Alzheimer’s Disease; MRI Images; Deep Learning; Model Basic; Model Branched; ResNet50.

Thermally Induced Vibration Suppression in a Thermoelastic Beam Structure

Kenan Yıldırım[1]

[1] Mus Alparslan University, Mus, Turkey
Email: k.yildirim@alparslan.edu.tr

Abstract

In this paper, the problem of thermally induced vibration suppression in a thermoelastic beam is studied. Physical equivalent of the present problem is that a thermoelastic beam is suddenly entering into daylight zone and vibrations are induced due to heating on the upper surface of the beam or thermoelastic beam in a spacecraft enters to intensive sunlight area just after leaving a shadow of a planet. Thermally induced vibrations are suppressed by means of minimum using of control forces to be applied to dynamic space actuators. Objective functional of the problem is chosen as a modified quadratical functional of the kinetic energy of the thermoelastic beam. Necessary optimality condition to be satisfied by an optimal control force is derived in the form of maximum principle, which converts the optimal vibration suppression problem to solving a system of distributed parameters system linked by initial-boundary-terminal conditions. Solution of the system is achieved via MATLAB© and simulated results reveal that thermally induced vibration suppression by means of dynamic space actuators are very effective and robust.

Keyword(s): thermal moment, maximum principle, vibration, thermoelastic

Two Stage System:

Autonomous Plate Detection and Recognition

Water Quality Measurement System in Iskenderun Technical University Pond

Ahmet Gökçen[1]

[1] Department of Computer Engineering, Iskenderun Technical University, Turkey
Email: ahmet.gokcen@iste.edu.tr

Abstract

Water is an indispensable resource for life. Both human beings, plants and animals… Considering that life begins in water, the importance of clean use of water resources is better understood. In this study, water quality measurements of the pond in Iskenderun Technical University were made. For the required measurement system is designed by using Arduino Mega with ATmega2560 series microcontroller, Gravity: Analog pH Sensor, Gravity: Analog Electrical Conductivity Sensor, Waterproof NTC. The obtained data were both written to the SD card and transferred to the database via wireless connection. As a result, the water quality in the pond will be measured in order to maintain the life cycle and necessary conditions will be improved.

Keyword(s): Water Quality, Embedded Systems, Water Sensors

References

1. Kondaveeti, H. K., Kumaravelu, N. K., Vanambathina, S. D., Mathe, S. E., & Vappangi, S. (2021). A systematic literature review on prototyping with Arduino: Applications, challenges, advantages, and limitations. Computer Science Review, 40, 100364.
2. Hong, W. J., Shamsuddin, N., Abas, E., Apong, R. A., Masri, Z., Suhaimi, H., ... & Noh, M. N. A. (2021). Water Quality Monitoring with Arduino Based Sensors. Environments, 8(1), 6.
3. Rao, A. S., Marshall, S., Gubbi, J., Palaniswami, M., Sinnott, R., & Pettigrovet, V. (2013, August). Design of low-cost autonomous water quality monitoring system. In 2013 International Conference on Advances in Computing, Communications and Informatics (ICACCI) (pp. 14-19). IEEE.
4. Salih, N. A. J., Hasan, I. J., & Abdulkhaleq, N. I. (2019). Design and implementation of a smart monitoring system for water quality of fish farms. Indonesian Journal of Electrical Engineering and Computer Science, 14(1), 44-50.
5. Wijnen, B., Anzalone, G. C., & Pearce, J. M. (2014). Open-source mobile water quality testing platform. Journal of Water, Sanitation and Hygiene for Development, 4(3), 532-537.

A review of TDMA Based MAC Protocols with Disjoint Time Slot

İpek Abasıkeleş-Turgut1[1]

[1] Department of Computer Engineering, Iskenderun Technical University, Turkey
Email: ipek.abasikeles@iste.edu.tr

Abstract

VANET is a promising technology used in recent years. With the communication between a vehicle with another vehicle or an infrastructure, both secure and non-secure applications are supported. The communication mostly relies on broadcast-based methodology, which requires a specific MAC design for these networks. MAC protocols for VANETs are divided into three categories included contention-free, contention-based and hybrid. In this article, initially the MAC protocols for VANETs are analyzed. Afterwards, the studies about TDMA based MAC protocols having disjoint time slots are investigated.

Keyword(s): Vehicular Ad Hoc Networks, Medium/Media Access Control, Time Division Multiple Access, Contention-Free, Disjoint Time Slot

1. Introduction

Nowadays, Vehicle Ad-hoc Networks (VANET), an important part of the Intelligent Transportation System (ITS), are attracting great attention in research and industry due to its ability to provide both secure and non-secure applications for various traffic scenarios. VANET is a customized version of Mobile Ad-hoc Networks (MANET) for vehicles, consisting of a set of fixed roadside units (RSU) and vehicles equipped with special embedded units (On-board Units - OBUs) to communicate [1].

VANET provides both secure and non-security services to improve the transportation system. Examples of security applications include key management, authentication, collision alert, and traffic alert. These applications are in charge of protecting the vehicle against any attack and problem that may come from the outside. File sharing applications that support mobility and usability are examples of non-security applications [2]. These applications often require broadcast-based communication of vehicles with each other or with the RSU. Different types of communication models are used according to the diversity of the elements communicating over VANET. In general, VANET communication is divided into vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I). While vehicles communicate directly with each other in V2V model, RSU is used for message transmission in V2I model. When RSU maintenance and distribution costs are taken into account, it can be said that the V2V model is more cost effective than the V2I model [3].

The success of the broadcasts depends on the communication between different units using a common channel within the guaranteed delay time without collision (i.e. collision avoidance) [4]. Media Access Control (MAC) protocols determine and regulate the rules required for units communicating over the network to access the common channel. In VANET, the problems including network topology that changes frequently due to high vehicle mobility, high bandwidth requirement, low latency constraint, hidden / exposed vehicle problem,

unbalanced load distribution in multi-channel use, inefficient use of single channel, collisions on communication links, etc., complicate the MAC protocol design in these networks [5]. Therefore, the MAC protocols proposed for MANET cannot meet the specific requirements of VANET. This situation necessitated the design of new MAC protocols specific to VANET.

2. A General View of MAC protocols for VANET

MAC protocols for VANET basically consist of 3 different categories [6]:

a) Contention-based: Vehicles using the network compete to access the common medium/channel. Access to the environment is random as there are no predetermined / allocated resources or time slots for the nodes. Although a vehicle needs to make sure that the channel is empty by listening before transmitting, if more than one vehicle wants to transmit at the same time, collisions may occur and data loss may occur. For this reason, it is not preferred to use these approaches, especially in heavy traffic.

b) Contention-Free - Based on Coordination (Contention-Free): In this method, the resource is distributed among vehicles by using synchronization schemes. Generally, frequency (Frequency Division Multiple Access - FDMA), time (Time Division Multiple Access - TDMA), code (Code Division Multiple Access - CDMA) and space (Space Division Multiple Access - SDMA) are used for partitioning.

c) Hybrid: It consists of two or more MAC approaches combined in the same protocol.

Although contention-based methods are easy to design, collisions from random channel access cause both data loss and extra delays, especially in high traffic situations [7]. Hybrid protocols make use of coordination-based protocols in addition to contention-based methods to alleviate this problem. Thus, it performs better in heavy traffic. However, this use increases the complexity of the network [8]. Contention-Free methods reduce transmission collisions by providing periodic message exchange and reservation of the channel for vehicles [9]. Therefore, methods based on coordination are examined in this article.

In coordination-based MAC protocols, the environment can be divided into 4 different ways: frequency, time, space and code. In Frequency Division Multiple Access (FDMA) protocols, the transmitter and receiver must be synchronized to the same channel frequency. Therefore, a frequency synchronization algorithm is used to match the communicating vehicles with each other. This algorithm increases both the complexity and the communication load on the network. In Code Division Multiple Access (CDMA) protocols, special codes are assigned to vehicles. At the beginning of each communication, the sender and receiver must agree on codes that will minimize the risk of collision. This requires the use of a CDMA code assignment algorithm to allocate codes on each communication. This algorithm, on the other hand, causes both overhead and increased complexity as in FDMA [10]. Space Division Multiple Access (SDMA) divides the medium geographically between nodes. Vehicles communicate based on their current location. In environments with high mobility, the load of the network increases as the node positions will constantly change [11]. In Time Division Multiple Access (TDMA) protocols, time is shared between nodes. Each node transmits only during its allotted time slot and keeps its radio turned off at other times. Thanks to this method, energy wastage caused by

both eavesdropping and collisions is prevented and the probability of collision is minimized [12]. For this reason, coordination-based TDMA MAC protocols will be examined in the article.

TDMA-based MAC protocols can be designed as centralized or distributed according to the coordinating device. In central management, all control in the network is performed by a coordinator. This coordinator is usually chosen as the cluster head (CH) in RSU or cluster-based architectures. However, this method causes transmission delays to increase, especially in heavy traffic loads [13]. In distributed network design, the elements in the network provide coordination among themselves, ensuring that delays can be tolerated even in high traffic density [14]. Distributed TDMA based MAC protocols can be collaboration based, discrete time slot based, prediction based, game based or enhanced [9]. In collaboration-based MAC protocols, nodes cooperate to communicate with their multi-hop neighbors. However, this method greatly increases the complexity of the network. In prediction-based protocols, collisions are predicted by certain algorithms and precautions are taken early. The algorithms used in collision prediction increase the computational load on the network. In game-based protocols, nodes play a game using certain parameters to determine the time slots to be assigned. Thus, the reservation collision problem in a high-density network is solved. However, this causes both latency and increased load on the network. Taking into account the disadvantages of other approaches, discrete time slot based and coordination based TDMA MAC protocols will be analyzed in this article.

In protocols with a discrete time slot frame structure, the time frames are divided into discrete time slot groups depending on the direction of the vehicle. With this method, merge collision is avoided. Merge collisions occur when nodes using the same time slot but belonging to different clusters approach each other [15]. While some approaches using this method in the literature allocate a special discrete time slot for RSU [1], there are also methods that dynamically change groups according to traffic density [16].

3. Literature Review

Discrete time slot based TDMA MAC studies that have been proposed in the literature in recent years have been examined.

The VeMAC [17] proposed by Hassan Aboubakr O., Weihua Z., Li Li, (2013) is a discrete-time slot TDMA-based MAC protocol that provides efficient and reliable broadcast service in the control channel by solving the hidden terminal problem. By assigning discrete time slots to vehicles moving in opposite directions and RSU, transmission collisions caused by node mobility in the control channel are reduced. In order not to increase the complexity on the network, a fixed time frame structure was preferred. The performance and effectiveness of the proposed method have enabled the VeMAC protocol to form the basis of many new approaches in the literature.

Vasileios D. , Konstantinos O. , Konstantinos G., Ioannis S. (2018) aimed to expand the d-TiMAC algorithm, which is based on the VeMAC protocol, which outperforms the VeMAC protocol in terms of both efficiency and time delay. For this purpose, they proposed the d-TiMAC [19] algorithm, which is an improved version of the TiMAC [18] protocol. Proposed by Imrich C., Andris F., (1994), TiMAC is a basic timing protocol developed for MANET

networks with multi-hop variable topology. The purpose of this protocol is to find solutions to problems arising from mobility. A hybrid design is presented by taking advantageous parts of two different protocol types based on coordination and contention. In the d-TiMAC method, discrete subframes called d-frames, specific to each vehicle's direction, are used. Each frame is split into two different d-frames. One of these frames is for the north-west direction; the other is for the south-east direction. In this way, each direction communicates in discrete time zones. Both TiMAC and d-TiMAC guarantee minimal immunity to motility. For this reason, while they provide an advantage in heavy traffic, they become disadvantaged in sparse traffic.

Ping C., Jun Z., Yuying W.(2017) proposed an adaptive tool MAC protocol for VANETs called A-VeMAC [16]. A-VeMAC is a multi-channel TDMA MAC protocol based on VeMAC. It supports one control channel and several service channels. For the control channel, each frame is divided into two discrete time slot groups, which are respectively associated with vehicles traveling in opposite directions. Unlike VeMAC, which splits each frame equally, A-VeMAC does not partition frames equally. Instead, it adapts to vehicle traffic conditions traveling in opposite directions. The aim is to better support unstable vehicle traffic conditions in opposite directions.

Shufing L., Yanheng L., Jian W. (2019) proposed ASTSMAC [20], an appropriate time slot shared TDMA MAC protocol, to provide solutions to packet delivery rate, collision rate and heavy traffic performance issues. They aimed to increase the performance of the network by reducing unnecessary bandwidth waste and meeting the needs of more vehicles. Basically, two methods were followed for this. In the first method, if the distance between two vehicles allows for no transmission collision, these vehicles are allowed to use the same time slot. Thus, it is ensured that more vehicles are supported with the same amount of time slots. In the second method, DAFA (frequency tuning algorithm) is presented to provide frequency tuning depending on the density of the nodes. The disadvantage of this approach is the use of wasted time slot, which increases in direct proportion to the number of collisions in the network, since both vehicles simultaneously empty the time slot when a collision occurs.

4. Conclusion

Due to specific characteristics of VANETs including high mobility and low latency requirement, new MAC protocols should be designed for these networks. Among all types of VANET MAC protocols, contention-free methods are the most preferable. Energy saving is gained while collisions are prevented in TDMA-based contention-free MAC protocols. In this article, discrete time slot based and coordination based TDMA MAC protocols are analyzed. The aim of the future work is to design a new discrete time slot based TDMA MAC protocol that outperforms existing studies on performance parameters such as packet transmission rate, delay rate, idle time slot usage in different traffic densities.

References

1. Y. Wang, J. Shi, L. Chen, B. Lu and Q. Yang, "A Novel Capture-Aware TDMA-Based MAC Protocol for Safety Messages Broadcast in Vehicular Ad Hoc Networks," in IEEE Access, vol. 7, pp. 116542-116554, 2019, doi: 10.1109/ACCESS.2019.2936144.

2. M.A. Razzaque, Ahmad Salehi S., and Seyed M. Cheraghi, “Security and Privacy in Vehicular Ad-Hoc Networks: Survey and the Road Ahead” Wireless Networks and Security, SCT, pp. 107–132, doi: 10.1007/978-3-642-36169-2_4
3. Ghulam A., Ziaul Haq A., Shahab H., Thar B., Saadi B., Fazal M. “PDMAC: A Priority-Based Enhanced TDMA Protocol for Warning Message Dissemination in VANETs” Sensors 2020, 20, 45, doi:10.3390/s20010045
4. F. Lyu et al., "MoMAC: Mobility-Aware and Collision-Avoidance MAC for Safety Applications in VANETs," in IEEE Transactions on Vehicular Technology, vol. 67, no. 11, pp. 10590-10602, Nov. 2018, doi: 10.1109/TVT.2018.2866496.
5. Ankit N. D. , Rachna M. P. , “A Review on MAC Protocols of Vehicular Ad Hoc Networks” IEEE WiSPNET 2016 conference, doi: 978-1-4673-9338-6/16/ 2016 IEEE
6. Zain, I. F. M., Awang, A., & Laouiti, A. (2017). “Hybrid MAC protocols in VANET: A survey. In Vehicular ad-hoc networks for smart cities” (pp. 3-14). Springer, Singapore.
7. Gaurav S., Ayalvadi G., and Peter K., “Performance Analysis of Contention Based Medium Access Control Protocols” IEEE TRANSACTIONS ON INFORMATION THEORY, VOL. 55, NO. 4, APRIL 2009 doi: 0018-9448
8. Vandana J., C. Hemanth and R.G. Sangeetha, “A survey on hybrid MAC protocols for vehicular ad-hoc networks” in Vehicular Communications m5G; v1.186 2016 doi: 2214-2096
9. Suchi Johari, M. Bala Krishna, “TDMA based contention-free MAC protocols for vehicular ad hoc networks: A survey” 10.1016/j.vehcom.2020.100308
10. Mohamed H., Paul M., Anis L., Rachid Z. and Leila Azouz S., “TDMA-based MAC Protocols for Vehicular Ad Hoc Networks A Survey, Qualitative Analysis and Open Research Issues” IEEE COMMUNICATION SURVEYS AND TUTORIALS , VOL. , NO. , MAY 2015 doi: 10.1109/COMST.2015.2440374
11. Zouina D., Sofian Ali B., Kayhan Zrar G., Iman L. and Samira M., “Load balancing aware SDMA-based beaconing approach in vehicular ad hoc networks” Institut Mines-Telecom and Springer-Verlag France 2017 doi: 10.1007/s12243-017-0562-y
12. Tarunpreet K. and Dilip K., “TDMA-Based MAC Protocols for Wireless Sensor Networks: A Survey and Comparative Analysis” in International Conference on Wireless Networks(WECON) 2016 doi: 10.1109/WECON.2016.7993426
13. Rongqing Z., Xiang C., Liuqing Y., Xia S., and Bingli J., “A Novel Centralized TDMA-Based Scheduling Protocol for Vehicular Networks” IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS doi: 10.1109/TITS.2014.2335746
14. Sailesh B., Weihua Z., Lakshmi V. T., and Fan B., “Link-Layer Cooperation Based on Distributed TDMA MAC for Vehicular Networks” IEEE Transactions on Vehicular Technology doi: 10.1109/TVT.2016.2634545
15. Tomotaka K. and Koji K., “Modified TDMA-Based MAC Protocol for Vehicular Ad Hoc Networks” The 12th IEEE International Workshop on Managing Ubiquitous Communications and Services, 2015 doi: 10.1109/PERCOMW.2015.7134000
16. Ping C., Jun Z., Yuying W., “A-VeMAC: An Adaptive Vehicular MAC Protocol for Vehicular Ad Hoc Networks” IEEE ICC 2017 Ad-Hoc and Sensor Networking Symposium doi: 10.1109/ICC.2017.7997358
17. Hassan Aboubakr O., Weihua Z., Li L., “VeMAC: A TDMA-Based MAC Protocol for Reliable Broadcast in IEEE TRANSACTIONS ON MOBILE COMPUTING, VOL. 12, NO. 9, SEPTEMBER 2013 doi: 10.1109/TMC.2012.142
18. Chlamtac, A. Farago, “Making transmission schedules immune to topology changes in multi-hop packet radio networks” IEEWACM TRANSACTIONS ON NETWORKING, VOL. 2, NO. 1, FEBRUARY 1994 doi: 10.1109/90.282605
19. Vasileios D., Konstantinos O., Konstantinos G., “A disjoint frame topology-independent TDMA MAC policy for safety applications in vehicular networks” Ad Hoc Networks 79 (2018) 43-52 doi: 10.1016/j.adhoc.2018.06.006
20. Shufing L., Yanheng L., Jian W. “ASTSMAC: Application Suitable Time-Slot Sharing MAC Protocol for Vehicular Ad Hoc Networks” in IEEE ACCESS VOLUME 7, 2019 doi: 10.1109/ACCESS.2019.2936612

Building a Hybrid Recommendation System for E-Commerce

Hakkı Şimşek[1], Mustafa Yeniad[2]

[1,2] Department of Computer Engineering, Ankara Yıldırım Beyazıt University, Turkey
Email: hakkisimsek92@gmail.com, myeniad@gmail.com

Abstract

With the technology occupying a large place in human life, our shopping habits have also changed drastically. Increasing product variety and alternatives have made it difficult for people to reach products that suit their own tastes. Thanks to the Suggestion Systems, it has made internet shopping easier by using the evaluation, comment and scoring criteria of other users who have used the same or similar products to the products they are interested in. By learning the tendencies of the users while choosing the products, the presentation of the appropriate products enabled the users to reach their requests easily and quickly. At the same time, it has become mandatory for e-commercial companies to learn the preferences of users in order to compete with other companies. In this study, the most popular recommendation systems and algorithms used in e-commerce platforms are compared.

Keyword(s): Recommendation systems, hybrid recommendation systems for e-commerce, e-commerce shoppings

1. Introduction

The reshaping of business practices and consumer habits together with the high development of technology has led to a rapid growth in e-commerce applications. One of the most important purposes of e-commerce is to ensure that products or services are delivered to users quickly and without any problems. Factors such as lower costs, savings in time, and reduction in storage costs have made e-commerce activities widespread. The efforts of e-commerce sites to increase their product and service sales potential have led to the emergence of the Recommendation System. Recommendation Systems is a system that aims to provide personalized products and services to users and provides easy purchasing opportunities by using various machine learning algorithms [1]. The increasing variety of products and the fact that each user has different experiences and tastes has created the question of which product will be recommended to whom. Suggestion Systems solve this problem by predicting the interests of the users and recommending the most suitable product for them.

Recommendation Systems learn from users' choices by taking into account the data provided by users and their preferences, interests and experiences. In this information, it finds the match between the user and the product-service, imposes the similarity between the user and the items, and offers them the most appropriate suggestion in line with the needs of the users [2].

Recommendation systems have also become very important as many large companies such as Youtube, Spotify, Netflix, Amazon have improved their service delivery practices. In 2009, Netflix offered a $1 million prize by holding a competition because it planned to produce a recommendation system that outperformed the algorithm it used [3]. The knowledge that another big company, Amazon, also generates 35% of its revenue from recommendation

systems, reveals how important recommendation systems are in being able to compete with other companies and increase their sales [4].

Recommendation Systems work with two basic types of information. These are characteristic information and user-item interactions. The characteristic information is information about the item and users. While the item consists of elements such as category and keyword, the information about the user includes the interests and likes of the users. The results reflecting the preferences are reached by taking into account the information provided by the users during registration and the actions and transactions they have made within the site afterwards. User-item interactions include elements such as the number of likes, ratings, comments and purchases on products by users. Ratings can be either 1-5 or 1-10 stars, or two, such as like-dislike [1].

There are four different recommendation systems. These; Collaborative Recommendation Systems, Content-Based Recommendation Systems, Popularity-Based Recommender Systems, and Hybrid Recommendation Systems. In this study, the use of these four different recommendation systems in e-commerce sites was examined and the algorithms were compared with each other.

Collaborative Recommendation Systems is one of the most widely used recommendation systems. It is purely based on past experience and the similarity between two users' preferences and tastes [5]. This recommendation system collects and analyzes information about users' interests and tries to predict which products users will like, taking into account the similarity with other users. When it finds similar users, it analyzes how much the likes overlap with each other and offers products accordingly [6].

In Content-Based Recommendation Systems, not only the user-item interaction but also the item information that the user has interacted with in the past is utilized [3]. This recommendation system is based on the idea of presenting items that are similar to users' past tastes. For example, it is recommended to the user by checking the movies that are similar to the content of the movie the user is watching [7].

Popularity-Based Recommendation Systems reveal popular or trending items among users and recommend them directly to the user. Unlike other recommendation systems, it is not affected by the user's preferences [2].

Hybrid Recommender Systems are based on the combined use of Collaborative Recommender Systems and Content-Based Recommendation Systems [7]. In order to achieve the best performance, a model can be created and used by adding content-based features to Collaborative Suggestion Systems or by applying vice versa scenarios [4]. By using these two systems together, it is aimed to get more efficient results by preventing the weaknesses of each of the systems. Big companies such as Netflix, Amazon, Spotify benefit greatly from this system. For example, Netflix offers suggestions both by checking users' past actions and taking into account the ratings made by users [8].

1.1.Problems in Recommendation Systems

The inclusion of new users and products in the system, the increase in user evaluations about the product, and the increased data may cause some problems.

Data sparsity is one of these problems. It occurs when similarity recognition between users is insufficient despite appropriate data [9].

Not having enough information about users who are new to the system creates an uncertain situation about the products to be offered to them [10].

In this study, we tried to reveal the right recommendation system to avoid certain problems encountered in the system and to make appropriate recommendations to the right users.

2. Material and Methods

Python libraries (pandas, NumPy and sci-kit learn) were used separately for Collaborative Recommendation Systems, Content-Based Recommendation Systems and Hybrid Recommendation Systems.

The NumPy library is like Python but more functional and used for scientific calculations. The Pandas library is a library that can be effectively used in operations with non-homogeneous sequences. Sci-kit learn is a convenient library for machine learning-oriented operations such as regression, matrices, and vectors. Cosine similarity is calculated by formulating the similarity between two different texts with the cosine function. SVD allows operations to be made with a complex matrix as the product of simpler matrices. In TfidfVectorizer, the processed words are converted to TF / IDF matrix and the similarity of the products is measured by considering the product features.

The live database of an e-commerce website was used as the dataset which includes a total of 5012 different orders and 7190 products. Firstly, no recommendation system was used for the first 15 days. Next, only the collaborative recommendation system was used. After that, a content-based recommendation system was used. Lastly, the hybrid recommendation system was used in 15-day periods.

In the first period, 344 orders were received without using any recommendation system, and a total of 501 products were sold in these orders.

In the second period, the collaborative recommendation system was used, 498 orders and 776 products were sold. When the whole dataset is examined, 36 orders and 45 products are matched with the collaborative recommendation system.

In the third period, the content-based recommendation system was used, 730 orders and 1105 products were sold. When the entire dataset is analyzed, 126 orders and 247 products match the content-based recommendation system.

In the third (last) period, the hybrid recommendation system was used, 1399 orders and 2241 products were sold. When the entire dataset is examined, 248 orders and 420 products match the hybrid recommendation system.

2.1. Comparison of Recommendation Systems

Users with at least 1 order were evaluated. Accordingly, in the dataset with 4694 users, the number of users with at least 1 order was 4260, and the number of user interactions with at least 3 orders was 1254, with a total of 5012 orders. The total number of products is 378.

In this study, TF/IDF technique (Term Frequency / Inverse Document Frequency) was used. However, for every product that the user interacts with;

- 30 products that the user does not interact with are sampled.
- The user is not aware of the existence of these products, because they have nothing to do with them.
- A list of 1 interactive and 30 non-interactive products is requested from the suggestion system.
- “Top N” accuracy measurements are made for the user and the interacted product from the list.
- All "first N grains" measurements are summed.

3. Results

For Collaborative Recommendation Systems;

Table 3.1 shows the number and proportions of products in orders and orders every other day, among 776 products out of 498 orders.

In the cooperation suggestion system, the order product ratio is between 1 and 2. This means that at least 1 and at most 2 products are ordered in 1 order.

In the collaborative recommendation system, the rate of purchasing a product in an order is 1.5582.

Table 3-1. Order and Products in Collaborative Recommendation System

Order	Number of Orders	Number of Products	Products / Orders
Order 1	2	2	1
Order 2	1	1	1
Order 3	1	2	2
Order 4	3	4	1,33
Order 5	1	2	2
Order 6	4	4	1
Order 7	2	3	1,5
Order 8	2	2	1
Order 9	2	2	1
Order 10	3	5	1,66

For Content Based Recommendation Systems;

Table 3.2 shows the number and proportions of the products in the order and every two days among 1105 products placed in 730 orders.

In the content-based recommendation system, the order product ratio is between 1 and 3. This means that at least 1 and at most 3 products are ordered in 1 order.

In the content-based recommendation system, the rate of purchasing a product in an order is 1.5136.

Table 3-2. Order and Products in Content Based Recommendation System

Order	**Number of Orders**	**Number of Products**	**Products / Orders**
Order 1	3	5	1,66
Order 2	3	4	1,33
Order 3	2	2	1
Order 4	1	3	3
Order 5	3	6	2
Order 6	2	3	1,5
Order 7	3	3	1
Order 8	3	4	1,33
Order 9	5	6	1,2
Order 10	4	5	1,25

For Hybrid Recommendation Systems;

Table 3.3 shows the number and proportions of products ordered every other day, among 2241 products placed in 1399 orders.

In the content-based recommendation system, the order product ratio is between 1 and 8. This means that at least 1 and at most 8 products are ordered in 1 order.

In the hybrid recommendation system, the rate of purchasing a product in an order is 1.6018.

Table 3-3. Order and Products in Hybrid Recommendation System

Order	Number of Orders	Number of Products	Products / Orders
Order 1	3	24	8
Order 2	4	11	2,25
Order 3	2	2	1
Order 4	1	1	1
Order 5	2	3	1,5
Order 6	2	3	1,5
Order 7	2	5	2,5
Order 8	3	4	1,33
Order 9	2	2	3
Order 10	1	2	2

4. Conclusion

The percentage of orders that do not use recommendation systems is 11.57 out of a total of 2971 orders in Table 4.1. The percentage of collaborative recommendation systems is 16.76. The percentage of content-based recommendation systems is 24.57. The percentage of hybrid recommendation systems is 47.08. The highest number of orders were received when using hybrid recommendation systems. In addition to the highest number of orders, the highest order/product ratio was also seen in hybrid recommendation systems. As future work, all e-commerce companies will be able to use recommendation systems. Currently, large companies (Amazon, Netflix) are successfully using recommendation systems. Medium-sized and small companies cannot use the recommendation systems due to the large size of the data and the high operating costs. Our renewed recommendation algorithms will make recommendation systems more accessible by managing big data with less cost.

Table 4-1. Comparison of Methods with Each Other

Recommendation Systems	Products / Orders
Collaborative Recommendation Systems	1,5582
Content-Based Recommendation Systems	1,5136.
Hybrid Recommendation Systems	1,6018.

References

1. G. Rodriguez, «Introduction to Recommender Systems,» 2018. [Online]. Available: https://tryolabs.com/blog/introduction-to-recommender-systems/.

2. R. Dwivedi, «What Are Recommendation Systems in Machine Learning?,» 16 April 2020. [Online]. Available: https://www.analyticssteps.com/blogs/what-are-recommendation-systems-machine-learning..
3. B. Rocca, «Introduction to recommender systems,» 3 June 2019. [Online]. Available: https://towardsdatascience.com/introduction-to-recommender-systems-6c66cf15ada.
4. İ. Çilingir, «Öneri Sistemleri (Recommendation Systems),» 24 June 2019. [Online]. Available: https://medium.com/@irmcilingir/%C3%B6neri-sistemleri-recommendation-systems-28a3f341c0a9.
5. J. Le, «The 4 Recommendation Engines That Can Predict Your Movie Tastes,» 1 May 2019. [Online]. Available: https://towardsdatascience.com/the-4-recommendation-engines-that-can-predict-your-movie-tastes-109dc4e10c52.
6. R. B. Allen, «User models: theory, method, and practice,» International Journal of man-machine Studies, cilt 32, no. 5, pp. 511-543, 1990.
7. H. Özkök, «Recommendation Engine (Tavsiye-Öneri Sistemleri),» 17 April 2020. [Online]. Available: https://www.datasciencearth.com/recommendation-engine-tavsiye-oneri-sistemleri/.
8. Ş. Güler, «Öneri sistemleri ve E-ticarette öneri sistemlerinin kullanımı,» (Master's thesis, Sakarya Üniversitesi), 2019.
9. M. Papagelis, D. Plexousakis ve T. Kutsuras, «Alleviating the Sparsity Problem of Collaborative Filtering Using Trust,» iTrust‟05 Proceeding of 3rd International on Trust Management, 2005.
10. J. Bobadilla, F. Ortega ve J. Bernal, «A collaborative filtering approach to mitigate the new user cold start problem,» Knowledge-based systems, pp. 225-238, 2012.

CyberLog: A Generalized IoT Platform Training to Protect Cyber Defence Agents and Algorithms

Bilal Olalere Ashade[1], Mustafa Yeniad[2]

[1,2] Department of Computer Engineering, Ankara Yıldırım Beyazıt University, Turkey
Emails: ashade.bilal1023@gmail.com, myeniad@gmail.com

Abstract

The world has advanced to the era of Internet of Things (IoT) which has enabled connectivity from traditional computing devices to smart devices. This development has ushered in various real-time services and smart solutions such as e-transport, e-health, e-manufacturing and other e-solutions that are made available either through data centers or edge devices on a scale that has never been seen before. Interestingly enough, IoT has also brought with it a lot of security challenges due to the number of attack surfaces, and the complexity, heterogeneity and number of participating devices and infrastructure. This has resulted in a rise in the cases and instances of cyber attacks which causes a lot of concern for Government, companies, and consumers alike. This study aims to leverage the recent advances in computational intelligence with the old and long-standing power of statistical analysis to combat this trend and restore confidence in the technology. To achieve this goal, an integrated internet-IoT attack simulation environment has been developed, which allows the configuration of attacks and simultaneous training of appropriate machine learning defense algorithms and can be easily integrated into the existing corporate infrastructure.

Keywords: Cybersecurity, machine learning, internet of things, attack vectors, knowledge discovery

1. Background

As seen in [3], the number of incidents in SCADA systems are increasing due to 0-day [1] attacks. Attackers find new ways of penetrating SCADA systems and these are viewed by nations as a new attack vector of cyber-warfare [6]. As a result of this, nations began to realize that this is a new force in their military operations and they are teaming up for cyber-warfare operations. Additionally, devices connected to the internet would be expected to exceed the number of human beings such that by 2020 we should have fifty billion devices connected [9].

IoT offers a platform whereby these connected devices blend with the actuators or physical objects in the environment in such a way as to enable seamless sharing of information and provision of better services [10].

IoT has enabled ubiquitous connectivity for a wide range of devices, services, and applications (eg. computers, smartphones, office equipment, autonomous cars, lighting systems, heating, ventilation and air-condition (HVAC) and household appliances) [10]. A good number of communication network technologies such as 3G, 4G, 5G, LTE, Wi-Fi, and Bluetooth have provided connectivity services for IoT deployment [10]. In addition, cloud computing is expected to provide the backbone for the worldwide information diffusion, data analytics and storage on IoT enabled devices [10]. This would eventually lead to intelligence being embedded in our environment.

The major concern in IoT-Cloud-edge device integration is that they are riddled with various security challenges such as application services attack, data integrity attack, privacy, trust, identity, standardization, Denial of service attack [7]. This will in turn threaten the security of Government, and organization assets while undermining the integrity of the services provided [10].

By providing a common internet-IoT-attack simulation environment we can better understand what might go wrong and how to avoid them. The knowledge gained from such environment would be packaged as a software API to be deployed in existing IoT infrastructures. This would help all stake-holders in understanding and preventing attacks while protecting public assets from damage caused by such attacks.

2. Agent Platform

Today, many manufacturers have software and hardware solutions in the fields of smart manufacturing, oil and gas refineries, nuclear power plants and vice versa. The usage of IoT edge devices is increasing everyday. On the other hand, the method of measuring any sensor on IoT devices and platforms are very hard and thus designing an efficient platform for getting measurements correctly, analytic processing and machine learning computations on these data in near-real time is a hot topic in real-world scenarios [4]. In order to solve this end to end distributed sensor reading and analytical computation on all sensor readings problem, multi-agent system [8] software is developed.

In CyberML ecosystem, with the help of multi-agent systems [8] , it will be ensured that the measurements made at more than one point are combined and then the inferences will be made using smart algorithms.

One of the multi-agent platform we will be using is Volltron [2]. It is an open-source platform for distributed sensing and control for performing simulations on smart-building energy consumption usage, and creating a more flexible and reliable energy sensor grid. On the contrary this platform is suitable for all types of multi-agent modelling and simulation scenarios especially for the SCADA based critical infrastructures.

Researchers are using Volttron for testing new control methods on a multi-agent setup and can serve as an ideal platform for rapidly making a prototype of new control applications. In addition, the open-source gives all freedom to the implement any research idea on this platform. Such openness also reflects Volttron's integration with other platforms and software.

The concept of an agent is defined as a computer system capable of autonomous action in a situated environment. An agent can react to information sent by the operator, the environment or other agents in real time. Volttron agents are generally divided into three classes. Platform agents serve as service providers in the platform. Cloud agents serve as bridges between the platform and other remote platforms. Control agents control the actual hardware devices.

3. Research Questions

It should be remarked here that there is no single integrated internet-IoT-attack standard simulation environment. In addition, present solutions are focused on traditional filterbased approaches that fall flat in the face of sophisticated attacks. The challenge ahead could be

possibly simplified when viewed through the lens of big data and machine learning. Possible questions to be answered are highlighted below:

- Can Big Data help understand attack signatures?
- Can Machine learning be effective in understanding an attack before the attack is launched?
- Could community detection algorithms give us ideas about the characteristics of good and bad signals or nodes?
- Can we understand the statistical properties of a normal versus compromised payload and connection profile?
- Can we understand the threat profile of a rogue device that is being inserted into an IoT network?
- Can we learn the statistical properties of compromised network sensor nodes?
- Can probabilistic graphical models be effective in modelling and structuring the dependency among different attack types?
- Can probabilistic graphical models help model the damage depth and rate after an attack?

The following analytical methods could be employed to answer some of the questions above:

- Machine learning (supervised, unsupervised and Reinforcement learning)
- Game theoretic approach
- Statistical and Network analysis

As remarked earlier a standard training environment is still lacking.

4. Research Methods

The basic approach to enable the solutions above would be to build a platform that enables the training and testing of algorithms. In addition, such platform should enable the configuration of different attack scenario. The next section describes the simulation environment.

Cyberlog: Simulation Platform

The simulation environment is intended to integrate internet web devices and sensor nodes. The core of the simulation environment is listed below:

- GNS3 (Traditional network architecture and simulation)
- Cooja simulator (Wireless sensor network) or TinyOS
- Docker
- Other physical devices and sensors as necessary.
- GNS3-WSN combo environment.
- IoT attack analysis API (new)
- Volttron Multi-Agent Platform

The analytical nodes gets the data and train the algorithms accordingly. The sensor nodes simulate IoT network and publish the data to exterior devices. A typical sensor network is depicted below:

In particular multiple nodes could be connected to GNS3. Consequently, each node could run a docker version of available wireless sensor network. Using Message Queuing Telemetry Transport (MQTT) [11] it would be possible to feed traffic from the sensor nodes in the virtual machine to external nodes or routers.

With this platform it is possible to build a specific attack scenario or general attack profile of interest.This would be followed by investigation of how to detect and prevent such attack. Even though there are internet-iot data dump as proposed in [5], the aim of the platform is to enable online learning. However, various data dumps could be employed to facilitate learning.

After the learning process it is the responsibility of the
Volttron platform for loading models and related algorithms for further runtime action on the actual real-world platform.

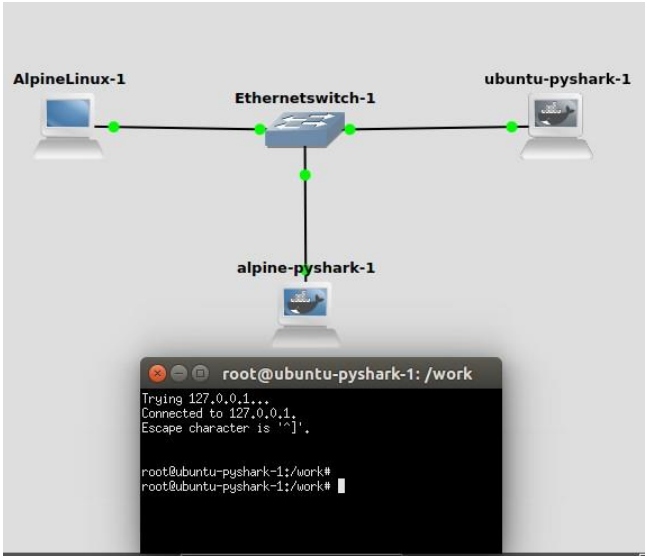


Figure 1. Simplified Network architecture.The ubuntu and alpine pyshark nodes are analytical nodes that will house the models and algorithms.

Figure 2. Simplified Sensor network (with 6LowPan and other IoT protocols). The sensor network would be connected to the internet architecture to simulate a true integrated internet-iot-attack environment

A minimum example of the platform is documented on Github[1]. Contribution to the project from the larger research community is highly welcome.

[1] https://github.com/bilalashade/IoTML-TestEnv-master.git

5. Conclusion

It can be said that, the importance of IoT is its influence. It goes beyond the simple exchange of data by connecting devices. However, there is global concern about the security challenges that such architecture poses. This study proposes a step with enormous potential for providing setups, models and algorithms that can be used to ensure the security of devices, structures, and services at all levels in realm of IoT.

References

1. V. L. Do, L. Fillatre, I. Nikiforov, and P. Willett. 2017. Security of SCADA systems against cyber–physical attacks*. IEEE Aerospace and Electronic Systems Magazine* 32, 5 (2017), 28–45.
2. J. Haack, B. Akyol, N. Tenney, B. Carpenter, R. Pratt, and T. Carroll. 2013. VOLTTRON™: An agent platform for integrating electric vehicles and Smart Grid. In *2013 International Conference on Connected Vehicles and Expo (ICCVE)*. 81–86.
3. Bill Miller and Dale Rowe. 2012. A Survey SCADA of and Critical Infrastructure Incidents. In *Proceedings of the 1st Annual Conference on Research in Information Technology (RIIT '12)*. Association for Computing Machinery, New York, NY, USA, 51–56. https://doi.org/10.1145/2380790.2380805
4. Valerio Morfino and Salvatore Rampone. 2020. Towards Near-RealTime Intrusion Detection for IoT Devices using Supervised Learning and Apache Spark. *Electronics* 9, 3 (Mar 2020), 444. https://doi.org/10. 3390/electronics9030444
5. Nour Moustafa. 2017. *Designing an online and reliable statistical anomaly detection framework for dealing with large high-speed network traffic.* Ph.D. Dissertation. University of New South Wales, Canberra, Australia.

A. Nicholson, S. Webber, S. Dyer, T. Patel, and H. Janicke. 2012. SCADA security in the light of Cyber Warfare. *Computers Security* 31, 4 (2012), 418 – 436. https://doi.org/10.1016/j.cose.2012.02.009

6. Jesus Pacheco and Salim Hariri. 2016. IoT security framework for smart cyber infrastructures. In *2016 IEEE 1st International Workshops on Foundations and Applications of Self* Systems (FAS* W)*. IEEE, 242– 247.
7. C. Savaglio, G. Fortino, and M. Zhou. 2016. Towards interoperable, cognitive and autonomic IoT systems: An agent-based approach. In *2016 IEEE 3rd World Forum on Internet of Things (WF-IoT)*. 58–63.
8. Statista. 2018. Internet of Things (IoT) connected devices installed base worldwide from 2015 to 2025 (in billions). (2018).
9. Samuel Tweneboah-Koduah, Knud Erik Skouby, and Reza Tadayoni. 2017. Cyber security threats to IoT applications and service domains. *Wireless Personal Communications* 95, 1 (2017), 169–185.
10. Wikipedia contributors. 2020. MQTT — Wikipedia, The Free Encycloped. https://en.wikipedia.org/w/index.php?title=MQTT&oldid= 974834571 [Online; accessed 31-August-2020].

Extraction of Lead Oxide from Car Batteries

Gürkan Ergün[1], Hüsnügül Yılmaz Atay[1]

[1]Department of Material Science and Engineering, Izmir Katip Celebi University, Turkey
Emails: hgulyilmaz@gmail.com, gurkan.erg@gmail.com

Abstract

The use of wasted lead–acid batteries in the recycling of lead oxide plays a very valuable role. Lead is a material very easy to recycle and, provided that adequate procedures are implemented, the final product (secondary lead) is indistinguishable from the primary lead produced from ores. About 50% of the lead consumed worldwide is derived from recycled and reused materials. [1]The recovery of metals from metal scrap has the advantage that it is easier and far less energy dependent than the production of primary lead from ores. The production of recycled lead requires 35–40% of the energy necessary to produce lead from ores. In addition, the recovery of lead decreases the lead dispersion in the environment and preserves the mineral reserves for the future. It is estimated that at least 85% of the consumed lead can potentially be recycled [8], [4], [5].

This study shows extraction of lead oxide from car batteries and the XRD patterns of the final outcome proves perfect results. Moreover, chemical analysis (Sulphate (SO4) precipitation process) and characterization results such as particle size analysis, Vibrating-Sample Magnetometer) and Scanning Electron Microscopy will be shown in details.

Keyword(s): Energy Saving, Recycling, Car Batteries, Lead Oxide

1.Introduction

The lead–acid batteries represent about 60% of batteries sold in the entire world. Lead is a material very easy to recycle and, provided that adequate procedures are implemented, the final product (secondary lead) is indistinguishable from the primary lead produced from ores. [3] About 50% of the lead consumed worldwide is derived from recycled and reused materials. The recovery of metals from metal scrap has the advantage that it is easier and far less energy dependent than the production of primary lead from ores. The production of recycled lead requires 35–40% of the energy necessary to produce lead from ores. [2] In addition, the recovery of lead decreases the lead dispersion in the environment and preserves the mineral reserves for the future. It is estimated that at least 85% of the consumed lead can potentially be recycled. [7]

2. Materials and Methods

2.1. Extraction of Lead Oxide Plates from Recycled Batteries

The procedure to extract of lead oxide plates from the battery was pretty straightforward. It can be summarized in 4 steps:

1. Cutting and removing the head of the battery.

2. Removal of Lead Oxide plates from PE separators.

3. Holding the plates under water supply (during 1 hour) and cleaning process.

4. The plates held and dried at 100 Celsius in the oven for 1 hour.

After those steps, dried Lead Oxide are ready to be extracted from their own plates. PbO is so gentle and soft that it can be extracted from the plates by an ordinary tweezer. The plate is divided into 3 level categories - Upper, lower & intermediate parts. The extracted powder of Lead Oxide is put into a plastic bag and the bag is located inside a crucible. After beating the bag with a mallet, the powder of 3 different Lead Oxide is ready to be used for characterization.

2.2. Characterization of Recycled Lead Oxide

2.2.1. Chemical Analysis

Precipitation of PbSO4

Sulfate analysis from plate; 10 g battery paste which is called “sample” was taken and put to moisture analyzer in order to get rid of moisture in the sample. After that, the sample was put into beaker. 100 ml Na2CO3 was added and stirred gently and then solution was boiled for a 20 minutes. At the end of the boiling process, the solution was taken tp the volumetric flask (500 ml) and some water added. Then the solution was filtrated and it was re-heated. Concentrated HCl was added to neutralize the solution. While the solution was boiling, 40 ml BaCl2 was added. After waiting for a while, the solution was filtrated. The precipitate was taken into crucible and kept in the oven at 600 - 700 Celsius for 3 hours. Lastly, remained ash content on crucible was measured and amount of BaSO4 was calculated. [10] All analysis were made in **İNCİ GS Yuasa Akü** Manisa, Turkey.

Color Indication

Where the upper side of the battery plate the current comes first has got more transformation because the magnitude of electricity transmission drops to the lower side, loss increases on the longer path. Therefore, the upper part works better and the lower part works less. There is also an acid density difference in the battery. The lower part has got more concentrated acid. As the acid density increases, the conversion of PbSO4 with current to PbO2 slows down. This reduces the conversion of the lower part from PbSO4 to PbO2, so the lower part goes to the mechanism we call “**Sulfation**”. Since lead sulphate is white, the lower part looks lighter as color. Since the mixture of black PbO2 and PbSO4 is more intense at the bottom, the color turns brown.

2.2.2. Scanning Electron Microscopy (SEM)

SEM technique was used in order to investigate particle size and close look at 3 different parts of the plate (upper, middle and bottom parts). All results were obtained in Central Research Laboratory in Katip Çelebi University.

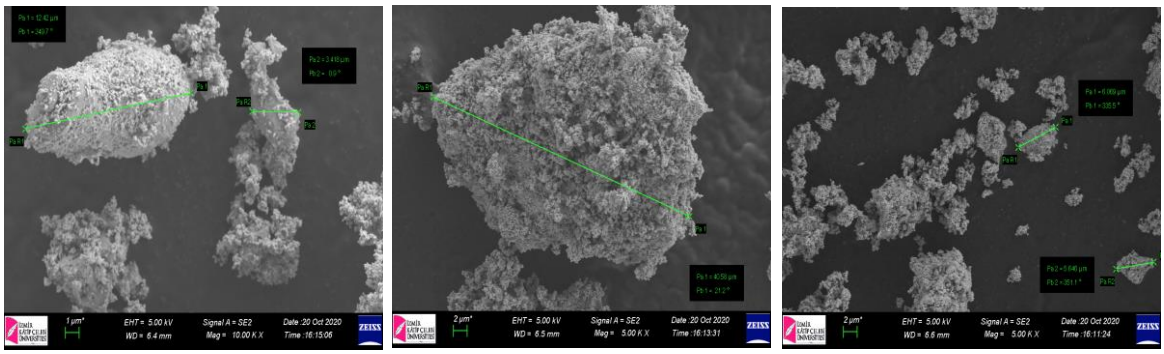


Figure 1. SEM images of Lead Oxide powder taken from the lower (left), the intermediate (middle) and the upper part (right).

2.2.3. VSM (Vibrating-Sample Magnetometer) Results

Hysteresis curves of 3 different levels of PbO_2 recycled from the plates and styrofoam (for comparison) were also recorded by sweeping the magnetic field in magnetometer. All observations made in EMUM in Dokuz Eylül University.

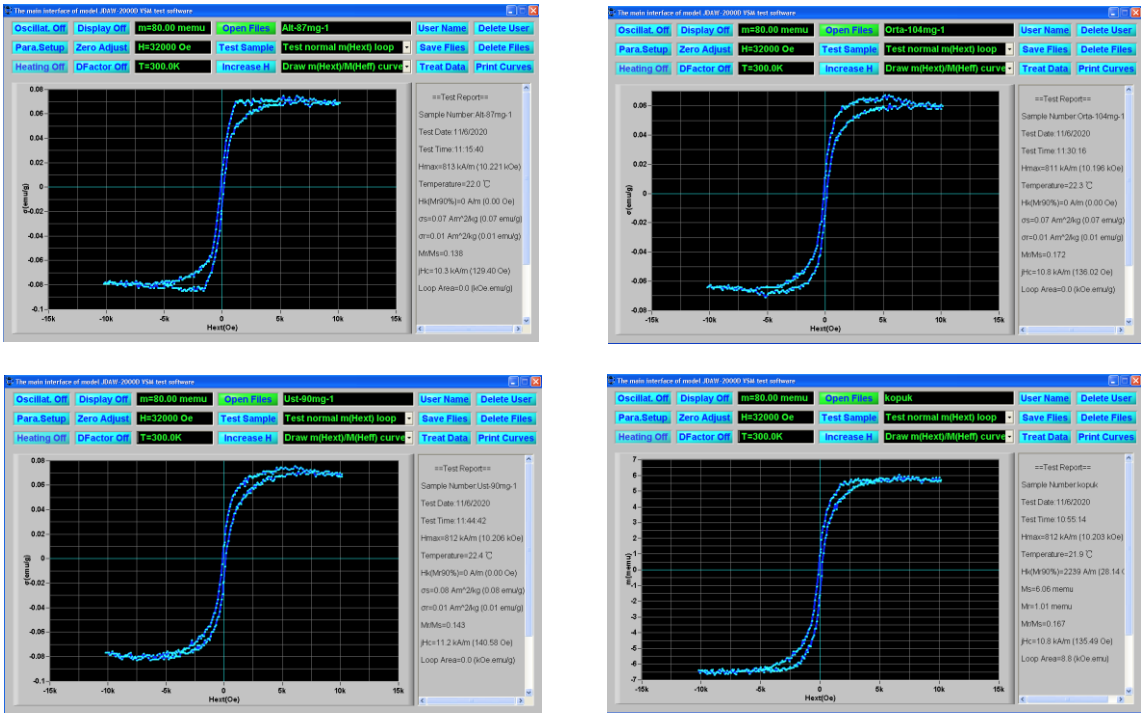


Figure 2. VSM results of Lead Oxide powder taken from the bottom (upper left), the intermediate (upper right), the top part (left bottom) and styrofoam (right bottom).

2.2.4. Particle Size Analysis Results

The particle size for the top and the bottom part of Lead Oxide from the plates also investigated by this characterization technique. All results were obtained in Central Research Laboratory in Katip Çelebi University.

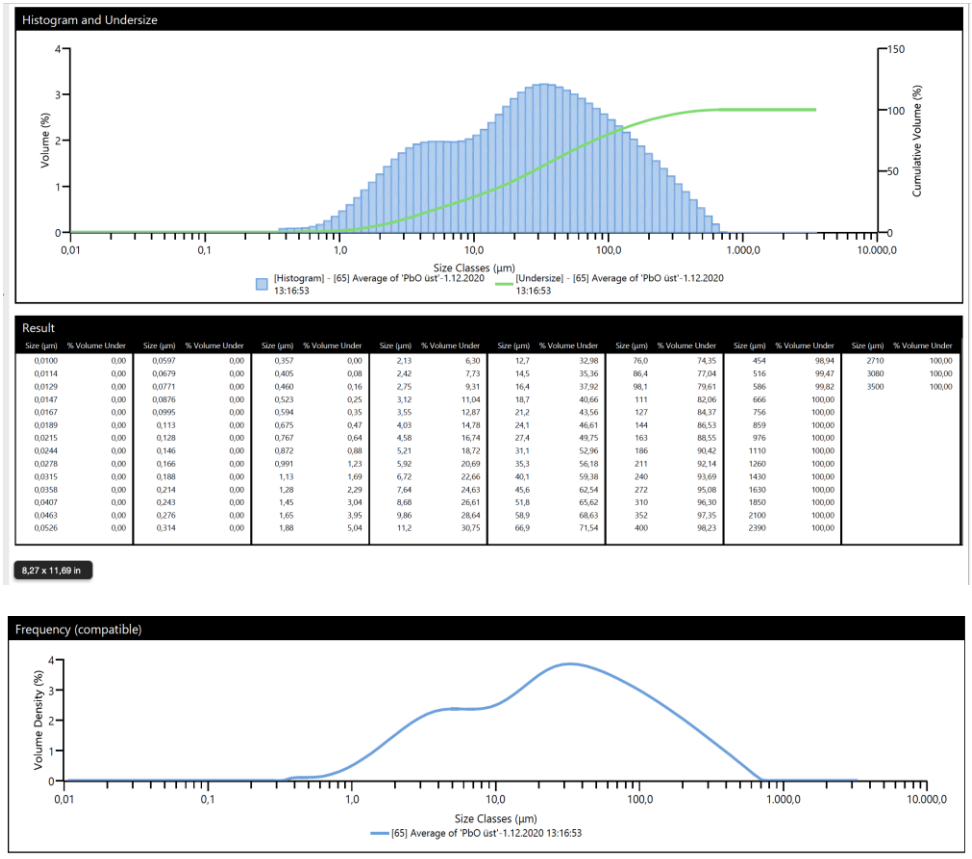
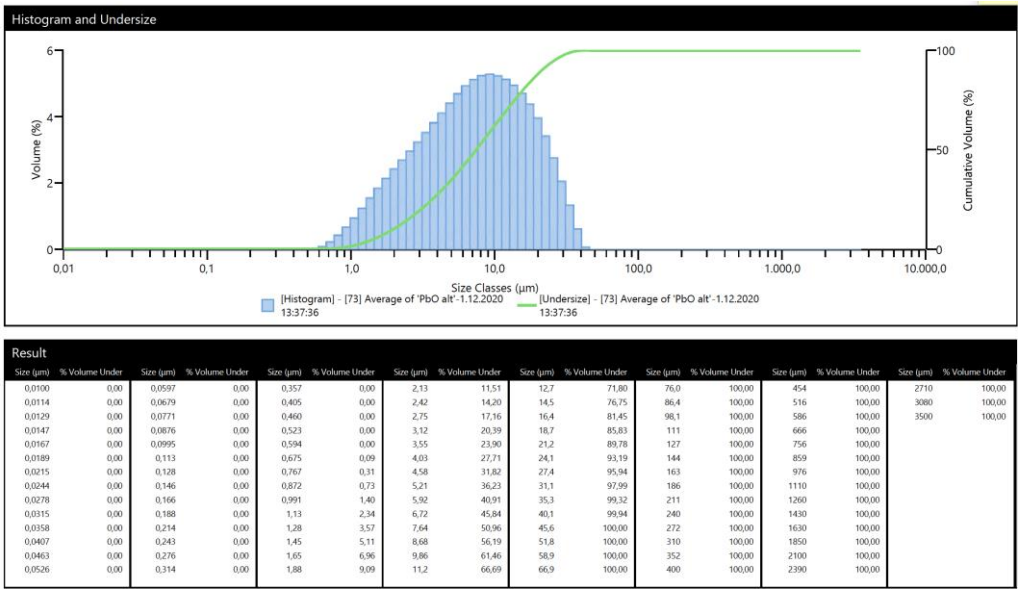


Figure 3. The partice size analysis of Lead Oxide powder taken from the top part.



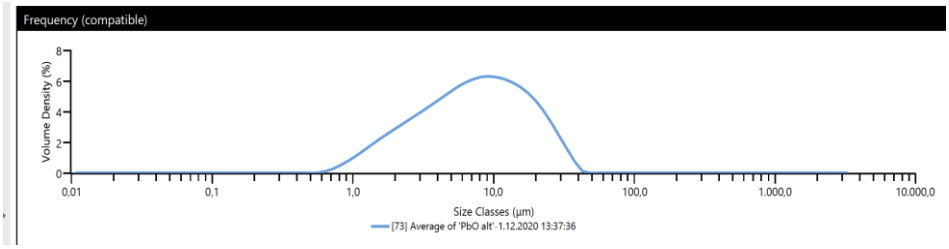


Figure 4. The partice size analysis of Lead Oxide powder taken from the bottom part.

2.2.5. X-Ray Diffraction Results

Lead sulfate, lead dioxide and lead oxide are the main components of lead paste in a used lead-acid battery. In order to make a detailed comparison, the literature source was used for XRD results of $PbSO_4$, PbO and PbO_2. [8]

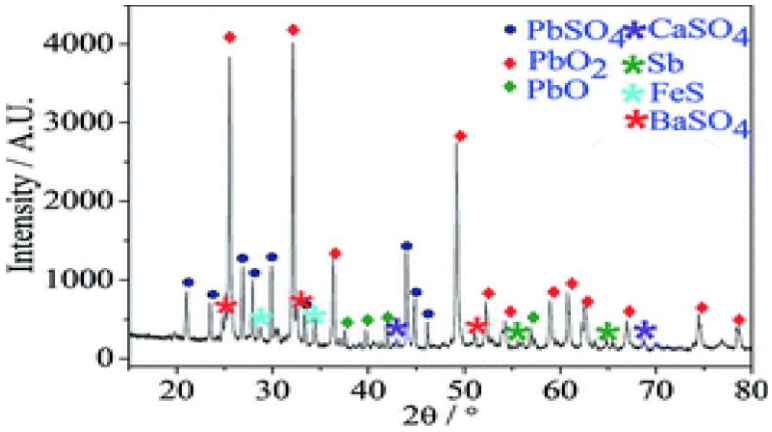


Figure 5. The literature source for XRD results of $PbSO_4$, PbO and PBO_2 [8].

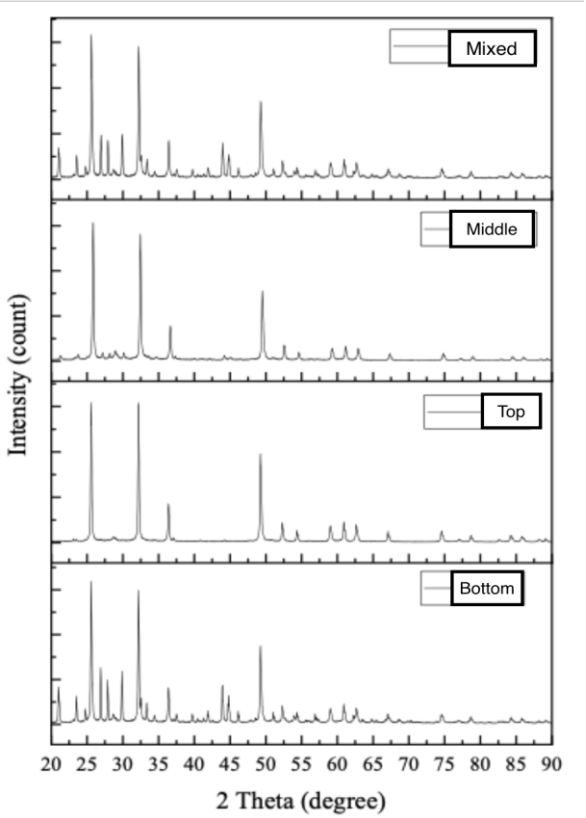


Figure 6. The XRD results of Lead Oxide powder recycled from the car battery in our project.

XRD results of the powder from recycled battery were obtained in Laboratories of Material Science and Engineering Department in İzmir Katip Çelebi University. PbO_2 peaks (25, 33 and 49 degree - in all plots), $PbSO_4$ peaks (between 25 and 30 degrees- in mixed and bottom plots) and PbO peaks (between 50 and 60 degrees, 35 and 45 degrees - in mixed and bottom plots) can be clearly seen in combined XRD plots.

3. Results and Discussion

Chemical analysis showed that there is inverse correlation between the abundance of $PbSO_4$ & PbO_2 and physical indication for $PbSO_4$ is color. All recycled powder contains 98 % Lead components + others. The percentages for $PbSO_4$: Lower part 16.7 , Intermediate part 10.3, Upper part 0.5 .

X-Ray diffraction results showed that recycled Lead Oxide powder had perfect purity. Our results matched with literature results perfectly. Since the powder will be used for RAM (radar absorbing material), the size comparison is important. SEM and Particle size analysis results indicated that the size of the recycled powder could be decreased below 10 micrometer. Optical and electronic properties of the powder can be changed by decreasing size and PbO can be a potential candidate to be a radar absorbing material.

Poor magnetization of Lead Oxide was shown in VSM analysis. Oxygen inside Lead which is a good conductor ensured the infinitesimal magnetic properties.

4. Conclusion

According to the EPA, about 80% of the lead and plastic in a lead-acid battery is recycled for reuse. Lead-acid batteries are also closed-loop recycled, which means each part of a battery is recycled into a new battery. A lead-acid battery contains lead and highly corrosive sulfuric acid electrolyte solution which are classified as hazardous wastes according to RA6969. [6] Contrary to other recycled products, lead quality is not lowered. [9] Many countries achieving recycling rates of more than 90 percent. In this work, we managed to recycle Lead Oxide from used car battery. In summary, Lead Oxide powders were successfully prepared for characterization. XRD results proved that recycled product had supreme purity.

5. Acknowledgements

The authors would like to acknowledge İNCİ GS YUASA AKÜ for their contribution in the extraction process and chemical analysis and also BAP project (2020-TDR-FEBE-0008).

References

1. R. Jolly, C. Rhin, The recycling of lead–acid batteries: production of lead and polypropylene, Resour. Conserv. Recycl. 10 (1994) 124–137.
2. Ei Liu, Beibei Ma, Yan Fu, Ke Zhang, Mohammed Adnan Mezaal, Fajun Li, Xiaoyuan Zhao, Lixu Lei,Electrochemical property of α-PbO prepared from the spent negative powders of lead acid batteries, (2017) 21:35–46.
3. Andrews, D.; Raychaudhuri, A.; Frias, C. Environmentally Sound Technologies for Recycling Secondary Lead. J. Power Sources 2000, 88, 124.
4. Ellis, T. W.; Mirza, A. H. The Refining of Secondary Lead for Use in Advanced Lead-acid Batteries. J. Power Sources 2010, 195, 4525.
5. Sonmez, M. S.; Kumar, R. V. Leaching of Waste Battery Paste Components. Part 1: Lead Citrate Synthesis From PbO and PbO2. Hydrometallurgy 2009, 95, 53.
6. I. Thornton, R. Rautiu, S. Brush, Lead—The Facts, Allan Printing, London, 2001.
7. G. De Angelis, F. Medici, M.R. Montereali, L. Pietrelli, Reuse of residues arising from lead batteries recycle: a feasibility study, Waste Manage. 22 (2002) 925–930.
8. "Qualitative lead extraction from recycled lead–acid batteries slag."Alessandra Smaniotto, Angela Antunes, Irajá do Nascimento Filho c, Luciana Dornelles Venquiaruto, Débora de Oliveira, Altemir Mossi, Marco Di Luccio, Helen Treichel, ,Rogerio Dallago (Journal of Hazardous Materials 172 (2009) 1677–1680)
9. Sutherland, Charles A.; Milner, Edward F.; Kerby, Robert C.; Teindl, Herbert; Melin, Albert; Bolt, Hermann M. "Lead". Ullmann's Encyclopedia of Industrial Chemistry. (Weinheim: Wiley-VCH. doi:10.1002/14356007.a15_193.pub2).
10. Alper Turhan, Ph.D. Thesis, "Electrochemically modified carbon fibers as an active mass additive in enhanced flooded lead acid battery" (June 2020)

Fine-tuning Deep Learning Models to Identify Pepper Leaf Diseases

Gokhan Altan[1]

[1] Department of Computer Engineering, Iskenderun Technical University, Turkey
Email: gokhan.altan@iste.edu.tr, ORCID: 0000-0001-7883-3131

Abstract

Deep Learning (DL) has become very popular, especially thanks to the towering achievements in classification, segmentation and identification in many image types by the use of transfer learning and detailed feature learning stages. Although DL has been criticized by some researchers due to down sampling procedure, it has created a new enlightenment to image processing approaches with its capabilities in generating feature activation maps. In this study, the pre-trained CNN architectures were evaluated for bell pepper leaf identification. The VGG-19 architecture with the best identification performance was fine-tuned by iterating various dropout factorization rates and fully-connected layers on different feature activation maps. The fine-tuned VGG-19 CNN architecture separated the healthy and diseased pepper leaves with an accuracy of 97.84%.

Keyword(s): Bell Pepper, Deep Learningdepht, VGG-19, Plant Leaf Diseases, PlantVillage

1. Introduction

Food production is a sector that is important in many areas, directly related to social life and economy. Planting, seed dressing, and harvesting are the main inception processes in the food industry. Especially in recent years, due to plant diseases and pandemics, the disruptions in production and the inability to make sufficient supply both impressed the national economy and prevented sufficient natural products from reaching the consumers. In this case, the necessity of preventing plant diseases and correct management of the necessary production policies has come to the forefront for the agricultural sector all over the world. Although it is difficult to predict some of these situations, plant diseases can contribute to the harvesting bodyborne leaf diseases, stem diseases, fruit diseases, and root diseases (Singh & Misra, 2017). Harvest loss caused by late detection of plant diseases and overdue in taking precautionary measures related to this loss is an application with conventional methods even in developed countries. Moreover, spraying for identified plant disease is applied to the entire harvest area. This causes unnecessary expenses for agricultural producers and even spoils the genetics of the plant due to chemical spraying. In this respect, locally and plant-based diagnosis of plant diseases is of great necessity to improve the earning in productivity. Therefore, computerized techniques for early detection of plant diseases and applying reliable preventives are the main focus of agriculture technologies.

The advancements on computer vision and the increasing computational capacities with the advantages of GPU have highlighted the artificial intelligence based applications. Moreover, common use of camera technologies even for mobile devices has contributed these standpoints on novel machine learning techniques. Use of computer vision in agriculture increases day by day with the prevalence of drone technologies. It enables the opportunity to configure plant disease identification models using real-time leaf, bacterial, fruit diseases using

hybrid techniques and robotic. in the light of these developments, computer-aided techniques with automated and hand-crafting have frequently been implemented to identify different plant leaf diseases. Jagan and Mohan handled the paddy spot disease using SIFT technique and conventional machine learning algorithms, including k-nearest neighbor and support vector machine (SVM) (Jagan et al., 2016). Phadikar focused on the rice leaf diseases feeding morphologic and histogram features into SVM with non-linear kernels (Phadikar, 2012). Islam et al extracted RGB-based features on spot areas to identify the rice leaf disease using Naive Bayes (Islam et al., 2018). Kumari et al. utilized contrast and statistical features of spots for tomato and cotton leaf diseases for unfolding to an artificial neural network (ANN) (Kumari et al., 2019). Chouhan et al. classified the similarity based correlation features for various plant species using ANN (Chouhan et al., 2018).

Deep Learning (DL) algorithms are popular hybrid techniques with feature learning and classification stages. Moreover, it is an easy-adaptable technique for many image types with high classification and segmentation performances on large-scale datasets using transfer learning on DL with pre-trained architectures. Hence, convolutional neural networks (CNN) algorithm is the first choice for many researchers without pre-processing and advanced image processing knowledge. Sladojevic et al. utilized CNN for various plant leaf diseases on their own large-scale database (Sladojevic et al., 2016). The studies applied transfer learning on pre-trained ImageNet weights using AlexNet (Amara et al., 2017; Brahimi et al., 2017; Chouhan et al., 2018; Kumar et al., 2020; Lee et al., 2017; Liu et al., 2018; Sladojevic et al., 2016), LeNet (Kumar et al., 2020), VGGNet (Amara et al., 2017; Lee et al., 2017), GoogleNet (Amara et al., 2017; Lee et al., 2017; Sladojevic et al., 2016), ResNet (Lee et al., 2017), and more. Lee et al. utilized feature learning on AlexNet using fine-tuning and presented high identification performances of 40 leaves (Lee et al., 2017). Amara et al. applied CNN with LeNet architecture to classify banana leaf disease on PlantVillage dataset with high accuracy rates (Amara et al., 2017). The researchers evaluated the performance of various pre-trained architectures on 9 tomato leaf diseases from PlantVillage dataset and apple leaf diseases on their own dataset (Lee et al., 2017). Ferentinos et al. utilized many pre-trained CNN architectures for separation of many leaf diseases on PlantVillage database (Ferentinos, 2018). Geetharamani and Pandian fine-tuned a novel CNN architecture with variations of classification parameters including, dropout rate, learning rate, and dense layers (Geetharamani & J., 2019). Kurup et al. fed the Capsule Networks and CNN architectures with plant leaves from PlantVillage to compare the identification performances (Kurup et al., 2020). Altan also analyzed the efficiency of Capsule Networks over plant leaf diseases (Altan, 2020). This paper aims at classifying bell pepper leaf diseases using pre-trained CNN architectures (VGG-16, VGG-19, and MobileNet), comparing the performance of fine-tuning on the best architecture for transferring dominant pixel information among convolutional blocks.

The remaining of the paper is organized as a general description of focused plant species in the PlantVillage database and architectural specifications of experimented CNN models in materials and methods. The experimental setup and performance evaluation metrics for the experimented architectures in experimental results. A complete comparison with state-of-the-art in terms of plant leaf disease identification performances and superiorities of the CNN models in discussion.

2. Materials and Methods

2.1. PlantVillage Database

PlantVillage Database is a large-scale plant leaf disease dataset. It enables enhancing the harvesting techniques with computer-assisted techniques to provide the productivity in harvesting as early diagnosis. Whereas the food industry is a main problem to supply the nutrition necessity for people all over the world, it is possible to qualify profitability in time dressing and watering. Therefore, it was presented as a challenge dataset by Land Grant University, USA (Hughes & Salathe, 2015). PlantVillage consists of a total number of 54K leaf images from thirteen species. Most of the plant species have disease spots for various leaf diseases and healthy leaves to develop diagnosis models for both binary and multi-classes.

Figure 1. Sample bell pepper leaves for healthy (left) and bacteria spot disease (right)

In this study, the bell pepper leaf bacteria disease is focused on with a binary classification. The PlantVillage database has a total number of 2.4K bell pepper leaf images including, 1K bacteria diseased and 1.4K healthy leaves. Fig. 1 indicates the randomly selected pepper leaf images for healthy and diseased. Each leaf image has a standardized square dimension with a standard background, but uncertain rotations with different angles. Hence, no cropping and background removal were applied to the inputs before feeding the CNN architectures.

Although several images have a shadow effect; none of the bell pepper images was excluded in the analysis. We augmented each leaf image with horizontal-, vertical-, and both-flipping before splitting training and testing folds.

2.2. Convolutional Neural Networks (CNN)

Deep Learning is an advancing hybrid technique with outstanding classification performances on various types of inputs. The main philosophy of Deep Learning is combining unsupervised algorithms in the feature learning and detailed supervised algorithms with novel factorization and regularization techniques. On the other hand, it has advantages of transferring the knowledge between related problems for reducing the training time. Transfer learning enables optimizing pre-trained architectures with a detailed convolutional integration on the adjacent convolution blocks. Moreover, transfer learning with popular pre-trained CNN

architectures provides a feature map extraction on a detailed convolutional layer. Hence, pre-trained CNNs enable reaching better globalization using regularization and factorization on large-scale datasets.

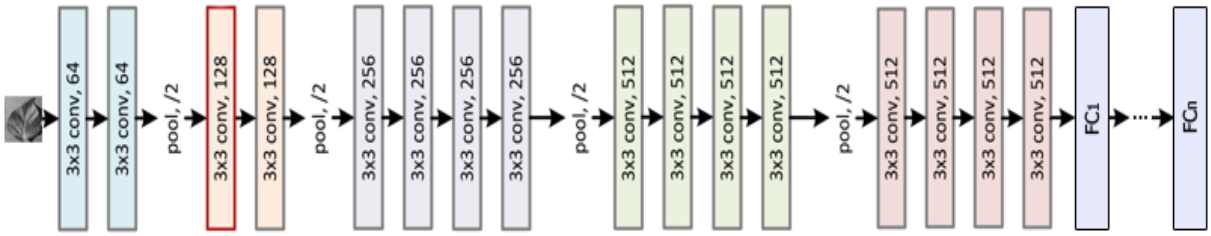


Figure 2. VGG-19 architecture and fine-tuning process in dense layers

The main benefits of CNN are feature learning, transfer learning, fine-tuning capability on pre-trained architectures, mobile architectures, and optimizing deep neural networks with high generalizations. The feature activation maps as a result of many convolution blocks with a pooling layer are fed into the fully-connected layers (FC) just as a multilayer perceptron. Whereas the primary convolution layers define low-level features, the subsequent layers generate high-level features by transferring the dominant pooling results (Altan, 2019). I evaluated the performance of transfer learning using popular pre-trained architectures, including VGG-16, VGG-19 (Simonyan & Zisserman, 2014), and MobileNet (Howard et al., 2017) on the identification of plant leaf diseases.

3. Experimental Results

The CNN has the ability to compose various combinations during the modeling process due to the sequence of convolutional layers, max pooling, and fully-connected layers. It is a challenging procedure for proposing novel CNN architectures. However, the popular pre-trained architectures have proven their usability and efficacy on different problems. Hence, transfer learning was chosen for the analysis of plant leaf diseases. VGG-16 and VGG-19 architectures were named according to their layer depth by Visual Geometry Group. Although VGG architectures have a small number of feature learning layers, it is not a mobile network due to the depth of the fully-connected layers. MobileNet is one of the most lightweight CNN architectures with the capabilities of adaptability to many data types in artificial intelligence. Even though very deep neural networks with complicated feature learning stages have many advantages on extracting detailed features, it is more important to reach approximate classification performances using simplistic architectures with fine-tuning on CNN architectures. That is why this study focuses on fine-tuning and pruning on simple CNN architectures.

Many plant leaf images have uncertain rotations, shadow views and sun shining. Nonetheless no image enhancing techniques such as background removal and histogram equalization were performed before CNN architectures.

The diseased leaf images have commonly small bacteria spots. Therefore, it is possible to diverging the loss of significance for disease separation on complex CNN architectures. It is an important point using shallow feature learning stages with low dimension of feature activations maps. Moreover, this point enables evaluating the success of CNN architectures for use as an

early diagnostic. Therefore, the best CNN architecture (among VGG-16, VGG-19, and MobileNet) was fine-tuned using various depths of fully-connected layers and neuron ranges in supervised learning stages. Thus it is aimed to reach the optimum model with highest generalization performance on leaf disease classification. Hereby, the best achievements were presented for each CNN architecture and compared with the state-of-art.

Each plant leaf image was converted to gray-scale image to obtain a quick analysis on one channel instead of three channels (Red, Green, and Blue). Whereas it is a standardized input size for VGG-16 and VGG-19, MobileNet supports all input sizes bigger than 32x32. The leaf images were resized to 224×224 for both CNN architectures. Each leaf image was augmented to enhance the dataset for supporting the pre-trained model with different angles and to avoid overfitting. Each image was increased by 4 times with vertical-, horizontal, and both-flipping. Hence, the training of the pre-trained CNN architectures for identification of plant leaf diseases were performed on a total number of 9900 plant leaf images.

The proposals were trained using 80% and 20% folds for train and test, respectively. The separate folds of the plant leaf images were utilized for training and testing the pre-trained CNN architectures. The performance of the proposals was evaluated using statistical test characteristics, including overall accuracy, sensitivity, specificity, and precision (Altan, 2019).

In the fully-connected layers of the CNN architectures were iterated for various dropout factorization rates (0.25, 0.5, 0.75) in the fine-tuning, a range of fully-connected layers (2 and 3), and neuron depth (256, 512, 1024, 2048, and 4096 neurons). The output functions of the last fully-connected layer is softmax.

Table 1. The classification performances (%) for the pre-trained CNN architectures

CNN architectures	Accuracy	Sensitivity	Specificity
VGG-16	82.25	77.17	85.72
VGG-19	91.12	90.35	91.64
MobileNet	90.27	89.51	90.80

The highest classification performances in terms of accuracy, sensitivity, and specificity are presented in Table 1. The experimental results indicate that the VGG-19 has the highest identification capability on the plant leaf diseases among the analyzed pre-trained architectures. Whereas MobileNet reached well enough classification performances, VGG-16 performed a miscarriage generalization on the re-training of the architectures on ImageNet weights. Therefore, VGG-19 architecture was fine-tuned using various fully-connected layer specifications. Table 2 presents the best five classification performances on fine-tuning of VGG-19 architecture.

Table 2. The best five classification performances (%) for fine-tuning of VGG-19 architecture with various fully-connected layers

Fully-connected layers	Dropout	Accuracy	Sensitivity	Specificity	Precision
FC1(512) **FC2(512)**	0.50	94.93	94.04	95.53	93.45
FC1(512) **FC2(1024)** **FC3(2048)**	0.25	94.95	94.11	95.52	93.44
FC1(1024) **FC2(1024)**	0.50	95.19	93.41	96.40	94.62
FC1(512) **FC2(512)** **FC3(2048)**	0.75	96.05	94.89	96.84	95.31
FC1(1024) **FC2(512)** **FC3(2048)**	0.50	**97.84**	97.48	98.09	97.19

The enhanced neuron architects in each fully-connected layer provide high identification accuracy using deeper models. It is experimented that using a big number of neurons for both fully-connected layers (4096 neurons) performed inefficacious generalization under the average classification performances for iterated range of dropout and training parameters. Using 1024 neurons at 1st dense layer, 512 neurons at 2nd dense layer, and 2048 neurons at 3rd dense layer with the proposed fine-tuning approach on VGG-19 architecture reached the best generalization using a dropout factorization rate of 0.5. The highest bell pepper leaf disease diagnosis performance was achieved with the rates of 97.84%, 97.48%, 97.49%, and 97.19% for accuracy, sensitivity, specificity, and precision, respectively.

4. Discussion

The pre-trained architectures are commonly utilized to analyze leaf, fruit, and root diseases for various plants. The easy-adaptability and denomination of CNN architectures are supported by transfer learning and pre-trained weights. Moreover, it enables fast convergence and robust generalizations for many related problems. Although novel architectures using capsule networks reported the advantages of using spatial information on direct images by excluding pooling layers, CNN architectures are still in the first position for consecutive convolutions to extract significant feature maps.

To the best of my knowledge, whereas a limited number of researchers focused directly on bell pepper leaf disease identification, most of them analyzed it among various plant species. However, these papers reported an overall identification accuracy rather than plant-based classification performances. The achievements in the related works on the PlantVillage database are presented in Table 3.

The achievements are obtained using the contingency table and number of the plant species for pepper leaves. Geetharamani and Arun Pandian highlighted a light-weight CNN model. It has nine layers including fully-connected layers and convolutional layers to identify the plant leaf diseases. They iterated with various batch size, epoch, and dropout factorization

index in the training of CNN. They reported the identification capacity of their proposal over AlexNet, VGGNet, Inception-v3 and ResNet. Their proposal achieved the highest pepper leaf disease classification rates of 93.00%, 92.00%, and 93.00% for accuracy, sensitivity, and specificity, respectively (Geetharamani & J., 2019). Kurup et al. proposed a capsule network on the PlantVillage dataset. Their proposal achieved pepper leaf classification accuracy rates among 91% and 96% (Kurup et al., 2020). Altan also applied a capsule network architecture on the augmented pepper leaf images. He reached classification performance rates of 95.76%, 96.37%, and 97.49% for accuracy, sensitivity, and specificity, respectively (Altan, 2020).

Although the literature reported well-enough classification performances with popular CNN architectures and novel capsule network architectures, nevertheless it is possible to reach better generalization capabilities by fine-tuning the architectures according to the problems. This study explores the effect of fine-tuning on well-known CNN architectures on pepper leaf disease identification using various fully-connected layer sequences and dropout factorization rates.

Table 3. The state-of-art for pepper leaf disease identification on PlantVillage

Related works	Classification	Accuracy	Sensitivity	Specificity
Geetharamani et al. (2019)	Own CNN architecture	93.00	92.00	93.00
Kurup et al.(2020)	Capsule Network	-	89.17	-
Altan (2020)	Capsule Network	95.76	96.37	97.49
This study	VGG-19 + Fine-tuning	97.84	97.48	98.09

The fine-tuned VGG-19 architecture overcomes the state-of-the-art using three fully-connected layers and 0.5 of dropout factorization. VGG-19 is the most responsible CNN architecture for the experimented ranges of classification parameters under the architecture-based evaluation metrics. Using a small number of neurons at the first fully-connected layer and detailing the previous fully-connected layers with a big number of neurons provided a better optimization in both training and testing of the proposal. The fine-tuned VGG-19 architecture achieved identification rates of 97.84%, 97.48%, and 98.09% for accuracy, sensitivity, and specificity, respectively.

References

1. Altan, G. (2019). DeepGraphNet: Grafiklerin Sınıflandırılmasında Derin Öğrenme Modelleri. European Journal of Science and Technology, Special Issue, 319–327. https://doi.org/10.31590/ejosat.638256
2. Altan, G. (2020). Performance Evaluation of Capsule Networks for Classification of Plant Leaf Diseases. International Journal of Applied Mathematics Electronics and Computers. https://doi.org/10.18100/ijamec.797392
3. Amara, J., Bouaziz, B., & Algergawy, A. (2017). A deep learning-based approach for banana leaf diseases classification. Lecture Notes in Informatics (LNI), Proceedings - Series of the Gesellschaft Fur Informatik (GI).
4. Brahimi, M., Boukhalfa, K., & Moussaoui, A. (2017). Deep Learning for Tomato Diseases: Classification and Symptoms Visualization. Applied Artificial Intelligence. https://doi.org/10.1080/08839514.2017.1315516
5. Chouhan, S. S., Kaul, A., Singh, U. P., & Jain, S. (2018). Bacterial foraging optimization based radial basis function neural network (BRBFNN) for identification and classification of plant leaf diseases: An automatic approach towards plant pathology. IEEE Access. https://doi.org/10.1109/ACCESS.2018.2800685
6. Ferentinos, K. P. (2018). Deep learning models for plant disease detection and diagnosis. Computers and Electronics in Agriculture. https://doi.org/10.1016/j.compag.2018.01.009
7. Geetharamani, G., & J., A. P. (2019). Identification of plant leaf diseases using a nine-layer deep convolutional

neural network. Computers and Electrical Engineering. https://doi.org/10.1016/j.compeleceng.2019.04.011
8. Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., & Adam, H. (2017). MobileNets: Efficient convolutional neural networks for mobile vision applications. In arXiv.
9. Hughes, D. P., & Salathe, M. (2015). An open access repository of images on plant health to enable the development of mobile disease diagnostics.
10. Islam, T., Sah, M., Baral, S., & Roychoudhury, R. (2018). A Faster Technique on Rice Disease Detectionusing Image Processing of Affected Area in Agro-Field. Proceedings of the International Conference on Inventive Communication and Computational Technologies, ICICCT 2018. https://doi.org/10.1109/ICICCT.2018.8473322
11. Jagan, K., Balasubramanian, M., & Palanivel, S. (2016). Detection and Recognition of Diseases from Paddy Plant Leaf Images. International Journal of Computer Applications. https://doi.org/10.5120/ijca2016910505
12. Kumar, S., Sharma, B., Sharma, V. K., Sharma, H., & Bansal, J. C. (2020). Plant leaf disease identification using exponential spider monkey optimization. Sustainable Computing: Informatics and Systems. https://doi.org/10.1016/j.suscom.2018.10.004
13. Kumari, C. U., Jeevan Prasad, S., & Mounika, G. (2019). Leaf Disease Detection: Feature Extraction with K-means clustering and Classification with ANN. 2019 3rd International Conference on Computing Methodologies and Communication (ICCMC), 1095–1098. https://doi.org/10.1109/ICCMC.2019.8819750
14. Kurup, R. V., Anupama, M. A., Vinayakumar, R., Sowmya, V., & Soman, K. P. (2020). Capsule network for plant disease and plant species classification. Advances in Intelligent Systems and Computing. https://doi.org/10.1007/978-3-030-37218-7_47
15. Lee, S. H., Chan, C. S., Mayo, S. J., & Remagnino, P. (2017). How deep learning extracts and learns leaf features for plant classification. Pattern Recognition. https://doi.org/10.1016/j.patcog.2017.05.015
16. Liu, B., Zhang, Y., He, D. J., & Li, Y. (2018). Identification of apple leaf diseases based on deep convolutional neural networks. Symmetry. https://doi.org/10.3390/sym10010011
17. Phadikar, S. (2012). Classification of Rice Leaf Diseases Based on Morphological Changes. International Journal of Information and Electronics Engineering. https://doi.org/10.7763/ijiee.2012.v2.137
18. Simonyan, K., & Zisserman, A. (2014). Very Deep Convolutional Networks for Large-Scale Image Recognition. 3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings. http://arxiv.org/abs/1409.1556
19. Singh, V., & Misra, A. K. (2017). Detection of plant leaf diseases using image segmentation and soft computing techniques. Information Processing in Agriculture. https://doi.org/10.1016/j.inpa.2016.10.005
20. Sladojevic, S., Arsenovic, M., Anderla, A., Culibrk, D., & Stefanovic, D. (2016). Deep Neural Networks Based Recognition of Plant Diseases by Leaf Image Classification. Computational Intelligence and Neuroscience. https://doi.org/10.1155/2016/3289801

Interpolation-based smart video stabilization

Semiha Dervişoğlu[1], Mehmet Sarıgül[2], Levent Karacan[3]

[1,2,3] Department of Computer Engineering, Iskenderun Technical University, Turkey
Emails: semihadervisoglu.mfbe19@iste.edu.tr, mehmet.sarigul@iste.edu.tr, levent.karacan@iste.edu.tr

Abstract

Video stabilization is the process of eliminating unwanted camera movements and shaking in a recorded video. Recently, learning-based video stabilization methods have become very popular. Supervised learning-based approaches need labeled data. For the video stabilization problem, recording both stable and unstable versions of the same video is quite troublesome and requires special hardware. In order to overcome this situation, learning-based interpolation methods that do not need such data have been proposed. In this paper, we review recent learning-based interpolation methods for video stabilization and discuss the shortcomings and potential improvements of them.

Keyword(s): Video stabilization, deep learning, unsupervised learning, interpolation methods

1. Introduction

Videos, which have become a habit of our daily life, are used in many fields such as military (unmanned aerial vehicles), education (scientific research, etc.), healthcare (in order to determine the size and location of the problem in endoscopy and colonoscopy videos) and film industry (in movie shootings, etc.). A video is defined as a set of consecutive images taken with a handheld camera or a camera positioned on a vehicle. The shaking of the hand or the shaking of the vehicle in these videos can cause visual problems. Although some hardware has been produced to solve this problem, this equipment can be expensive. For this reason, the interest in the field of digital video stabilization has increased.

Digital video stabilization can be applied in three steps. Motion path estimation, motion path smoothing and producing the stable video (Figure 1). In the majority of the digital video stabilization studies, motion is detected by using the similarity and feature extraction between the frames, camera path is corrected and video is stabilized. Pixel and block mapping-based motion detection methods use inter-frame distance metric and similarity measurement to estimate the motion of each pixel and block. In the pixel-based approach a pixel is represented with 3 color values and an invariant brightness value. It is hard to estimate the pixel movement between the consecutive frames due to the similarity of neighboring pixels. A pixel can match with several pixels in the next frame. To overcome this problem a holistic approach is applied. A 2-dimensional transform is applied to one of the frames as a whole and the number of matching pixels is tried to be maximized. Block-based mappings reduce the negative effects of pixel-based methods and block-pixels are used for displacements between two scenes. It also only considers blocks at a certain distance from the block to be matched to avoid over-matching. On the other hand, the feature-matching methods find the easily recognizable points in the scene. Only the displacements of the relevant points are calculated. All videos are processed frame-by-frame, and the positions of these relevant points are determined by tracking the

selected features in the video frames, and the motion trajectories are determined. There are several feature extraction algorithms such as SIFT, SURF and FAST. SIFT does direction-independent feature extraction (Chiu et al 2013, Rublee et al. 2011, Lowe, G. 2004), SURF approach is an optimized version of the SIFT algorithm. (Rublee et al. 2011, Bay et al. 2006), and FAST is a corner detection algorithm reported to give more successful results than alternatives (Chiu et al. 2013). Although feature matching methods are successful, they have limitations. Feature extraction algorithms may not be able to extract enough features in certain parts of the scene and this may lead to poor results in stabilization. This is one of the reasons learning-based methods become popular in video stabilization. Supervised learning-based methods require stable and unstable pairs for each video in the learning set. Although some datasets are suggested on this subject, the quality of these studies is limited as it is dependent on the training dataset (DeepStab). On the other hand, self-supervised or semi-supervised methods may reach a better generalization quality (Choi et al. 2019; Xu et al. 2020; Lee et al. 2021). Recently, semi-supervised or unsupervised methods based on interpolation are frequently encountered in the literature.

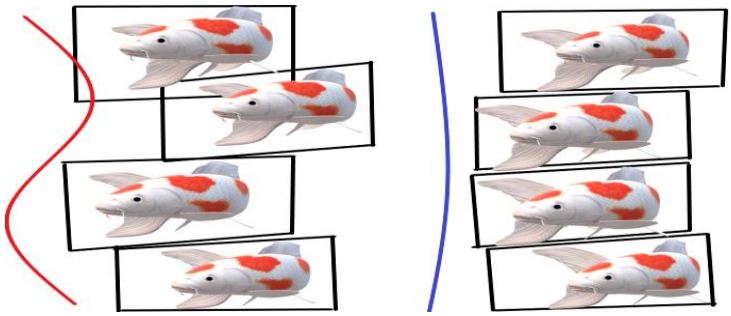


Figure 1. Principle of video stabilization

2. Development of Interpolation-Based Video Frame Interpolations

Interpolation is a technique that creates pixels instead of unknown pixels in image amplifications. There are widely used image interpolation techniques. The first of these methods is the Nearest-Neighbor interpolation. In this method, the distances between the nearest K neighbors and the unknown p pixel are found and the closest value is set to p. The second is bilinear interpolation. Four pixels are indexed as A, B, C, D according to the distance to unknown pixel, p. Using these 4 close points, the value of the unknown pixel is adjusted. Third, Bicubic interpolation is calculated in the similar way to bilinear interpolation. However, it is calculated by evaluating 16 close pixels instead of 4. In this way, it spreads the effect with more pixels. Another method, Cubic B-SPline (Eq.1) makes the interpolation curve smoother and creates better image edges. (Han, D. 2013, March).

$$f(x) = \sum_{k=-\infty}^{\infty} \quad B_{k,n+1}(x).f(x_k) \qquad \text{(Eq.1)}$$

Mahajan et al. proposed a path-based interpolation method. This method preserves the operating frequency content, which takes two input images and generates an intermediate frame, and produces a natural-looking animation sequence. If this sequence is considered as a single scan line among the input frames A and B (shifted form of A), the pixels are copied from

A initially and the image starts to be copied from B at some point. This is the interpolation path for unknown pixel p. In this way, a smooth image is obtained. (D. Mahajan et al. 2009).

Classical video frame interpolation algorithms include optical flow and frame interpolation. In these studies, the quality of frame interpolation is largely dependent on optical flow. Werlberger et al obtained natural looking intermediate frames (t+1/2) between two input (t, t+1) frames by propagating the optical flow using linear interpolation. This study provides advantages such as finding lost frames and partially correcting the distorted frame. However, optical flow cannot be found in the restored parts. Interpolation techniques are used for the restored parts of the image (Werlberger, et al. 2011, July). Similar work based on optical flow was performed in 3D (Yu,et al 2013). There are also different approaches without optical-flow. The phase-based method for estimating the motion is the alternative method which uses Euler as a phase-based approach for interpolation. This method reduces the limitations of motion analysis that could be interpolated. It was successful at frame interpolation and retiming of high-resolution high frame rate video. (Meyer et al. 2015).

There are also studies combining the video frame interpolation steps of motion estimation and pixel synthesis into a single local convolutional network. Pixel interpolation is performed by convolution on pixel patches instead of optical flow formulation. Its advantages are to handle occlusion, blur and sudden brightness change and it provides high frame interpolation. This method is more flexible than optical flow methods. However, it is not effective for large motions (Niklaus et al. 2017a). Another study used large kernels to handle large motions. Due to this memory demand, it limits the number of pixels the kernels can predict in a frame. To solve this problem, the model has formulated the frame interpolation as local separable convolution on input frames using synchronous 1D kernel pairs. (Niklaus et al. 2017b). Classical optical flow-based methods fail when flow estimation is difficult. Recent neural network-based methods handling pixel values as hallucinations produce blurry solutions. Deep Voxel flow combines two methods. In this work, a network that learns to synthesize video by streaming pixel values from existing ones is trained. It showed that the Deep Voxel flow approach improves the latest CNN techniques to interpolate and extrapolate both optical flow and video. This method is still limited in adapting to inaccuracies in the motion/voxel flow estimation. (Liu, et al. 2017). Niklaus et al proposed another method solving the occlusion problem of video frame interpolation by estimating and using bidirectional flow. This flow is used to warp and blend the input frames. This approach warps not only the input frames but also the pixel-based contextual information. It used a pre-trained neural network to extract contextual information specifically. Finally, the difference from other methods is that pre-warped frames and context maps are fed into a video frame synthesis network to generate the interpolated frame in a context sensitive manner (Niklaus et al. 2018).

3. CNN-Based Development of Video Stabilization

In order to achieve more efficient and consistent video stabilization, learning-based studies have been suggested, recently. Wang created the stable/unstable dataset called DeepStab and proposed a solution with a Siamese convolutional neural network called Stabnet (Wang et al. 2018). With a new formulation of the stabilization problem, instead of estimating and correcting a virtual camera path, it produced a stable output in an online fashion that gradually learned the transformation parameters along the timeline for each unstable scene. This approach

has been a breaking point in the deep learning approach for video stabilization. However, the 2D homomorphic transformations used in this study may fail on sequential images with heavy movements. Another study proposed a new online deep learning framework to learn the stabilization transformation for each given unstable frame and the generated stable frames, instead of classical stabilization. (Xu et al. 2018). The network consisted of a generative network with spatial transformer networks embedded in different layers. A stable frame was created for the incoming unstable frame by calculating an appropriate affine transformation. It has limits for severe tremors. Yu et al. (2019) eliminated the training step by learning enough 2D models with a convolutional network (CNN) used only as a regression tool. However, since CNN is a computationally difficult approach. Choi et al. (2020) proposed an unsupervised approach focusing on image interpolation, which has a significant effect on video stabilization to avoid the cropping effect. This method takes two consecutive frames of the video, performs a one-sixth transformation over one of these frames, and a network is trained to predict this transformation. With the help of this network, they proposed an algorithm that takes two frames and creates an intermediate-frame by iterative interpolation method and performs the stabilization process in this way. However, it causes blurring while stabilizing the process. (Guilluy et al. 2021). Myungsub Choi et al. (2020) In the CAIN approach are used channel attention end-to-end splice video frame interpolation network. They proposed a simple and low computational method without using the optical flow method which reduces the effects of large movements involving occlusion. Instead of optical flow, PixelShuffle (Shi et al. 2016) was used with channel attention. PixelShuffle progressively distributes motion-related information across channels. Creates the transformed feature map to capture variations between linked frames combined with a channel attention to capture motion. In this way it generalizes movements that are not visible. It provided video frame interpolation, which synthesizes high-quality images without motion estimation. Ali et al. (2020) used DIFRINT motivation instead of optical flow to detect local and global movements in video. Perspective inconsistency in video pairs in the DeepStap (Wang et al.) dataset has been detected. For this, training video pairs with similar perspectives but different movements were created. With the help of this dataset, motion-blind video has been tried to be stabilized. It offers unsupervised and expandable video frame interpolation-based strategy to produce iso-perspective videos. The first motion blind deep video stabilization network is presented with the help of iso perspective dataset.

4. Conclusion /Inferences Obtained

Interpolation-based stabilization methods stabilize between frames, reducing inter-frame shake without cropping and distortion. However, the quality of the images created due to the iterative approach may decrease as the number of iterations increases. The most important advantage of these methods is that the scenes can be trained end-to-end without supervision.

In situations where flow estimation is difficult, traditional optical flow-based methods fail, while recent neural network methods that view pixel values as hallucinations can produce blurry solutions. Some recent interpolation approaches have used deep network structure, which learns to synthesize video scenes by streaming pixel values from existing fields of neighboring frames. Although these methods are similar to the optical flow method, they are used as an intermediate layer, therefore, their accuracy does not directly affect the result and these methods do not require surveillance like optical flow. However, 2D transformations may also create blur in the images.

Existing methods, handling 2D information, create gaps in distant and near objects. For this reason, perspective can be insufficient. Newer methods processing 3D information promise more efficient results in video stabilization.

As a result, there are shortcomings in video stabilization, such as occlusion, inaccurate estimation of the motion, and poor frame synthesis. For this reason, video stabilization has been up to date for almost 20 years.. The fact that the problems in the video stabilization mentioned above have not been fully resolved yet is proof that the method will be popular in the near future.

5. Acknowledgments

This study was supported by project number 120E447 from the The Scientic and Technological Research Council of Turkey (TUBITAK).

References

1. Han, D. (2013, March). Comparison of commonly used image interpolation methods. In Proceedings of the 2nd International Conference on Computer Science and Electronics Engineering (ICCSEE 2013) (Vol. 10). (2), 313-318.
2. Mahajan, D., Huang, F. C., Matusik, W., Ramamoorthi, R., & Belhumeur, P. (2009). Moving gradients: a path-based method for plausible image interpolation. ACM Transactions on Graphics (TOG), 28(3), 1-11.
3. Werlberger, M., Pock, T., Unger, M., & Bischof, H. (2011, July). Optical flow guided TV-L 1 video interpolation and restoration. In International Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition (pp. 273-286). Springer, Berlin, Heidelberg.
4. Yu, Z., Li, H., Wang, Z., Hu, Z., & Chen, C. W. (2013). Multi-level video frame interpolation: Exploiting the interaction among different levels. IEEE Transactions on Circuits and Systems for Video Technology, 23(7), 1235-1248.
5. Meyer, S., Wang, O., Zimmer, H., Grosse, M., & Sorkine-Hornung, A. (2015). Phase-based frame interpolation for video. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1410-1418).
6. Niklaus, S., Mai, L., & Liu, F. (2017). Video frame interpolation via adaptive convolution. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 670-679).
7. Niklaus, S., Mai, L., & Liu, F. (2017). Video frame interpolation via adaptive separable convolution. In Proceedings of the IEEE International Conference on Computer Vision (pp. 261-270).
8. Liu, Z., Yeh, R. A., Tang, X., Liu, Y., & Agarwala, A. (2017). Video frame synthesis using deep voxel flow. In Proceedings of the IEEE International Conference on Computer Vision (pp. 4463-4471).
9. Niklaus, S., & Liu, F. (2018). Context-aware synthesis for video frame interpolation. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1701-1710).
10. Choi, J., & Kweon, I. S. (2020). Deep iterative frame interpolation for full-frame video stabilization. ACM Transactions on Graphics (TOG), 39(1), 1-9.
11. Sarıgül, M., & KARACAN, L. Stabil ve stabil olmayan videoların derin evrişimli ağlarla sınıflandırılması Classifying stable and unstable videos with deep convolutional networks.
12. Guilluy, W., Oudre, L., & Beghdadi, A. (2020). Video stabilization: Overview, challenges and perspectives. Signal Processing: Image Communication, 116015.
13. Hore, A., & Ziou, D. (2010, August). Image quality metrics: PSNR vs. SSIM. In 2010 20th international conference on pattern recognition (pp. 2366-2369). IEEE.
14. Wang, M., Yang, G. Y., Lin, J. K., Zhang, S. H., Shamir, A., Lu, S. P., & Hu, S. M. (2018). Deep online video stabilization with multi-grid warping transformation learning. IEEE Transactions on Image Processing, 28(5), 2283-2292..
15. 991 IEEE Transactions on Image Processing, 28(5):2283–2292,.
16. Xu, S. Z., Hu, J., Wang, M., Mu, T. J., & Hu, S. M. (2018, October). Deep video stabilization using adversarial networks. In Computer Graphics Forum (Vol. 37, No. 7, pp. 267-276).

17. Jiyang Yu and Ravi Ramamoorthi.(2019) Robust video stabilization by optimization in CNN weight.
18. In IEEE Conference on Computer Vision and Pattern Recognition, CVPR, Long Beach, CA, USA, June 16-20, 2019, pages 3800–3808.
19. Yu, J., & Ramamoorthi, R. (2020). Learning video stabilization using optical flow. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 8159-8167).
20. Chiu, L. C., Chang, T. S., Chen, J. Y., & Chang, N. Y. C. (2013). Fast SIFT design for real-time visual feature extraction. IEEE Transactions on Image Processing, 22(8), 3158-3167.
21. Lowe, G. (2004). Sift-the scale invariant feature transform. Int. J, 2(91-110), 2.
22. Bay, H., Tuytelaars, T., & Van Gool, L. (2006, May). Surf: Speeded up robust features. In European conference on computer vision (pp. 404-417). Springer, Berlin, Heidelberg.
23. Rublee, E., Rabaud, V., Konolige, K., & Bradski, G. (2011, November). ORB: An efficient alternative to SIFT or SURF. In 2011 International conference on computer vision (pp. 2564-2571). Ieee.
24. Ali, M. K., Yu, S., & Kim, T. H. (2020). Learning Deep Video Stabilization without Optical Flow. arXiv preprint arXiv:2011.09697.
25. Guilluy, W., Oudre, L., & Beghdadi, A. (2021). Video stabilization: Overview, challenges and perspectives. Signal Processing: Image Communication, 90, 116015.
26. Lee, Y. C., Tseng, K. W., Chen, Y. T., Chen, C. C., Chen, C. S., & Hung, Y. P. (2021). 3D Video Stabilization With Depth Estimation by CNN-Based Optimization. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 10621-10630).
27. Xu, Y., Zhang, J., Maybank, S. J., & Tao, D. (2020). DUT: Learning Video Stabilization by Simply Watching Unstable Videos. arXiv preprint arXiv:2011.14574.
28. Choi, M., Kim, H., Han, B., Xu, N., & Lee, K. M. (2020, April). Channel attention is all you need for video frame interpolation. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 34, No. 07, pp. 10663-10671).
29. Shi, W., Caballero, J., Huszár, F., Totz, J., Aitken, A. P., Bishop, R., ... & Wang, Z. (2016). Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1874-1883).

Learning-Based Image Rendering

Muhammed Kerim SOLMAZ[1], Mehmet SARIGÜL[2], Levent KARACAN[3]

[1,2,3] Department of Computer Engineering, Iskenderun Technical University, Turkey
Emails: msolmaz.mfbe20@iste.edu.tr, mehmet.sarigul@iste.edu.tr, levent.karacan@iste.edu.tr

Abstract

Image rendering is essential study field for computer science, robotics, and augmented reality. In the last decade, the increase in the graphics processing power of computers and the widespread use of deep learning networks have led to deep learning networks being at the heart of the studies on image rendering. The use of deeper networks improves the visual representation ability of the trained models and gives them the ability to render high quality images. In this study, various information is given about subjects such as image rendering, obtaining 3D data from 2D data, 3D image rendering, differentiable rendering and recent studies on this subject.

Keyword(s): computer vision, depth image, point cloud generation, differentiable rendering

1. Introduction

In computer science, data is kept in matrices of different sizes. Image rendering aims to extract meaningful visuals from these matrices. Recent studies with image rendering, 3D rendering, and differentiable rendering in the computer science community are mentioned in this study.

The minimum number of coordinates required to specify a point on an object is called dimension. Lines are one-dimensional because they lie on a single axis. Visuals are 2-dimensional because a point on the image is specified with two coordinates, x, and y-axis. Three-dimensional objects are x,y and z it is specified with three coordinates. What distinguishes 2-dimensional and 3-dimensional objects from each other is that 3-dimensional objects are additionally specified with z coordinates. Three-dimensional space; is the name given to the medium that has width, height, and depth information.

Datasets with known depth information and camera intrinsic parameters are more suitable for use for 3D studies. This is because they can be represented more accurately and easily in 3D space by utilizing depth information. By using depth information, the object can be represented in 3D space. Objects represented in 3D space can be transferred back to 2D space after the desired transformation processes are performed.

Rendering is the process of obtaining an image from 2D or 3D data. The resulting image is called a rendered image. It is one of the sub-topics of computer graphics and has an essential place in computer graphics. It is a method needed and used in many different sectors, not only in computer graphics but also in architecture, marketing, construction,3D mapping, video games, etc. Usually, a 2D render is obtained from 3D data. Information such as the geometry of the 3D data, camera perspective, and lighting must be known.

One of the missing aspects of the rendering process is that it is not possible to propagate back through the render output after the render output is taken. As a solution to this problem, the differentiable rendering [4] process has been developed. In the differentiable rendering process, it is possible to return over the render output.

2. Materials and Methods

2.1. The image depth and depth completion

In computer vision, the concept of depth is the information about the distance of an object. The depth map is the image that indicates the depths of all things in an environment. Today, it is possible to create a depth map with various methods (laser scanning devices, cameras that can measure depth, and similar tools). It is essential for studies. It is so important because it is much easier to represent 2D data with depth information 3D. Not all cameras capable of measuring depth give successful results. These cameras are sufficient to help get an idea about depth information, but they create some missing regions in the depth map for critical studies. These missing regions cause the 3D representation to be missing during the 3D representation for 3D studies and reduce the work's efficiency. Due to this problem, various depth map completion algorithms have been developed [1],[2],[3]. These smart algorithms have been developed to complete the missing regions in the depth map. Depth maps completed using these algorithms are more suitable for 3D studies.

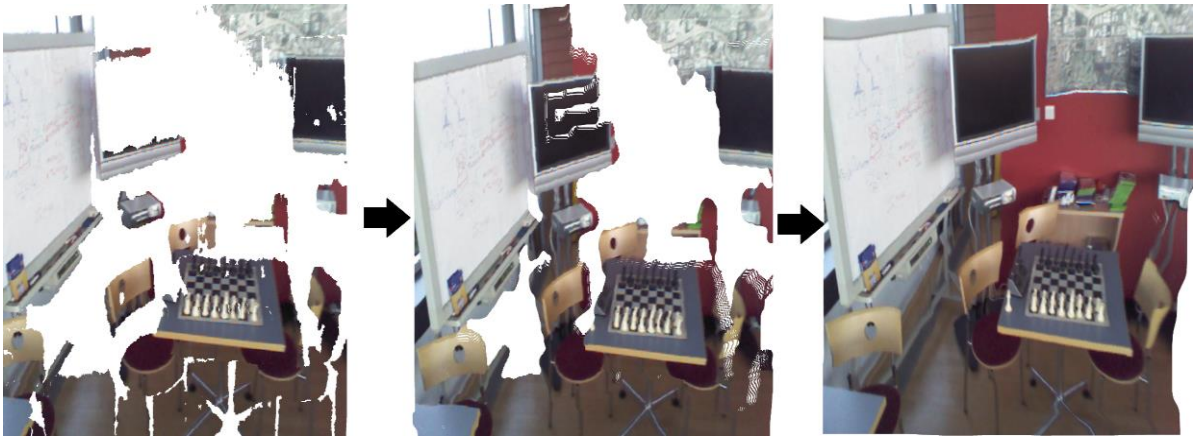


Figure 1. Depth compilation

2.2. Point Cloud

A point cloud is a representation of an object as points in 3D space.Points are positioned according to depth information and camera intrinsic values in the 3D coordinate system. By using the all points in 3D space, the whole object can be observed. A point cloud is a very nice tool for examining and studying a 3D object. The object can be examined from all angles, and points can be manipulated according to the work to be done. It is possible to create a colored point cloud [9] in data with RGB-D (color and depth) knowledge. Nowadays, the process of creating a point cloud is relatively easy when enough data is available. Various algorithms are

available for this process [5],[6]. These algorithms ensure that 3D data is represented with points.

Figure 2. RGB-D image to point cloud

2.3. Differentiable Rendering

Image rendering is one of the cornerstones of computer vision. This is because data can be represented visually. The work area is pervasive and is used in almost all studies related to computer vision. Although it is used frequently, image rendering has some problems. At the beginning of these problems are the losses that occur when the 2D render output passes back to the 3D space. Since there is no back-propagation after rendering, it is impossible to return to the 3D space without loss. This problem has been solved with differentiable rendering [4]. With the differentiable rendering process, it becomes possible to return to the 3D space when necessary after rendering by relating the properties of the 3D scene. Differentiable rendering has become a very popular field of the study recently. It has been used in many studies.Some of these studies are, NeRF[7] work that makes 3D scene estimation by spreading back the data such as light, opacity from the 2D images, SynSin[8] is the estimation of the image that will be formed when the camera perspective of the 2D image is changed, Pytorch3D[9] work that creates colored point clouds in 3D space and returns from these point clouds back to 2D space, deep learning based 3D rendering work using Cuda cores.[10]. It seems that its popularity will continue to increase.

3. Future Directions

These days, when the importance of learning-based methods is better understood, computer vision, deep learning, machine learning, and computer graphics have become fields of study where it is inevitable to use them together. Differentiable rendering strengthens the connection by increasing the relationship between these fields of study. Differentiable rendering will be the main subject of computer vision studies to be made in the future. The reason for this is that it allows learning-based studies because it pushes back-propagation possible. With this advantage, it is likely that studies on such as scene estimation, pose estimation, 3D visualization errors, and 3D optimization will increase. The development of these fields of work will allow the direct action of many industries (architecture, marketing, construction, 3D mapping, video games, etc.). Therefore, it will be a method that will raise many sectors, especially learning-

based 3D rendering studies and computer vision studies, to much more advanced levels than they are now.

4. Conclusion

As a result, differentiable rendering and learning-based image rendering should be studied for 3D studies in these times when computer vision is gaining importance. Anyone working in computer graphics, deep learning, or computer vision should know about differentiable rendering and learning-based rendering.

5. Acknowledgments

This study was supported by project number 120E447 from the The Scientic and Technological Research Council of Turkey (TUBITAK).

References

1. Zhang, J., Ramanagopal, M. S., Vasudevan, R., & Johnson-Roberson, M. (2020, May). Listereo: Generate dense depth maps from lidar and stereo imagery. In 2020 IEEE International Conference on Robotics and Automation (ICRA) (pp. 7829-7836). IEEE.
2. Imran, S., Long, Y., Liu, X., & Morris, D. (2019, June). Depth coefficients for depth completion. In 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 12438-12447). IEEE.
3. Zhang, Y., & Funkhouser, T. (2018). Deep depth completion of a single rgb-d image. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 175-185).
4. Kato, H., Beker, D., Morariu, M., Ando, T., Matsuoka, T., Kehl, W., & Gaidon, A. (2020). Differentiable rendering: A survey. arXiv preprint arXiv:2006.12057.
5. Zhou, Q. Y., Park, J., & Koltun, V. (2018). Open3D: A modern library for 3D data processing. arXiv preprint arXiv:1801.09847.
6. Luo, S., & Hu, W. (2021). Diffusion probabilistic models for 3d point cloud generation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 2837-2845).
7. Mildenhall, B., Srinivasan, P. P., Tancik, M., Barron, J. T., Ramamoorthi, R., & Ng, R. (2020, August). Nerf: Representing scenes as neural radiance fields for view synthesis. In European conference on computer vision (pp. 405-421). Springer, Cham.
8. Wiles, O., Gkioxari, G., Szeliski, R., & Johnson, J. (2020). Synsin: End-to-end view synthesis from a single image. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 7467-7477).
9. Shinohara, T., Xiu, H., & Matsuoka, M. (2021). Point2color: 3D Point Cloud Colorization Using a Conditional Generative Network and Differentiable Rendering for Airborne LiDAR. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 1062-1071).
10. Ravi, N., Reizenstein, J., Novotny, D., Gordon, T., Lo, W. Y., Johnson, J., & Gkioxari, G. (2020). Accelerating 3d deep learning with pytorch3d. arXiv preprint arXiv:2007.08501.

Low Cost Multifunctional Security System That Does Not Require Hardware Design

Tahir Alpay[1], Ersin Özdemir[2]
[1,2] Department of Computer Engineering, Iskenderun Technical University, Turkey
Emails: tahiralpay.mdb16@iste.edu.tr[1], ersin.ozdemir@iste.edu.tr[2]

Abstract

This study was started with the aim of developing a low-cost city security system. In the study, a computer vision-based smart security system was implemented by using software developed on Raspberry Pi and Arduino. The designed system is motion sensitive, takes photos and video images from the environment and sends data about theft, fire, earthquake, gas leakage, air quality to the user via WIFI and/or GSM. Due to the size, ease of use and wireless design of the system, it can safely collect data in houses and workplaces without attracting attention. The security system has also been designed taking into account the outdoor conditions. It can be used in outdoor research, security applications, early detection of forest fires.

Our work continues in order to add artificial intelligence algorithms that recognize the user and operate accordingly to the security system.

Keyword(s): Vision based security system; Photo trap; WIFI, GSM; Raspberry Pi; Arduino

1. Introduction

Our homes and workplaces are environments, where the objects we value materially and morally are the most. The safety of these places is important. According to the Statistical Data of the Penitentiary Institution, the number of people involved in crime on 31 December 2019 increased by 10.1% compared to 2018. The most committed crime is theft with 15.2% (Penal Institution Statistics, 2019). On the other hand, incidents such as gas leaks and fires in residences increased by 11.8% compared to 2008 (Kılıç, M. 2018). As can be understood from these statistical data, security problems are increasing in our environment. For this reason, it is tried to take the highest level of precautions.

The study consists of two modules. The first module is the Raspberry-controlled motion detection system. The system has 4 different functions in itself. Functions are selected according to the availability of the Internet and the characteristics of the environment. In case of theft, the system takes snapshots and videos at specified times and sends them to the user via e-mail and SMS. It is important that the data collected in security systems are delivered to the user quickly. For this reason, communication is provided in a fast and high quality manner using Simple Mail Transfer Protocol (SMTP) and sim800l.

In the second part, events such as shaking, gas leakage, sudden temperature change, air quality and fire in the environment where the system is installed are read through sensors and processed with an algorithm, and when necessary, the user is informed via SMS.

With the developments in technology, security systems are also developing. Intelligent security systems are increasing their popularity day by day, as computers are reduced to the size of credit cards and are easily accessible (Godbole, S., et al. 2016). In the studies, security systems are generally presented as two different studies in which motion-detection camera and

fire, earthquake, and air quality data are evaluated. In the motion detection camera system, systems are not designed for different environmental conditions and the user is only warned by SMS (Göçmen, O., & Kıvanç, G. 2021) In the warning method, the warning system is insufficient except for certain features (Jones S.R, et al. 2018). In this study, the two systems were combined and designed to be used with or without internet for the safety of environments such as streets, houses, workplaces, and natural protection areas.

2. Materials and Methods

Arduino, Raspberry were used in the design of the system and SIM800L GSM card was used for communication. The algorithm of the project is given in Figure 1.

2.1. Arduino

Arduino is an open source development platform that provides ease of use compared to many microcontrollers. It provides great convenience in terms of coding through the libraries in Arduino. Thanks to the analog and digital inputs on the Arduino board, it provides the opportunity to process analog and digital data. For this reason, it is preferred for electronic projects today.

2.2. Raspberry Pi

Raspberry Pi is a single board computer with open source software. It can be used in all kinds of computer projects (such as software development, embedded systems and operating system applications). In addition, stable structure, easy use, cost are some of the advantages of Raspberry Pi.

2.3. SIM800L

The SIM800L module is a development board that allows the use of GSM communication and GPRS with an antenna. This card provides the opportunity to forward and receive calls, receive and send SMS, location tracking, and remote control of devices. The board contains Universal Asynchronous Receiver Transmitter (UART) pins for connection with the microcontroller. The band speed of the module can automatically detect the band between 1200bps and 115200 bps.

2.4. Sensors

The sensors, circuit boards and protocols used in the developed system are explained in order.

DHT11 is a temperature and humidity sensor with an 8-bit processor. DHT11 can measure the temperature in the range of 0-50°C with an accuracy of ±2°C.

The PIR sensor uses a pyroelectric sensor to detect motion and detects infrared light. The working principle is based on the difference change between the sensors sensitive to the IR wavelength in the sensor.

The MQ9 Sensor Board measures Carbon Monoxide and Combustible Gas, detects the presence of 100 - 10,000 ppm combustible gas concentration and 10 - 10,000 ppm Carbon Monoxide (CO) in the environment.

The SW-420 is a high-sensitivity, non-directional vibration sensor. Although its usage area is wide, it is generally used in applications such as burglar alarm and earthquake alarm. It gives digital output (0/1) when it detects vibration.

Fire Detector Sensor Board (Flame sensor) is a sensor board that can detect flame at wavelengths between 760 nm and 1100 nm thanks to the IR sensor on it.

Radio Frequency Identification technology (RFID) is a method of recognizing objects using radio frequency. RFID basically consists of a tag and a reader. RFID tags can be programmed to read, write and store object information such as product code, password, name (Nayak, M., & Dash, P., 2014)

2.5. Simple Email Transfer Protocol (SMTP)

SMTP is the protocol that determines the way of communication between the server and the client to send an e-mail. Free e-mail service providers have started to support the SMTP protocol. (Hoffman, P., 2002)

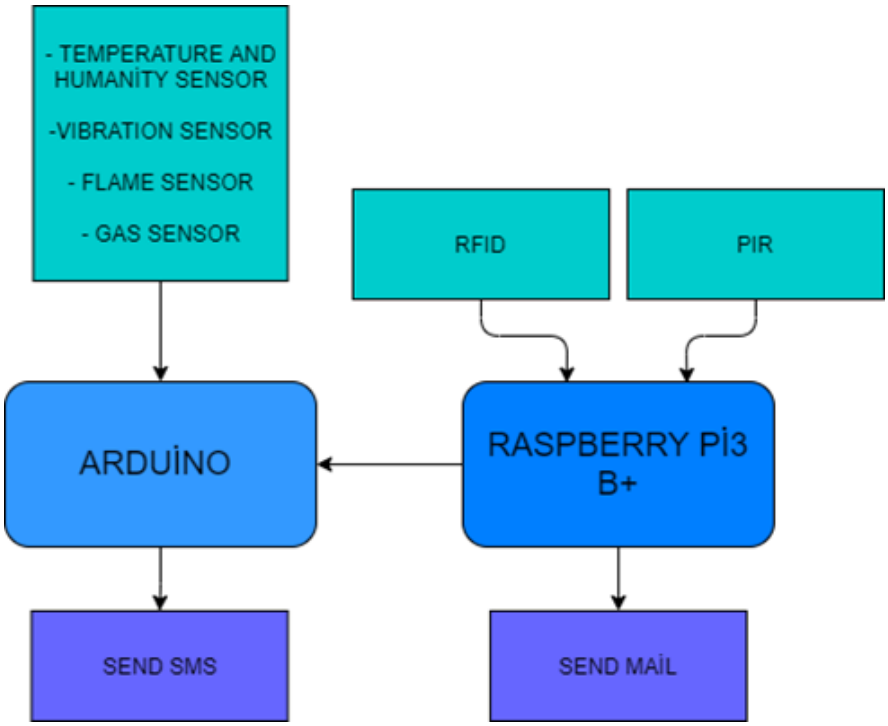


Figure 1. General flow diagram of the system

3. System Design and Implementation

The operation of the system is modular. In this section, the modules are explained in order.

1. General working principle of the master module

There are 4 different functions in this module. The functions are selected according to the environment in which the system will work and whether the internet is available. Switching between functions is provided by master RFID. In general, the functions; After detecting the first motion, two different counters, motion_on and motion_off, run. motion_on counter starts

counting as soon as PIR sensor is active and stops counting when motion stops. The movement_off counter starts counting as soon as the movement stops. If motion is detected during the counting of motion_off counter, motion_on counter becomes active again. “1” second after motion is detected, the photo mode is activated and the image is taken. At this stage, if motion is not detected for “1” minute, the motion_on counter is reset and the program returns to the beginning. “6” seconds after the motion is detected, the video mode is activated and a “10”-second image is taken, the motion_on and motion_off counters are reset and the SMS pin is activated.

The sub-functions of the system are described below. The sub-functions of the system are described below.

1. Function

This function is designed for continuous recording environments where there is no internet. This mode, which is suitable for outdoor environments, allows to use the system as a photo trap. It can be used in environments such as fields, vineyards, gardens and forests. Received images are saved in folders according to date and time and the user is warned via SMS. By making instant WIFI connection to the security system, the user can log in with the code and password of the security system and access the data. Figure 2 shows the flowchart of the 1st and 2nd functions.

2. Function

It is designed for continuous recording where there is an internet connection. This mode, which is suitable for outdoor environments, allows the system to be used as a security camera. It can be used in places such as home, apartment. In this function, the images collected with the camera (Raspberry Pi camera ver1.3) are recorded with a name and the images are sent to the user via e-mail from an e-mail address defined via the SMTP protocol, and also the user is warned via SMS. Since there will be a delay during the mail sending, the time to enter the photo or video is adjusted accordingly in the software.

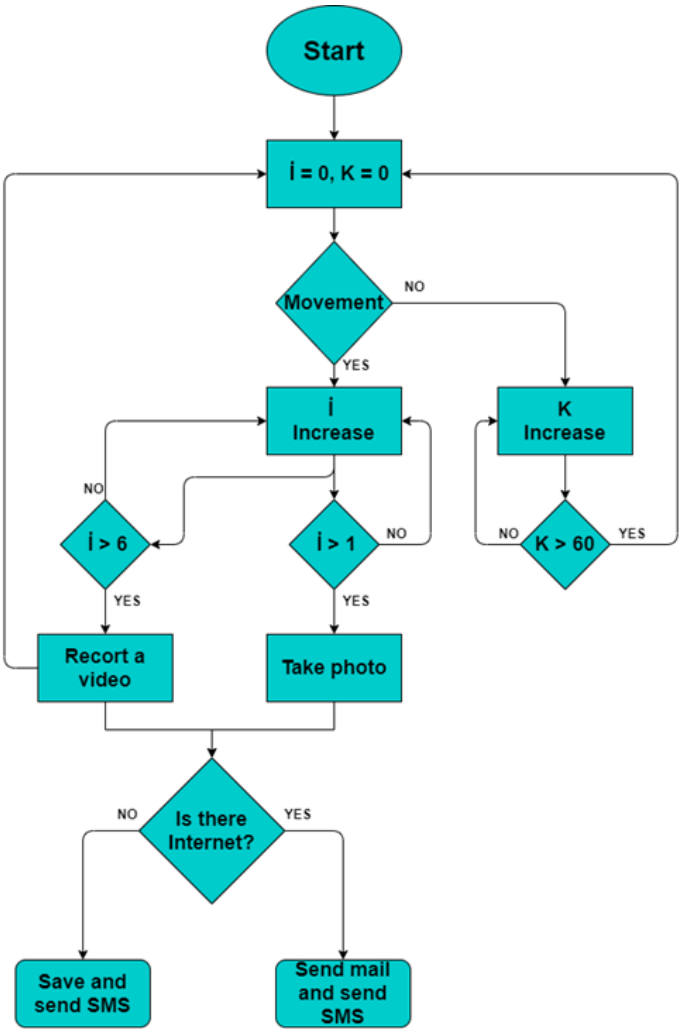


Figure 2. Flowchart of 1st and 2nd function

3. Function

This function is activated by detecting the defined RFID card of the system, images (.jpg, and/or .h264) recorded according to date and time are sent to the user via SMS. By making instant WIFI connection to the security system, the user can access the data by logging in with the code and password of the system with a remote access software. (Yüksel, M. E., & Odabaşı, Ş. D. 2009)

4. Function

After detecting the RFID card, the data is sent to the user via e-mail, and the user is warned via SMS. RFID lock is used to prevent warnings that may occur due to mobility at home or work places. The device remains on standby until the user presents the RFID card to the device.

In case the internet is active (2nd and 4th functions), the data is sent directly to the user by e-mail and warned by SMS. In Figure 3, a screenshot of the message sent to the user is given on the python shell screen. Figure 4 and Figure 5 show the photo and video sent to the user's e-mail address.

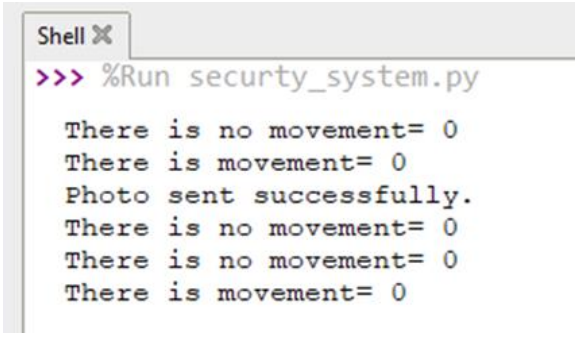


Figure 3. Python Shell Screen

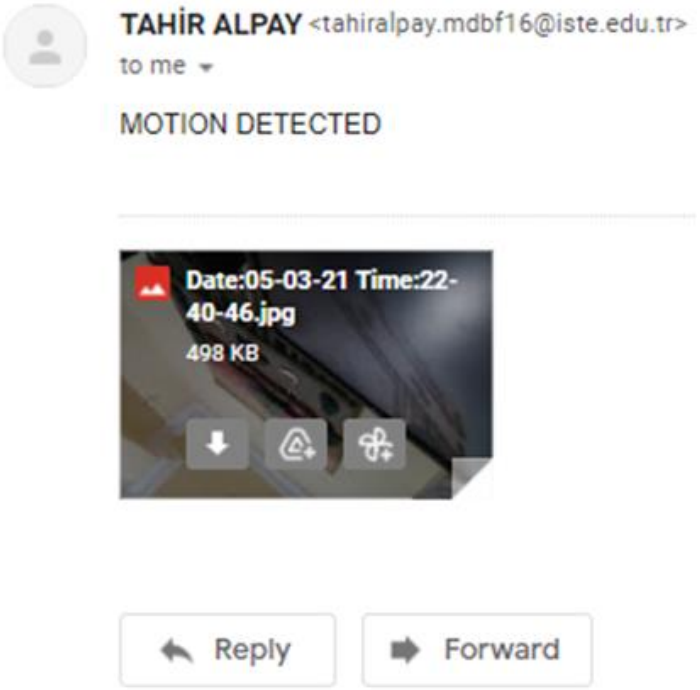


Figure 4. Sending the Photograph to the Recipient by Mail

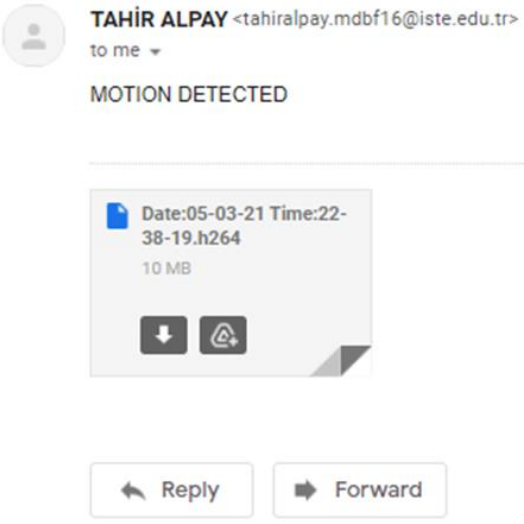


Figure 5. Delivering Videos to Recipients by Mail

In the 1st and 3rd functions where there is no internet, data can be obtained by logging into the system via remote desktop applications. With the remote desktop application, the data in the sample folders shown in Figure 6 and Figure 7 were viewed or copied.

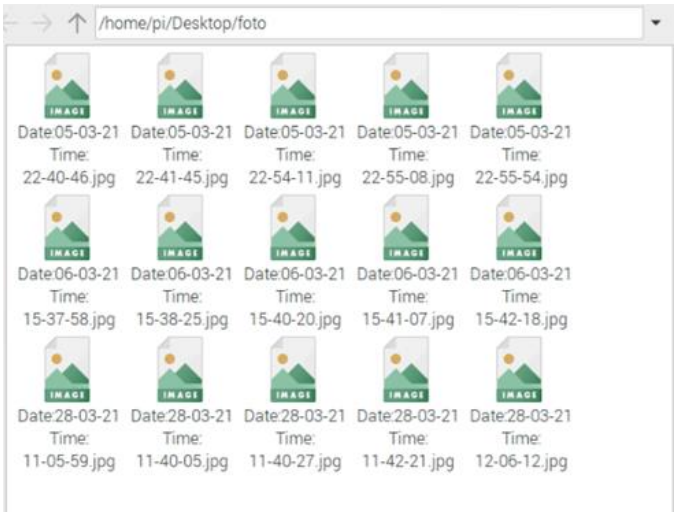


Figure 7. Photos saved in the Security System

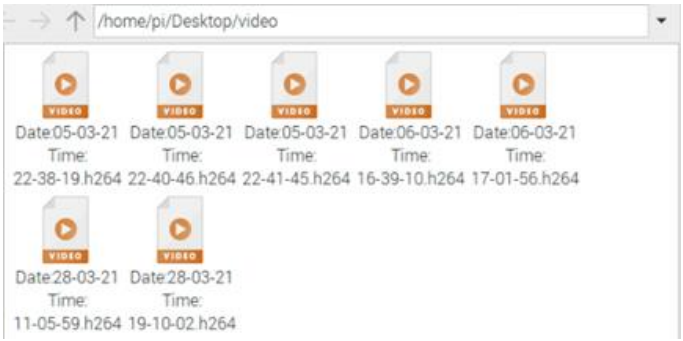


Figure 7. Videos Stored in the Security System

General Working Principle of the System

The Arduino reads and compares the data from the connected sensors. When it is alert or active on the SMS pin, the SMS command block is activated and the user is warned.

While the SW-42 vibration module gives a digital 0 output under normal conditions, it gives a digital 1 output in case of any earthquake and shaking. The SMS function, when Arduino detects this digital 1 output, is activated and the user is warned via SMS as shown in Figure 8.

Figure 8. Ground Shaking - Earthquake warning

When the analog value read from the MQ-9 Carbon Monoxide and flammable gas sensor exceeds the specified threshold value (threshold value = 400), the SMS function is activated and the user is warned via SMS as shown in Figure 9.

Figure 9. Gas leak alert to the Client

In this project, two different methods were used to detect the fire hazard. These;

1- Flame sensor gives digital 1 output under normal conditions. When it detects a flame, it gives a digital 0 output and the SMS function is activated.

2- The data obtained from the carbon monoxide and combustible gas sensor card (Mq-9) and the temperature and humidity sensor (DHT11) cards are evaluated with the algorithm. If there is an increase in temperature and chemical gases before the flame goes out, the SMS function is activated and the user is warned.

A screenshot of the warning message sent for the fire hazard is given in Figure 10.

Figure 10. Alerting the customer to the fire hazard

Motion detection is done with the image processing program running on Raspberry Pi. When a motion is detected in the image, SMS is activated via the 32nd pin and it is aimed for the user to check their mail as soon as possible. (Figure 11)

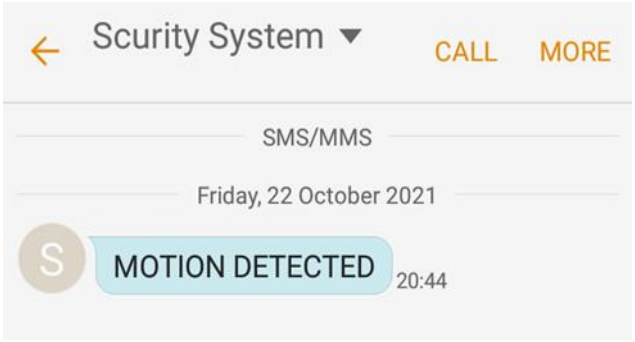


Figure 11. Screenshot of motion detection alarm

As an additional feature to the developed security system, studies on keyless password entry with the help of computer vision have been completed at a modular level. In this way, the user and the security system communicate with a kind of sign language. The user's fingers are detected by the camera, the number of fingers and the digit of the password are determined and the password is entered. Tests on the compatibility of this module with the system are ongoing. (Figure 12 and 13) With this method, it is aimed to activate or deactivate the system without using the keyboard. (Ravikiran, J., et al., 2009)

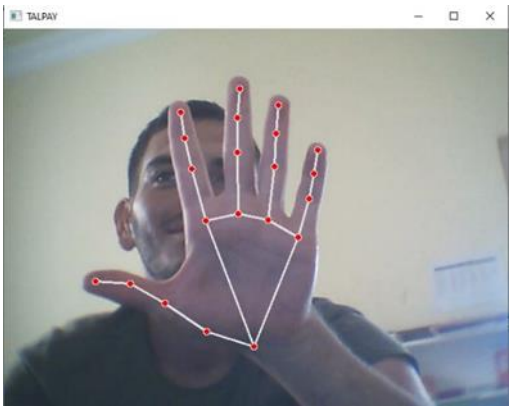


Figure 12. Detection and counting of fingers from the image

Figure 13. Writing the number of fingers on the LCD screen

5. Results and Discussion

The designed security system has been tested for 4 different scenarios. When motion is detected, the system performed the operations determined in the algorithm chosen in accordance with the scenario. Event images were recorded in defined places depending on the scenario, and the user was warned via email and SMS. In addition, in the developed system, events such as decrease in air quality, fire early warning, concussion and earthquake were created artificially and the user was warned by the security system.

6. Conclusion

With this study, a security system that is more comprehensive than other security systems on the market, does not require hardware design, consists of easily available modules, is open source to development and has a low cost.

Studies on the subject continue on learning the past behavior of the user, training artificial neural networks with this data and guiding the algorithm accordingly.

References

1. Ali, H, Erçelebi, E . (2021). Development of an Embedded System for Building System Management Based on PV-Powered. Balkan Journal of Electrical and Computer Engineering, 9 (3) , 249-254 . DOI: 10.17694/bajece.933353
2. Ceza İnfaz Kurumu İstatistikleri, (2019). TÜİK - Veri Portalı https://data.tuik.gov.tr/Bulten/Index?p=Prison-Statistics-2019-33625, the address has been reached at 20.10.2021.
3. Hoffman, P. (2002). SMTP service extension for secure SMTP over transport layer security. RFC 3207, February.
4. Kılıç, M. (2018). Gelişmiş ülkelerde ve Türkiye'de yangın istatistikleri. Yangın ve Güvenlik Dergisi, 199, 8-10.
5. Godbole, S., Deshpande, S., Barve, N., & Galim, S. (2016). Review on Theft Prevention System using Raspberry Pi and PIR Sensor. International Journal of Computer Applications, 155(11).
6. Göçmen, O., & Kıvanç, G. i-RoHEAD. (2021) "Eğitim, Eğlence ve Güvenlik Amaçlı Kullanılabilen İnteraktif Robotik Kafa", https://docplayer.biz.tr/51859885-I-rohead-egitim-eglence-ve-guvenlik-amacli-kullanilabilen-interaktif-robotik-kafa-ogr-gor-ogul-gocmen.html, the address has been reached at 20.10. 2021
7. Nayak, M., & Dash, P. (2014). Smart surveillance monitoring system using Raspberry Pi and PIR sensor.
8. Ravikiran, J., Mahesh, K., Mahishi, S., Dheeraj, R., Sudheender, S., & Pujari, N. V. (2009). Finger detection for sign language recognition. In Proceedings of the international Multi Conference of Engineers and Computer Scientists (Vol. 1, No. 1, pp. 18-20).
9. Jones S.R, Beardmore J, Gillott M, Boukhanouf R, Walker G. A, (2018), control methodology for Building Energy Management Systems (BEMS) in heat networks with distributed generation. Energy Procedia;153:295– 302. doi:10.1016/j.egypro.2018.10.011.
10. Yüksel, M. E., & Odabaşı, Ş. D. (2009). Nesneler İzlenebilir ve Yönetilebilir mi? Çözüm: RFID. Akademik Bilişim, 11-13.

Mobile EEG Measurement System Design

Enver Kaan Alptürk[1*], Yakup Kutlu[2]

[1]Iskenderun Technical University, Institute of Engineering and Natural Sciences, Hatay, Turkey
[2]Technical University, Faculty of Engineering and Natural Sciences, Hatay, Turkey

E-mail: [1] kaanalpturk@outlook.com [2]yakup.kutlu@iste.edu.tr

Abstract

Electroencephalography (EEG) are small amplitude and low frequency electrical signals received from the brain through electrodes placed in the skull. EEG signals can be easily disturbed due to external factors. For this reason, it must undergo a series of filtering processes at both the hardware layer and the software layer. EEG signals are already low amplitude signals. Since they also lose power during filtering, the signal output must be strengthened, that is, the amplitude must be increased. For this, Instrumentation Amplifiers have been used.

In this study, an Instrumentation Amplifier circuit has been designed for electrodes that allow EEG recording, and the received signals are passed through a series of filtering processes in both the hardware layer and the software layer. It is aimed that the received brain signals can perform motor activities of disabled individuals by using the channels determined for "Motor Activities" in previous studies.

Keywords: EEG, signal filtering, instrumentation amplifier, motor activity

1. Introduction

EEG signals are biological electrical signals obtained from the brain. EEG signals are affected by the electrical and magnetic field in the environment and by all physical and mental activities of the individual to whom the electrode is attached (Alptürk & Kutlu, 2021). Therefore, before working on EEG signals, the received signals must be processed through a series of processes. These operations are basically filtering and amplitude amplification. Since EEG signals are of very low amplitude, they are difficult to detect. Because environmental factors, the tension resistance of the human body, and the intensity of muscle signals make it difficult to detect brain signals. The most important environmental factor is the mains voltage. The frequency of this voltage, which is 60Hz in some regions, is 50Hz in some regions. The signals must be filtered so that the mains voltage does not affect the EEG signals. This filtering can be done at the hardware layer or at the software layer. The filter may need to be applied at the hardware layer in order to get instant results and minimize data processing time. EEG signals are examined in five main groups according to their frequency ranges (Table 1). Frequency ranges provide information about the psychological and biological states of the brain (Kumar & Bhuvaneswari, 2012).

Group	Frequency	The cases where is seen
Delta	0.5-3.5 Hz	Adults: Occurs in sleep mode.
Tetha	4-7 Hz	Mostly occur in stressful situations
Alpha	8-12 Hz	When awake, it intensifies during the transition to listening mode.
Beta	13-30 Hz	Seen in situations of concentration, stress, anxiety.
Gamma	30+ Hz	Seen in certain motor brain functions.

Table 1. Grouping of EEG signals according to frequencies.

EEG signals are collected by some special electrodes and recorded by passing certain pre-processing with special circuits.

Since the amplitudes of the EEG signals taken from the electrodes are below 100 μV, it is very difficult to detect and distinguish them. In addition, the noise intensity is quite high. Therefore, the signals need to be filtered and their amplitudes amplified.

Bo Luan et al. recorded EEG signals through skin screw electrodes in 2012. They used INA118 as first stage amplifier and TLC277CP as operational amplifier. They have been increased the signals obtained by means of a potentiometer between 5.76 and 101 times (Luan, Sun, & Jia, 2012).

In 2013, L. Zhang et al. have been conducted a portable Brain-Computer Interfaces (BCI) study using two bipolar silver chloride electrodes. They have been filtered the 50 Hz mains voltage by applying a mode rejection filter, and then increased the signal 13.5 times. Finally, they have been filtered low-frequency noise with a 0.5 Hz passive RC High-Pass filter and increased the signal 51 times with the main amplifier. They have been used INA118 as amplifier (Zhang, Guo, Wu, & Zhou, 2013).

Marzieh Moradi et al. have been used a CMOS-based amplifier on the portable EEG circuit they designed in 2021. The amplifier achieves a midband gain of 70 dB and a bandwidth of -3dB in the 0.1-212 Hz range. The adjustable LPF has a cut-off frequency of 100 Hz. They have been used a continuous time second-order 100 Hz Gm-C Low-Pass Filter to reduce the ripples at the output of the amplifier. They have been achieved an output of 92 dB at the amplifier output (Moradi, Dousti, & Torkzadeh, 2021).

In this study, brain signals were obtained from the FP1 channel using 3 silver electrodes. The signals have been passed through the passive RC Low-Pass Filter and their amplitudes were amplified and transferred to the computer. The signals have been recorded during 3 different states and the accuracy of the study performed as a result of the frequency analysis of the signals is shown in the results section.

2. Materials And Methods

In this study, disposable electrodes, which are also widely used in ECG systems, have been used to detect brain signals. (Figure 1). Disposable electrodes have been preferred because they already contain gel that can eliminate body resistance and are completely hygienic. Disposable electrodes contain a silver/silver chloride (Ag/AgCl) electrolyte and a gel attached to this electrolyte. (Biyomedikal Cihaz Teknolojileri - Elektrotlar, 2007). Silver electrodes are the

most suitable electrodes for capturing brain signals. It provides the advantage of low voltage resistance and high conductivity.

Figure 1. Disposable electrodes.

Electrodes placed at 3 points of the skull enabled data to be obtained from the FP1 channel in the frontal cortex. The voltage differences have been transferred to the circuit by connecting the electrode used for the main signal to the FP1 channel, the negative electrode to the FP2 channel and the electrode used for the grounding to the TP10 channel.

48.23 Hz passive RC Low-Pass filter have been used to minimize the noise generated in the signals. In this way, both meaningful signals for EEG can be distinguished and the effects of noise caused by 50 Hz city voltage are reduced. A 100 µF capacitor and a 33 Ohm resistor have been used for the RC Low-pass filter (Figure 2).

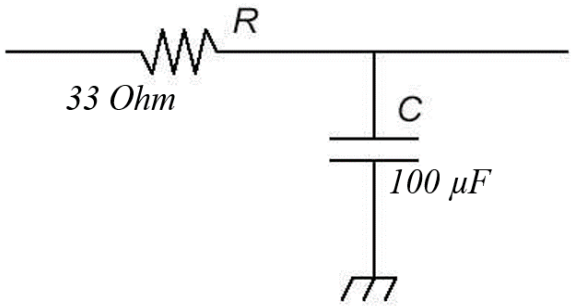


Figure 2. RC Low-Pass filter

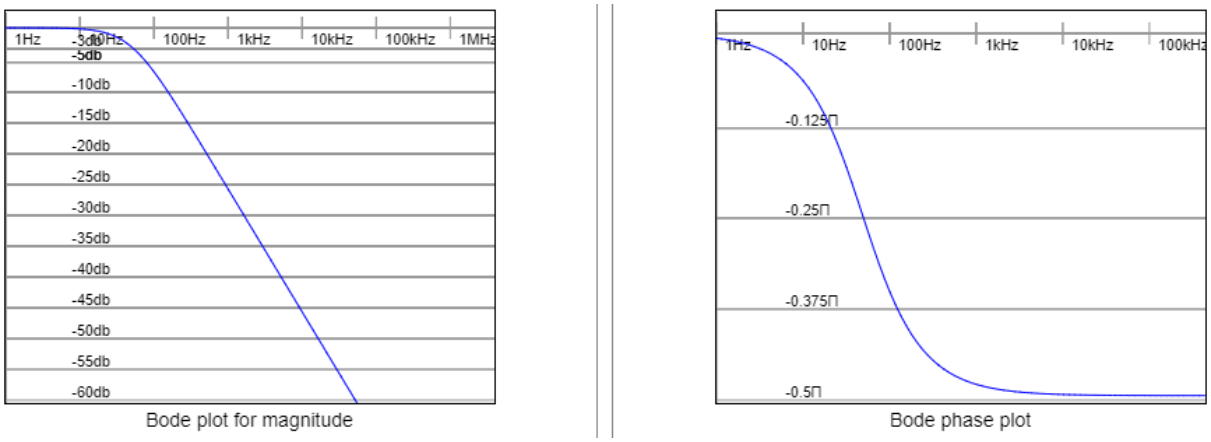
The output value of the signal filtered by the RC Low-Pass Filter is calculated using the formula in Equation 1 (Rao & Mothiki, 2006). The resistance (R) is calculated in ohms and the capacitor (C) in Farads.

$$F = \frac{1}{2\pi RC} \qquad \text{(Eq. 1)}$$

Since a 33 Ohm resistor and a 100 µF capacitor are used in this study, an output frequency of

$$F = \frac{1}{2\pi RC} = \frac{1}{2\pi \; x \; 33 \; x \; 0.0001} = 48.23 \; Hz$$

is obtained as a result of applying the equation. The magnitude and phase values (Graph 1) resulting from the 48.23 Hz RC Low-Pass Filter show how high frequency signals are attenuated.



Graph 1. Magnitude and phase values

With the simulation, sine waves of 20 Hz, 40 Hz, 60 Hz and 100 Hz were sent to the RC Low-Pass Filter, respectively, and it was observed that the filter reduced the amplitude of the high-frequency signals (Figure 3).

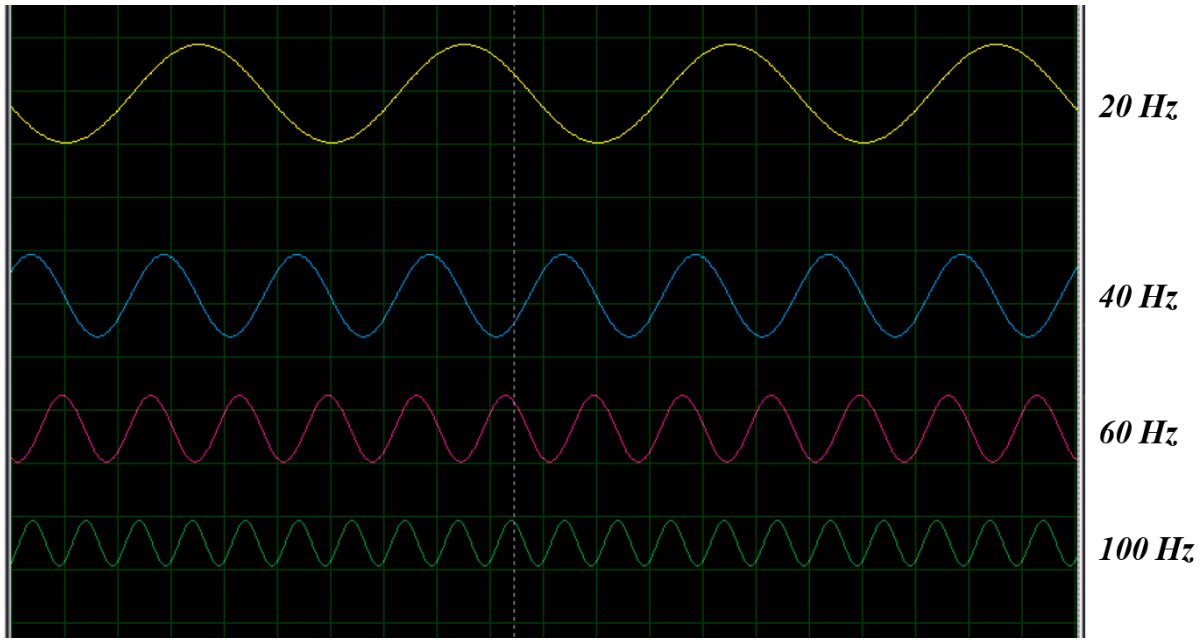


Figure 3. RC Low-Pass filter simulation

Since the amplitudes of frequencies below 48.23 Hz also decrease a little after the filtering of the signals, it will be difficult to distinguish the low amplitude EEG signals. For this reason, the AD620ARZ op-amp, which is an amplifier, have been connected to the circuit. The AD620 series provides the possibility to amplify the signal approximately 10000 times. AD620 gain is calculated with the formula in Equation 2 (Analog Devices Inc., 2011).

$$G = \frac{49.4K}{R_G} + 1 \qquad \text{(Eq. 2)}$$

Signal amplification in AD620 is done with the resistor (R_G) connected to the circuit. In this study, a 4.7 Ohm resistor was connected to the circuit and the gain was calculated as

$$G = \frac{49.4K}{R_G} + 1 = \frac{49.4K}{4.7} + 1 = 10.5K$$

In this direction, it is aimed to increase EEG signals smaller than 100 µV to 0.5 - 1 V. The circuit diagram model is shown in Figure 4.

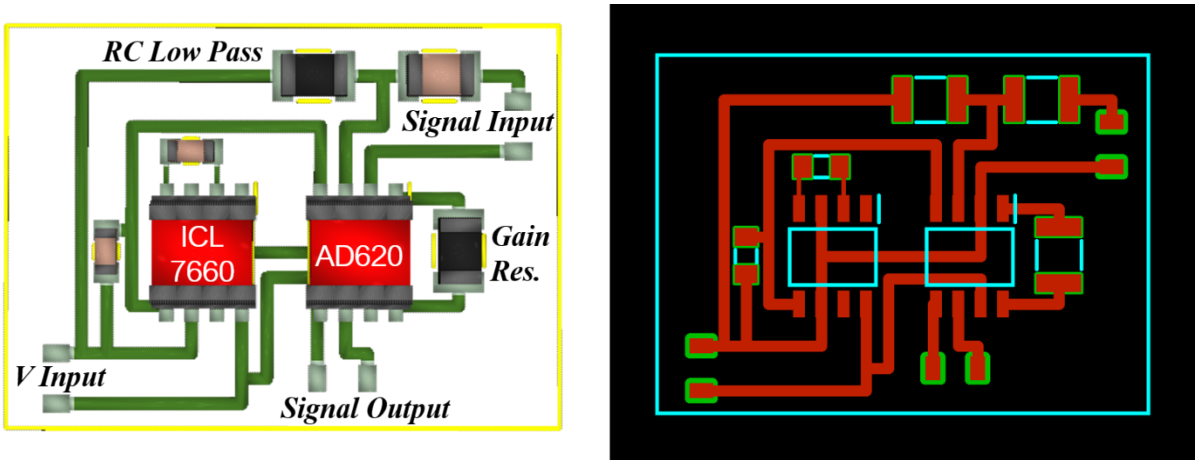


Figure 4. Circuit diagram.

3. Results

With the created circuit model, the EEG signals received through the electrodes were transferred to the Arduino UNO via analogue pins. Since the signals lost their amplitude after filtering, it was observed that their amplitudes were in the range of 0.3 - 0.5V as a result of the amplification process.

In this study, signals have been recorded based on three basic positions. The signals received from the Fp2 channel in the frontal cortex were transferred to the computer during nervous, resting state (eyes closed) and blinking movements, and analysed by Fast Fourier Transform (FFT) method. The results are shown in Figure 5, Figure 6 and Figure 7, respectively.

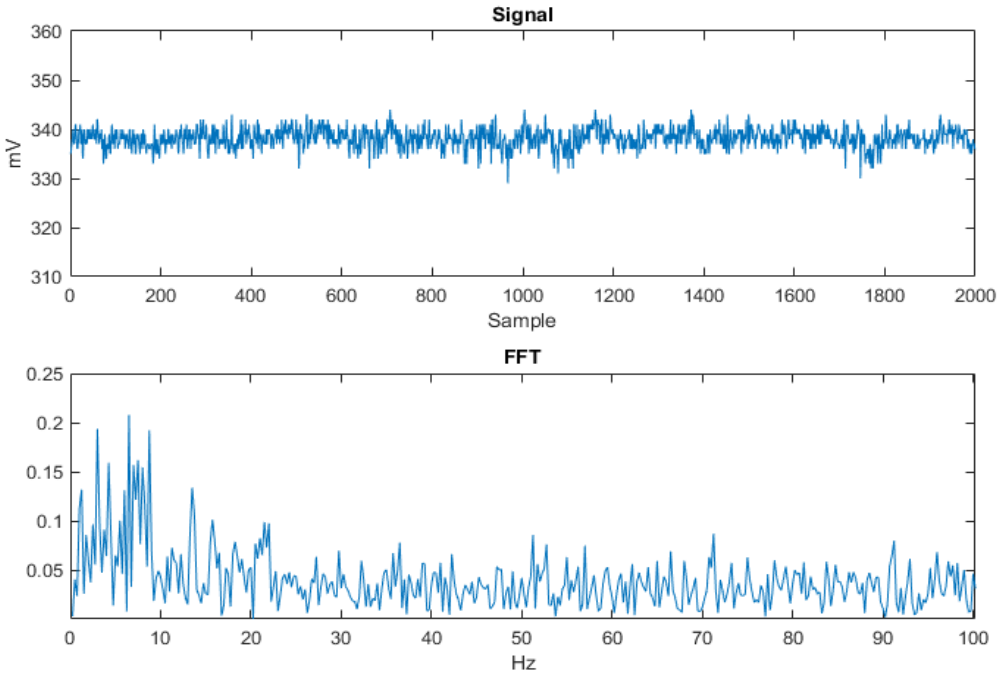


Figure 5. EEG signals received at the time of nervousness.

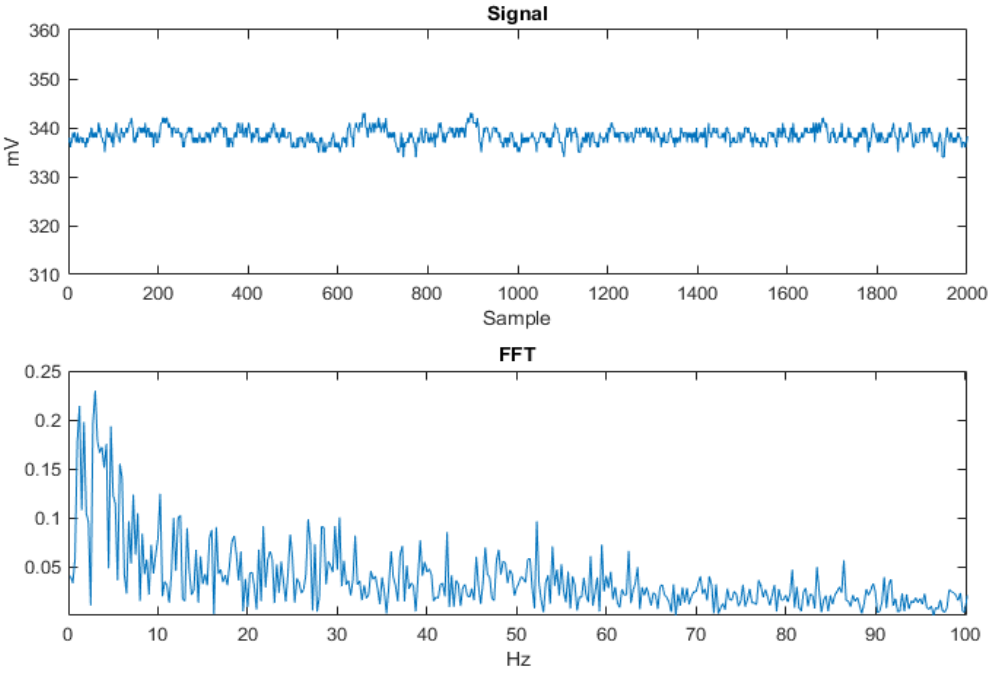


Figure 6. EEG signals received during resting (Eyes Closed).

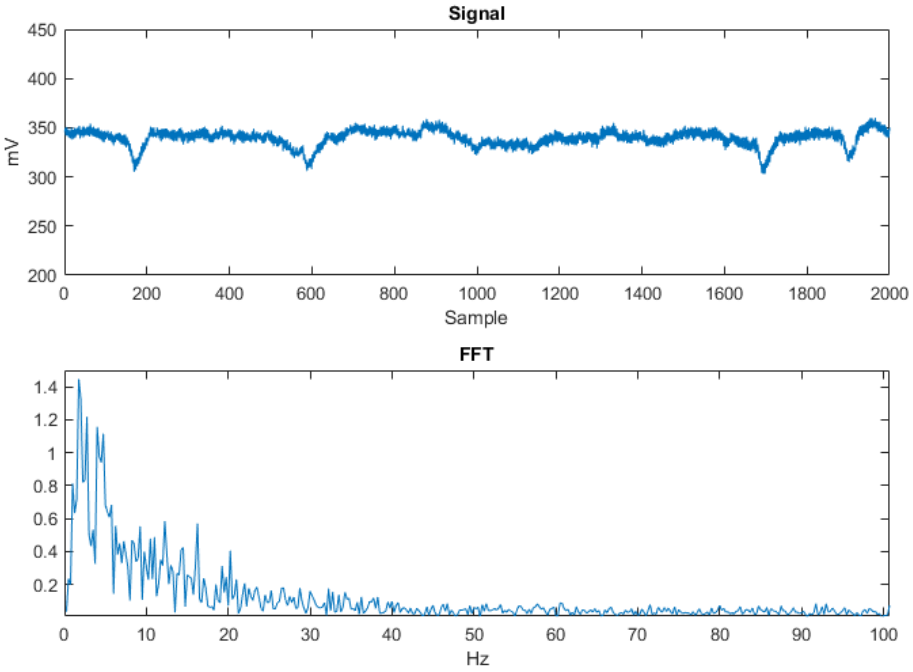


Figure 7. EEG signals received during eyes blinking.

As a result of the examination of the recordings taken during 3 different states, the distinctive features of the signals obtained from the nervous state, resting state and blinking activities could be clearly observed. While it was observed in Figure 5 that the signal frequency increased considerably during the nervous state, both the amplitudes and frequencies of these signals decreased during the resting state in Figure 6. In Figure 7, each eyes-blink created peaks of approximately 50 mV in the signal.

In line with the results obtained, it was seen that different signals were obtained with their distinctive features after different brain thought states. Frequency analysis doing with FFT demonstrated the accuracy of the materials and methods used.

4. Discussion

In this study, a 3-electrode mobile EEG measurement system was designed. In line with the examination of the signals obtained with the design, it has been observed that the study has achieved its purpose and will contribute to future studies. EEG signals come to a very important point to understand the structure of the human brain and to popularize the use of brain-controlled devices. It is aimed to carry out specific studies with different electrode combinations that will be replaced in the future.

References

1. Alptürk, E. K., & Kutlu, Y. (2021). Analysis of Relation between Motor Activity and Imaginary EEG Records. *Journal of Artificial Intellicence with Application*, 5-10.
2. Analog Devices Inc. (2011). *AD620 Datasheet*. Retrieved from Analog Devices: https://www.analog.com/media/en/technical-documentation/data-sheets/AD620.pdf
3. Biyomedikal Cihaz Teknolojileri - Elektrotlar. (2007). In *Mesleki Eğitim ve Öğretim Sisteminin Güçlendirilmesi Projesi (MEGEP)* (pp. 19-22). Ankara: T.C Milli Eğitim Bakanlığı.
4. Kumar, J. S., & Bhuvaneswari, P. (2012). Analysis of Electroencephalography (EEG) signals and its categorization–a study. *Procedia engineering*(38), 2525-2536.
5. Luan, B., Sun, M., & Jia, W. (2012). Portable amplifier design for a novel EEG monitor in point-of-care applications. *38th Annual Northeast Bioengineering Conference (NEBEC),*, (pp. 388-389).
6. Moradi, M., Dousti, M., & Torkzadeh, P. (2021). Designing a Low-Power LNA and Filter for Portable EEG Acquisition Applications. *IEEE Access, 9, 71968-71978*, (pp. 71968-71978).
7. Rao, P., & Mothiki, S. (2006). Pulse and Digital Circuits. In P. Rao, *Pulse and Digital Circuits* (pp. 69-96). New Delhi: Tata McGraw-Hill Education.
8. Zhang, L., Guo, X. J., Wu, X. P., & Zhou, B. Y. (2013). Low-cost circuit design of EEG signal acquisition for the brain-computer interface system. *6th International Conference on Biomedical Engineering and Informatics* (pp. 245-250). IEEE.

Prediction Diabetes using Machine Learning Techniques

Şeyma Kızıltaş [1], Dr. Mustafa Yeniad [2]

[1] Computer Engineering, Ankara Yıldırım Beyazıt University, Turkey, kiziltasseyma@gmail.com
[2] Computer Engineering, Ankara Yıldırım Beyazıt University, Turkey, myeniad@ybu.edu.tr

Abstract

Technology is continuously developing to improve the way people live and technology used in healthcare industry is also changing quickly. For instance, different technological devices are used for the diagnosis and treatment of diseases. It has been revealed that diagnosis of disease can be made by computer systems with developing technology. Machine learning algorithms are used in the field of health, as like as sales and marketing, logistics and finance areas, due to their high performance. The aim of this study is to explore different types of machine learning classification algorithms that can be used in diabetes diagnosis and to show their comparative analysis. In the study, seven classification algorithms were used in the literature. These algorithms are Logistic Regression, K-Nearest Neighbor, Multilayer Perceptron, Random Forest, Decision Trees, Support Vector Machine and Naive Bayes. Firstly, classification performance of algorithms are compared. These comparisons are based on accuracy, sensitivity, precision, and F1-score.

Keywords: Artificial neural network, Decision Tree, Diabetes Disease, Machine learning, Naive Bayes

1. Introduction

Diabetes mellitus is generally well-known as diabetes. Diabetes mellitus occurs when the pancreas cannot produce enough insulin or the one it produces cannot be used effectively. Insulin provides that sugar is stored as glycogen in the cell. Diabetics cannot use the glucose that passes from food to the blood, so blood sugar level rises. This event is called hyperglycemia. This situation causes damage to many tissues and organs in the long term such as eyes, kidneys, nerves, heart and blood vessels. Diabetes can be classified into three different types such as diabetes 1, diabetes 2, and gestation diabetes. Diabetes is a major health issue that has reached alarming levels in the world. Recent researches by World Health Organization (WHO) showed a great increase in number of diabetic patients and the deaths that are attributed to diabetes each year. In 2014, 8.5% of adults that aged 18 years and older had diabetes. In addition, according to the International Diabetic Federation (IDF) Diabetes Atlas 9th Edition 2019, 578 million people will have diabetes in 2030. Diabetes is a chronical disease. Therefore, it is significant that diabetes might be uncover at an early stage.

Machine learning is a branch of artificial intelligence (AI) and computer science which provides systems with the ability to learn and improve from its own experience. Machine learning methods are used in the health sector because the number of data is very large and the analysis takes time. Researches have shown that machine learning techniques can be used for diabetes prediction such as [1]. In this study, machine learning classification algorithms which are Naive Bayes (NB), Logistic Regression (LR), Multi-Layer Perceptron (MLP), Support Vector Machine (SVM), K-Nearest Neighbor (K-NN), Decision Tree (DT) and Random Forrest (RF) were implemented. The dataset is Pima Indians Diabetes Data (PIDD).

2. Literature Review

There are many studies for diabetes in the field of machine learning. Because diabetes is a chronical disease and should be diagnosed at an early stage. Various machine learning-based methods have been proposed on diabetes disease recognition.

Deepti Sisodia and Dilip Sisodia applied several machine learning algorithms to diagnose diabetes disease [2]. Naïve Bayes (NB), Decision Tree (DT) and Support Vector Machine (SVM) methods were used by the authors. Pima Indian Diabetes Dataset (PIDD) was used and the performances of the classifications were tested by the WEKA tool. The highest accuracy was obtained by the Naive Bayes which was 76.30%.

Islam and Jahan applied various machine learning methods which can be used in diabetes prediction [3]. Naïve Bayes (NB), Logistic Regression (LR), Multilayer Perceptron(MLP), Support Vector Machine (SVM), Decision Tree(DT), Random Forest (RF), AdaBoost, One Rule, K-Nearest Neighbor (K-NN) methods were used by the authors. Pima Indian Diabetes Dataset (PIDD) was used. The highest accuracy was obtained by the Logistic Regression which was 78.01%.

Kaur and Kumari also applied five machine learning algorithms to predict diabetes mellitus [4]. These algorithms are Radial Basis Function (RBF), Multifactor Dimensionality Reduction (MDR), K-Nearest Neighbor (kNN), Kernel Support Vector Machine (SVM) and Artificial Neural Network (ANN). Pima Indian Diabetes Dataset (PIDD) was used and highest accuracy was obtained by the Kernel Support Vector Machine (SVM) which was 89%.

Saru and Subashree analyzed machine learning techniques to predict diabetes [5]. Pima Indian Diabetes Dataset (PIDD) was used as the dataset. Logistic regression with Support Vector Machine (SVM), Decision Tree (DT), K-NN (k=1) and k-NN(k=3) are the classifiers. The highest accuracy was obtained by the Decision Tree which was 94.4%.

Kumari and Chitra applied Support Vector Machine (SVM) to diagnosis of diabetes [6]. Also Pima Indian Diabetes Dataset (PIDD) was used as the dataset. The accuracy was recorded as 78%.

Al-Zebari and Sengur compared machine learning algorithms for diabetes disease detection [7]. Decision Trees (DT), Logistic Regression (LR), Discriminant Analysis (DA), Support Vector Machines (SVM), k-Nearest Neighbors (k-NN) and ensemble learners techniques was used in the work. In this study, Pima Indian Diabetes Dataset (PIDD) was used and the Logistic Regression method had the best accuracy with 77.9%.

Table 1. *Comparison of the best models of the papers.*

Paper	Best Model	Accuracy
Sisodia and Sisodia [2]	Naive Bayes	76.30%
Islam and Jahan [3]	Logistic Regression	78.01%
Kaur and Kumari [4]	Kernel Support Vector Machine	89%

Saru and Subashree [5]	Decision Tree	94.4%
Kumari and Chitra [6]	Support Vector Machine	78%
Al-Zebari and Sengur [7]	Logistic Regressions	77.9%

Table 1. shows the best model and accuracy in the studies described above.

3. Materials and Methods

3.1. Data

The dataset used in the study is Pima Indian Diabetes (PIDD). The dataset can be found on the Kaggle website [8]. The original owner of this dataset is the National Institute of Diabetes and Digestive and Kidney Diseases. PIDD has been gathered among the Pima Indian female population aged at least 21 years near Phoenix, Arizona. This dataset contains 768 samples with 9 attributes. This dataset has 8 specific variables. The description of the data is given below in Table 2. [3]

Table 2. *Attribute's name and their types.*

Attribute	Description	Type
Pregnancies	Number of times pregnant	Numeric
Plasma-Glucose	Plasma glucose concentration a 2 hours in an oral glucose tolerance test	Numeric
BloodPressure	Diastolic blood pressure (mm Hg)	Numeric
SkinThickness	Triceps skin fold thickness (mm)	Numeric
Insulin	2-Hour serum insulin (mu U/ml)	Numeric
BMI	Body mass index (weight in kg/(height in m)^2)	Numeric
DiabetesPedigreeFunction	Diabetes pedigree function	Numeric
Age	Age of the patient (years)	Numeric
Outcome	Class variable (0 or 1)	Nominal

There are 500 instances of class 0 and 268 class 1. Waikato Environment for Knowledge Analysis (WEKA) tool has been used to categorize the data in this paper. WEKA is developed at University of Waikato. WEKA version 3.8 was used in this study.

3.2. Data preprocessing

Data may not always be complete and there may be abnormal values, impossible data combinations, missing values, duplicate data in the data set. Data preprocessing is required task to clean data and increase the accuracy and effectiveness of a machine learning model.

When the Pima Indian Diabetes dataset was analyzed, it was found that many attributes had impossible values with 0's. The numbers of missing values for each attribute are as follows:

- Pregnancies: 110
- Glucose: 5
- BloodPressure: 35
- SkinThickness: 227
- Insulin: 374
- BMI: 11

To eliminate missing values in this study, zero values for the pregnant attribute were left assuming they were real values and 234 samples that have at least two impossible value for attribute of the glucose, blood pressure, skin thickness, insulin and bmi were removed. After the pre-processing 534 instances are remain out of 768.

3.3. Performance Evaluation

While classifying the data in this study, 10-fold cross validation was applied as a test option. Cross-validation is preferred for overfitting problem and small datasets. In 10-fold cross validation, the data file is divided into ten and nine parts are used for training and one part for testing, this process is repeated ten times. In this study, WEKA tool was used and results of algorithms were compared. To compare the results, the number of values which are true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) were used in the confusion matrix and applying the following equations with these numbers, accuracy, sensitivity, precision and F1-score ratios were calculated.

- Accuracy: Shows ratio of correctly classified samples to the total number of tested samples. It is defined as in Equation 1.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

- Sensitivity (Recall): Shows ratio of positive classification of instances i.e. TP to the sum of TP and FN as shown in Equation 2.

$$Sensitivity = \frac{TP}{TP+FN} \quad (2)$$

- Precision: Shows ratio of positive sample that were correctly classified to the total number of positive predicted samples as shown Equation 3.

$$Precision = \frac{TP}{TP+FP} \quad (3)$$

- F1-score: It is a way of combining the precision and recall of the model as shown in Equation 4.

$$F1 - score = \frac{2 \times Precision \times Recall}{Precision+Recall} \quad (4)$$

4. Classification Outcomes

The results of the different classification methods tested with the WEKA tool are shown in Table 3. SVM had the highest accuracy (78.65%) whereas K-NN obtained lowest accuracy (71.16%). Total accuracy is above 71% in all cases. The second highest accuracy (77.71%) obtained from LR. Sensitivity and precision are quite gladsome. Also, RF acquired third highest accuracy (76.77%).

Table 3. *Result of algorithms.*

Algorithm	Accuracy (%)	Sensitivity (%)	Precision (%)	F1-score (%)
Naïve Bayes	75.65	82.35	81.44	81.89
Support Vector Machine	78.65	89.91	80.45	84.91
Decision Tree	74.71	78.99	82.45	80.68
Logistic Regression	77.71	89.07	79.89	84.23
Random Forest	76.77	85.99	80.57	83.19
K-Nearest Neighbor	71.16	80.39	77.35	78.84
Multilayer Perceptron	74.71	82.35	80.32	81.32

5. Conclusion

Millions of people around the world suffer from diabetes. However most of these people don't even know if they have the disease. Early diagnosis of diabetes can abate long-term complications and cost. Therefore, multiple machine learning algorithms applied and analyzed for PIDD. The results show that the best performance was produced by an SVM algorithm. Generally, all techniques produced an accuracy score of around 70%. Further analysis of attributes and different combination of feature selection is necessary to achieve higher accuracy. Also, much more datasets can be generated, real datasets can be taken or deep neutral networks can be applied to consider the real impact of the performance of the algorithms.

References

1. Q. Zou, K. Qu, Y. Luo, D.Yin, Y. Ju and H. Tang, “Predicting diabetes mellitus with machine learning techniques,” Frontiers in Genetics, vol. 9, Nov., 2018.
2. D. Sisodia and D. S. Sisodia, “Prediction of diabetes using classification algorithms”, Procedia Computer Science, vol. 132, pp. 1578-1585, 2018.
3. M. A. Islam, and N. Jahan, “Prediction of onset diabetes using machine learning techniques,” International Journal of Computer Applications, vol. 180, no. 5, Dec., pp.7-11, 2017.
4. H. Kaur, and V. Kumari, “Predictive modelling and analytics for diabetes using a machine learning approach,” Applied computing and informatics, 2020.
5. S. Saru, and S.Subashree, “Analysis and prediction of diabetes using machine learning,” International Journal of Emerging Technology and Innovative Engineering, vol. 5, no. 4, Apr., 2019.
6. V. A. Kumari and R.Chitra, “ Classification of diabetes disease using support vector machine,” International Journal of Engineering Research and Applications, vol. 3, no. 2, pp. 1797-1801, 2013.
7. A. Al-Zebari, and A.Sengur, “Performance comparison of machine learning techniques on diabetes disease detection,” In Proc. IEE International Informatics and Software Engineering Conference ’11, 2019, pp. 1-4.
8. “Pima Indians Diabetes Database,” Kaggle, Oct. 6, 2016. [Online]. Available: https://www.kaggle.com/uciml/pima-indians-diabetes-database. [Accessed Oct. 10, 2020].

Regional Signal Recognition of Body Sounds

Osman Balli[1], Yakup Kutlu[2]

[1,2] Department of Computer Engineering, Iskenderun Technical University, Turkey
Emails: osman.balli50@gmail.com, yakup.kutlu@iste.edu.tr

Abstract

Many diseases can be diagnosed from body sounds. But the perception of different sounds when recording these sounds or resting by doctors makes it difficult to diagnose the disease from body sounds. In addition to isolating these sounds from the external environment, it is necessary to separate the sounds coming from different regions within the body during the analysis. Separation of heart, lung and abdominal sounds makes it easier to analyze them, especially in digital analysis. In this study, a data was created with the lungs, heart and abdomen. The coefficients of MFCC was taken to the data. The coefficients obtained were trained in the CNN model. The aim is to determine which region of the sounds. When looking at the results, the educational achievement is around 98% and the test achievement is around 85%.

Keyword(s): Body sounds, MFCC, Deep Learning

1. Introduction

Auscultation is the process of listening to intra-body sounds with the help of a stethoscope for the diagnosis of a disease, for checking the functioning of organs. Auscultation is a procedure that provides basic information about the condition of the body due to an ailment or during a routine examination. The sounds of the heart, lungs, abdomen and other organs are tested in a kind of resting way. As a result of listening, a general knowledge of the person's condition is obtained.

We can listen to many sounds from the body. But although we know a little about which region they belong to, we may not be able to understand how much other sounds in the body interfere with the sounds we listen to. On the other hand, it is more convenient and advantageous to understand and distinguish body sounds in a digital environment than to diagnose with a stethoscope.

There are many studies on body sounds in the literature. In the study conducted on the diagnosis of COPD using lung sounds, the second degree difference plot (SODP) analysis method was used. Deep Belief Networks (DBNs) are combined with this method. The binding of 3D-SODP quantification properties together with DBN separated lung sounds from different COPD levels with high classification performance rates of 95.84%, 95.84%, 93.34% and 93.65% for accuracy, sensitivity and specificity, respectively. (Gokhan Altan vd., 2018). Another study focused on the most common lung sounds, wheeze and crackle. The data was collected with a custom mobile phone application and an electronic stethoscope. 284 Data received from patients with was detected ROC curves with AUCs of 0.86 for wheeze and 0.74 for crackle. According to the data taken from 284 patients, 0.86% of the results were obtained. It has been shown that semi-supervised deep learning can be successful in large datasets. (Chamberlain vd., 2016). Aykanat et al., on the other hand, have shown how different classifications in lung sounds will affect the results. There are 2 types of methods used; mel

frequency spectral coefficient (MFCC) features in a support vector machine (SVM) and spectrogram images in the convolutional neural network (CNN). He achieved the best result in the training of healthy and pathologically classified voices. He achieved 86% success in SVM and 86% success in CNN. (Aykanat vd., 2017). Altan et al. analyzed the severity of COPD from 12 channel lung sounds. The Deep ELM model was used and has separated five COPD severities with classification performance rates of 94.31%, 94.28%, 98.76%, and 0.9659 for overall accuracy, weighted-sensitivity, weighted-specificity, and area under the curve (AUC) value, respectively (Gökhan Altan vd., 2020). In another study, 100 diseased and 100 normal heart sound data were used. The Time-Frequency Distribution (TFD) analysis of these data and the heart sound analysis based on Mel Frequency Spectrum Coefficient (MFCC) were performed. The coefficients obtained from these methods are trained in artificial neural networks. The system is able to produce the accuracy up to 90.0% using the TFD and 80.0% using the MFCC (Kamarulafizam vd., 2007). In another study on heart sounds, it was found that MFCC was applied to 1381 data sets of real and simulated, normal and abnormal domains. The classification rates for normal and abnormal heart sounds were found to be 95.7% for continuous murmurs, 96.25% for systolic murmurs, and 90% for diastolic murmurs using a probabilistic comparison approach (Chauhan vd., 2008).

In this study, a dataset consisting of sounds taken from the heart, lungs and abdomen was prepared. The aim of the study is to develop a model that can decompose sounds. MFCC was first applied to audio data. The coefficients obtained as a result of Mfcc were trained in the CNN model.

2. Materials and Methodss

2.1. Database

In this study, audio recordings were taken from 12 people. A total of 12 records were obtained from 1 person, including four from different regions of the lungs, heart and abdomen. The recordings are 20s and have a sampling frequency of 4000. The sounds were divided into 2s and a total of 10 audio data were obtained from 1 audio recording.

2.2. Mel Frekans Cepstrum Katsayısına (MFCC)

MFCC serves to create attributes using audio signals. There are many studies in which attribute extraction is performed using MFCC in audio signal (Palaniappan vd., 2014)(Nogueira vd., 2019)(Kamarulafizam vd., 2007). It is very important to create meaningful attributes of body sounds for educational work to be done. In this study, Mel Frequency Cepstrum Coefficients were used. The processes used in MFCC are shown in Figure 2.

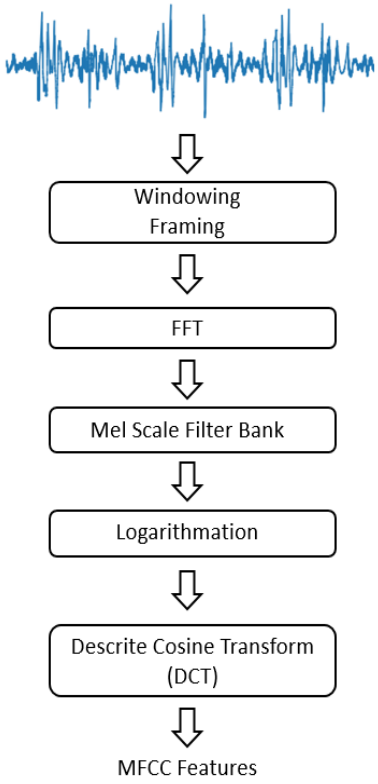


Figure 1. Block Diagram of MFCC

MFCC applies the window size to the signal first. Each window implements a fft and allows it to pass to the frakans domain. Then convert it from the frequency scale to the Mel logarithmic scale with the Mel filter bank. Logarithmic applies DCT to the data obtained as a result of the conversion, and the results obtained are converted to a time scale. As a result of these operations, 1287 attributes were obtained from every 2-second signal. The obtained data were trained in the CNN model.

2.3. Convolution Neural Network (CNN)

Deep learning is a branch of artificial intelligence that has started to be used frequently with the spread of high-performance hardware. One of the most popular methods of deep learning is CNN. It is a method used in different fields such as signal processing, classification, recognition.

The data provided to CNN must be in matrix format. It consists of layers called convolution and pooling. The number of these layers is determined by the person whose method is to be applied. The Convolution layer is the layer where the number and size of filters to be applied to the data are determined. The results obtained from this layer are again given to the convolution layer or the pooling layer. Different methods such as max, min, average can be used in the pooling layer. After the pooling process is applied to the data coming to the pooling layer, the number of data decreases. According to the method we have determined, it moves meaningful data to the next layer. After a certain number of convolution and pooling operations are performed, the data is moved to the flatten layer. Here the results are transformed into a

one-dimensional matrix. The matrix is transferred to artificial neural networks. Weights in training are constantly updated. The highest performance gives the optimal weight values.

In this study, the data were converted into a 35x35 square matrix. In the study, 64-, 128- and 256-layer Convolution layers were applied. In addition, 3x3 average pooling was performed 2 times. The general scheme of CNN is shown in Figure 2.

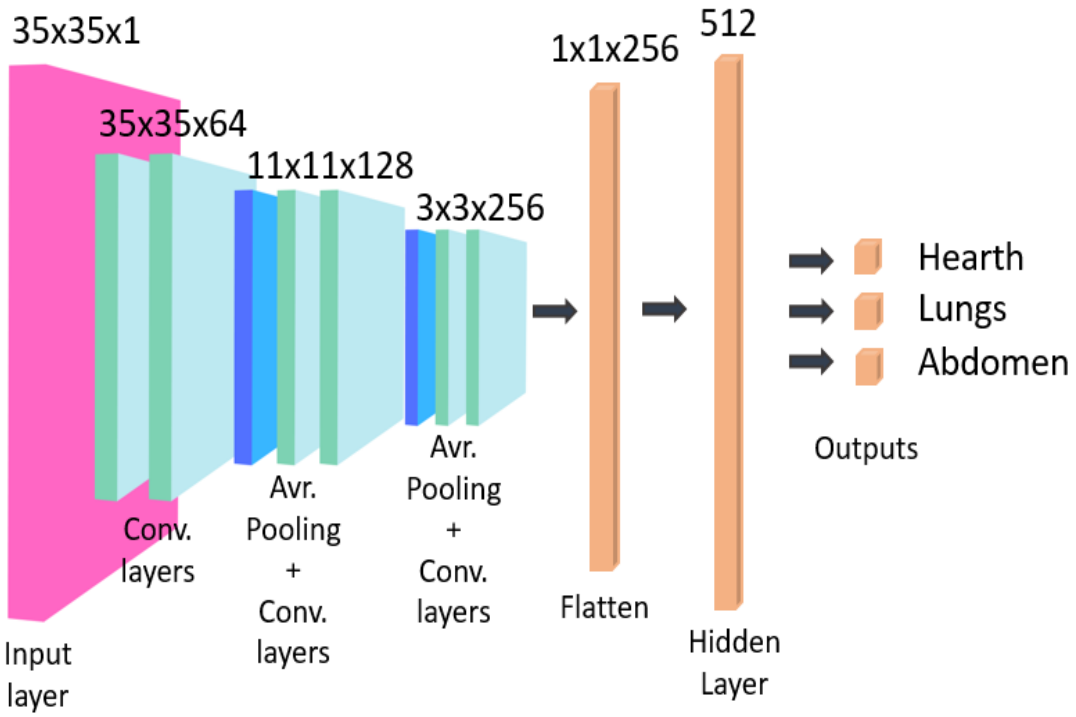


Figure 2. Shema of Convolution Neural Network

3. Results and Discussion

Obtaining useful results in the analysis of body sounds will pave the way for early diagnosis in many diseases. In addition, the creation of ready-made models is ethical in the faster operation of diagnostic systems. It is very important to get high performance values in these models. We focus on 2 achievements: training and test success. As a result of this study, these 2 parameters were examined and analyzed.

As a result of CNN training with attributes obtained by applying MFCC to audio data, train and test achievements were examined. a model was prepared that used 1440 voice data consisting of 2s and determined which of these sounds belong to the heart, lung and abdominal regions. A processing time of 130 epoch was determined in the training. It is seen that when the educational achievement reaches around 98%, the test success is around 85%.

It is seen that regional recognition has been successful with MFCC and CNN. But in order for the results to be even more reliable, the number of data must be increased. In addition, increasing the test performance will be effective in making the model work more accurately.

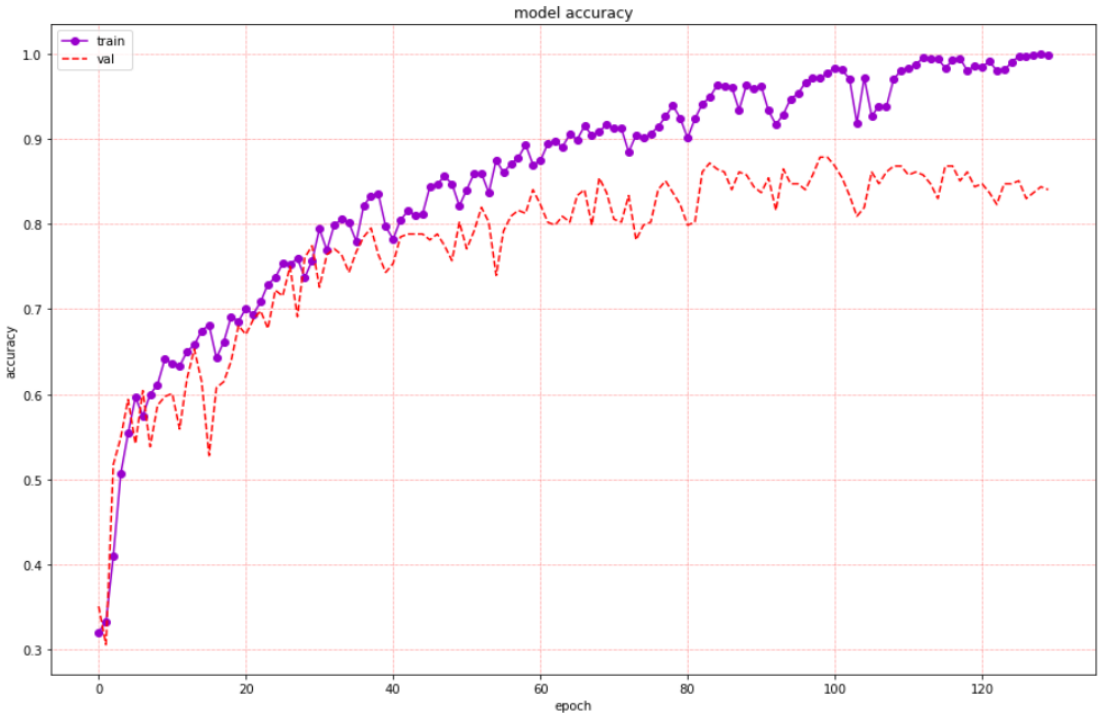


Figure 3. Result of Model Accuracy

References

1. Altan, Gökhan, Kutlu, Y., Gökçen, A. 2020. "Chronic obstructive pulmonary disease severity analysis using deep learning on multi-channel lung sounds". Turkish Journal of Electrical Engineering and Computer Sciences, *28*(5), 2979–2996.
2. Altan, Gokhan, Kutlu, Y., Pekmezci, A. Ö., Nural, S. 2018. "Deep learning with 3D-second order difference plot on respiratory sounds". Biomedical Signal Processing and Control, *45*, 58–69.
3. Aykanat, M., Kılıç, Ö., Kurt, B., Saryal, S. 2017. "Classification of lung sounds using convolutional neural networks". Eurasip Journal on Image and Video Processing, *2017*(1).
4. Chamberlain, D., Kodgule, R., Ganelin, D., Miglani, V., Fletcher, R. R. 2016. "Application of semi-supervised deep learning to lung sound analysis". Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS, *2016-October*, 804–807.
5. Chauhan, S., Wang, P., Sing Lim, C., Anantharaman, V. 2008. "A computer-aided MFCC-based HMM system for automatic auscultation". Computers in Biology and Medicine, *38*(2), 221–233.
6. Kamarulafizam, I., Salleh, S. H., Najeb, J. M., Ariff, A. K., Chowdhury, A. 2007. "Heart sound analysis using MFCC and time frequency distribution". IFMBE Proceedings, *14*(1), 946–949.
7. Nogueira, D. M., Ferreira, C. A., Gomes, E. F., Jorge, A. M. 2019. "Classifying Heart Sounds Using Images of Motifs, MFCC and Temporal Features". Journal of Medical Systems, *43*(6), 186–203.
8. Palaniappan, R., Sundaraj, K., Sundaraj, S. 2014. "A comparative study of the svm and k-nn machine learning algorithms for the diagnosis of respiratory pathologies using pulmonary acoustic signals". BMC Bioinformatics, *15*(1), 1–8.

Scene Construction from Depth Map Using Image-to-Image Translation Model

Hasan Avşar[1], Mehmet Sarıgül[2], Levent Karacan[3]

[1,2,3] Department of Computer Engineering, Iskenderun Technical University, Turkey
Emails: hasanavsar.mfbe20@iste.edu.tr, mehmet.sarigul@iste.edu.tr, levent.karacan@iste.edu.tr

Abstract

In recent years, deep learning approach to solve the image and video processing problems have become very popular. Generative Adversarial Networks (GANs) are one of the most popular deep learning-based models. GANs form a generative model utilizing two sub-models, namely, generator and discriminator. The generator tries to generate indistinguishably realistic outputs where the discriminator tires to classify the outputs of the generator as real or fake. These two models work together to achieve a successful generation of realistic outputs. This study aims to reconstruct daytime image of a given depth map data recorded with a camera or a sensor which can capture the depth map data during night time or in a lightless environment. Our model was used for reconstructing the 2D images for a given depth map representation of a known scene. The model was trained with the chess scene from 7-scenes dataset and realistic 2D images were successfully generated for the given input maps.

Keyword(s): Deep Learning, Generative Adversarial Networks, Pix2pixHD

1. Introduction

Photo-realistic image rendering for a specific scene or an object problem and there are many studies focusing on it. Representing a scene in a model learned from a data set makes it possible to generate successful outputs. The scene can be modelled with different kind of data inputs. Camera location and orientation, point clouds, scene coordinate maps and depth maps can be used along with the corresponding real RGB images to represent a scene. After a successful training, photo-realistic images of a scene can be constructed for desired inputs. Devices such as a LIDAR camera or a Kinect sensor can capture depth map of a scene even in the lightless environment. Using these kinds of devices, captured depth map data in a dark place or during the night time can be converted to daytime images. This conversion problem can be solved with the use of deep learning (Schmidhuber, 2015; LeCun et al., 2015) algorithms. Deep neural networks are solid structures and also are popularly used in different problem domains. This popularity enabled the development of many different variations of deep learning algorithms. Generative Adversarial Networks (GANs) (Goodfellow et al., 2014) are a deep learning based generative algorithm popularly used for applications such as image-to-image conversion, style transfer, scene construction. GANs are basically algorithms that utilize two sub-networks, namely generator and discriminator. These networks challenge each other to update and upgrade the generated outputs of the main algorithm. The generator generates outputs and the discriminator classifies whether the output is real or fake. This challenge continues until the generated results are indistinguishable from the real-life data.

Pix2pixHD (Ting-Chun et al., 2018) is a type of GAN which performs image-to-image transformation. It aims to solve the generated image quality problem at higher resolution images with the use of multi-scale generator and discriminator architectures. Pix2pixHD is inspired

from a previous study named Pix2pix (P. Isola et al., 2017) and with its multi-scale approach it can generate better results. In this study, Pix2pixHD was used for modelling a scene with depth maps and corresponding real images. Then de photo-realistic images were successfully generated for unseen depth map data inputs gathered from the same scene.

2. Materials and Methods

2.1. Generative Adversarial Networks

Generative Adversarial Networks, abbreviated as GANs, are basically a pair of models competing with each other in order to be able to generate outputs which are indistinguishable from the real ones. The term "adversarial", comes from this nature of models in competition. Starting with randomized input data, the generator model generates fake images which become more and more realistic as the training continues. The discriminator model is fed with the fake images along with the real images from training dataset and makes a binary prediction whether the generated image is fake or real. At the end of each competition round, both generator and discriminator models are updated based on the success of generated (fake) images. The basic flowchart of a GAN is given in Figure 1.

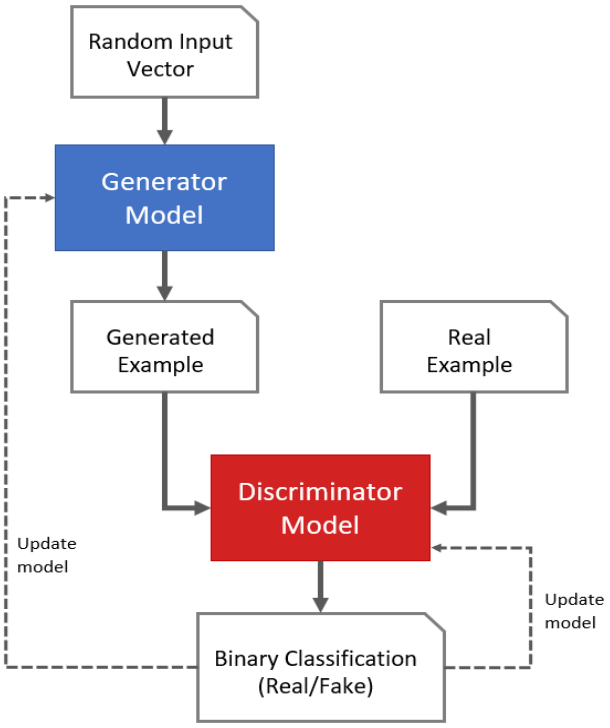


Figure 1. Generative Adversarial Network Structure

2.2. Pix2PixHD

Pix2pixHD is an extended GAN model which bring new ideas to the basic structure in order to overcome the problems arising as the resolution of generated images increase. It includes multi-scale generator and discriminator architectures which makes it possible to generate realistic images at 2048 x 1024 resolution. A generator network G1 is trained on lower resolution (coarse) and then it is appended to another network G2 to be jointly trained for higher

resolution (fine) scale. The 2-scale structure of generator model G = {G1, G2} is given in Figure 2. As the target resolution increases, the generator can be modified to have more scale levels (G = {G1, …, Gn}).

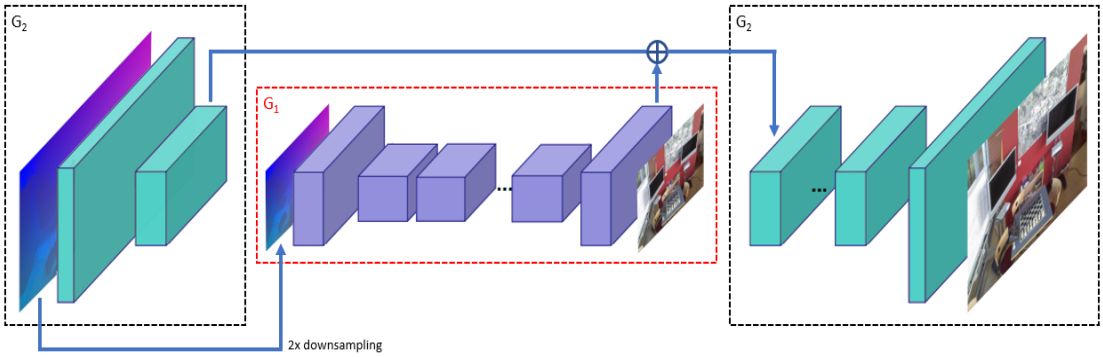


Figure 2. Multi-Scale Generator Model Structure of Pix2pixHD

Various number of discriminator networks are utilized to handle the discrimination task on different scales of given input images. The images are down sampled and are fed to into discriminator networks to be processed in a coarse-to-fine manner. The result is then used to update individual discriminator networks and the generator network.

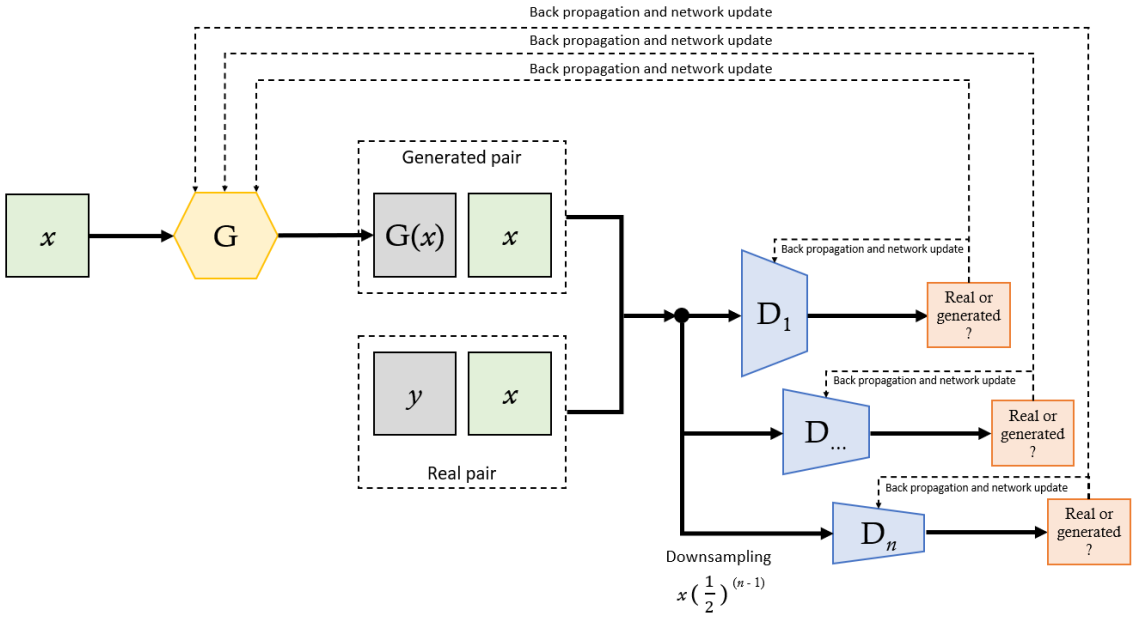


Figure 3. Multi-Scale Discriminator Model Structure of Pix2pixHD

2.3. 7-Scenes Dataset

Chess scene of 7-Scenes dataset was used for this study. The training dataset includes 4000 consecutive frames of a video stream where the test data set includes 2000 frames.

2.4. Experiment

The Pix2pixHD was structured with 4 levels of generator models and 2 layers of discriminator models. Training was performed with 44 epochs.

3. Results and Discussion

Using the Pix2pixHD model formed after training, realistic images are generated successfully. Figure 4 shows the comparison of generated images with the real images. Figure 5 shows the loss trends achieved during the training.

Images generated by Pix2pixHD | Real Images

Figure 4. Comparison of generated images with real images

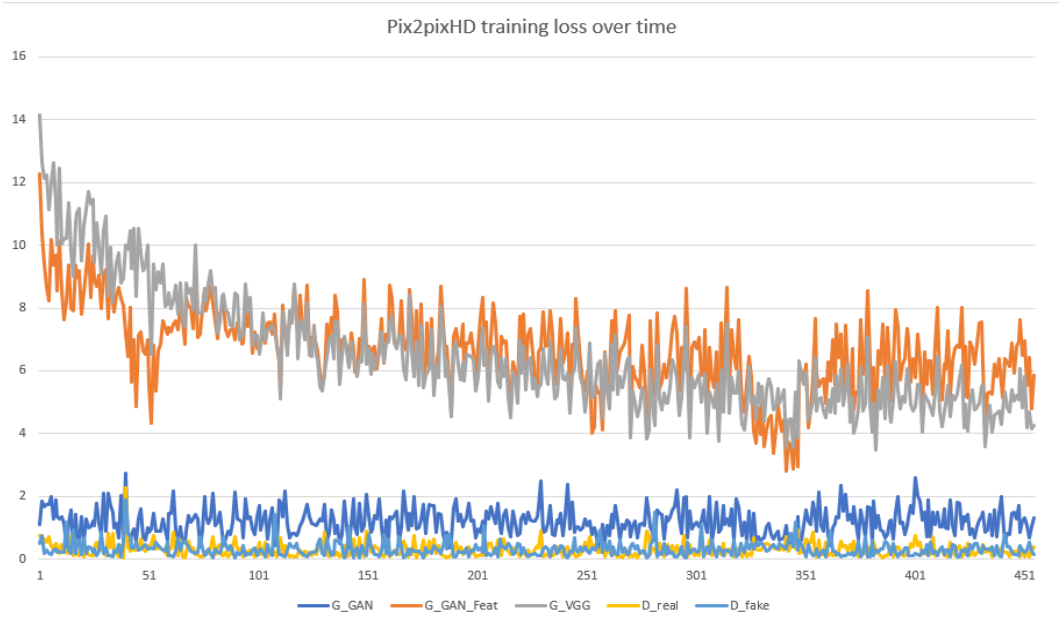


Figure 5. Training loss trends

4. Conclusion

GANs are one of the most preferred deep learning-based algorithms used for image-to-image conversion. Pix2PixHD is a type of GAN which targets to increase the generated output efficiency even for the images with high resolution. Pix2PixHD uses multiscale network structures enabling coarse-to-fine approach to the realistic image generation problem. In this study, Pix2pixHD was used to generate realistic daytime images of a given depth map from a known scene. Although there are imperfections in the generated images, the realistic images were successfully generated by the model.

The success of this study shows that the depth map data recorded by a camera or a sensor can be converted to daytime images. LIDAR cameras, Kinect sensors or other similar devices can capture depth map data in lightless environment. The model in this study can be used to generate daytime images of a scene during the night time or when the scene is in the complete darkness. On the other hand, the model can also be further developed for generating daytime video of a scene just by converting the depth map data captured in the darkness.

5. Acknowledgments

This study was supported by project number 120E447 from the The Scientic and Technological Research Council of Turkey (TUBITAK).

References

1. Schmidhuber, J. (2015). Deep learning in neural networks: an overview. Neural networks, 61, 85-117.
2. LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. Nature 521, 436–444 (2015).
3. Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., & Bengio, Y. (2014). Generative adversarial nets. In Advances in neural information processing systems, 2672-2680.
4. Ting-Chun Wang, Ming-Yu Liu, Jun-Yan Zhu, Andrew Tao, Jan Kautz, and Bryan Catanzaro. "High-Resolution Image Synthesis and Semantic Manipulation with Conditional GANs", in CVPR, 2018.

5. P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros. Imageto-image translation with conditional adversarial networks. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.
6. Jeong, Hyun-Jin & Moon, Yong-Jae & Park, Eunsu & Lee, Harim. (2020). Solar Coronal Magnetic Field Extrapolation from Synchronic Data with AI-generated Farside. The Astrophysical Journal. 903. 10.3847/2041-8213/abc255.

KEYWORDS

U

V

W

Y