VIGYAAN'21-22

7th NATIONAL CONFERENCE

ON

EMERGING TRENDS IN CLOUD COMPUTING

Saturday, 19 February 2022

PROCEEDINGS

Naipunnya
To reach the unreachable

Organised by :

Department of Computer Science

NAIPUNNYA INSTITUTE OF MANAGEMENT AND INFORMATION TECHNOLOGY (NIMIT)

VIGYAAN 2021-2022

Volume 3, Issue 3

Proceedings of Seventh National Conference on Emerging Trends in Cloud Computing

Vigyaan 2021-2022

The Conference Proceedings-*"Emerging Trends in Cloud Computing"*

Manager

Fr.(Dr).Paul Kaithottungal
Executive Director & Principal
Naipunnya Institute of Management and Information Technology

Editor

Ms. Teresa Parackal
Vice Principal ,Acadamics
Naipunnya Institute of Management and Information Technology

Mr.Robert Fernandez
Vice Principal, Campus Management
Naipunnya Institute of Management and Information Technology

Editorial Advisory Council

Mr. Jayakrishnan S
Head of the Department
Department of Computer Science
Naipunnya Institute of Management and Information Technology

Editorial Board

Mr.Deepak K V
Assistant Professor
Department of Computer Science
Naipunnya Institute of Management and Information Technology
Ms. Siji Jose
Assistant Professor
Department of Computer Science
Naipunnya Institute of Management and Information Technology
Ms. Nithya paul
Assistant Professor
Department of Computer Science
Naipunnya Institute of Management and Information Technology
Ms.Joicy Joy
Assistant Professor
Department of Computer Science
Naipunnya Institute of Management and Information Technology

Editorial and Administrative Office
Naipunnya Institute of Management and Information Technology
Pongam,Korraty East,Trissur,Kerala-680 308,Ph:0480 2730340,2730341 Web: www.naipunnya.ac.in,Email:mail@ naipunnya.ac.in

Publisher:
Naipunnya Institute of Management and Information Technology Pongam, Korraty East, Trissur,Kerala 680308, Ph:0480 2730340,2730341,Web: www.naipunnya.ac.in,Email:mail@ naipunnya.ac.in

ISBN 978-81-949892-5-7

All Rights Reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission of the publisher.

Foreword

We are here with the third edition of "VIGYAAN – 2021-22". The theme for the edition is "Emerging Trends in Cloud Computing". Cloud Computing is making a significant impact on Information Technology (IT). It is not merely a technology concept but also a new approach of implementing electronic commerce. An idea that computing is used similarly to the way of utility consumption, is to revolutionize the development and delivery of IT services.

Cloud Computing has recently been recognized as one of the most emerging technology. A considerable amount of research has been carried out to explore different areas in Cloud Computing. Few areas including reliability, security and business value of Cloud Computing are yet to be explored. We would like to presents a brief summary on the analysis of current gaps and new trends in cloud computing based on information systems literature, industry reports, and practical experience reflections. Moreover, it highlights the significance of cloud computing and its implications for practitioner and academics.

Research adds to the stock of knowledge and provides the source of new ideas, methods, techniques, and findings across a whole range of disciplinary and multi – disciplinary areas. 'VIGYAAN' has been focusing on addressing the developing areas of computer science. To document this intellectual vibrancy will always be the key aspect of VIGYAAN. We focus to promote knowledge and make the various academic developments in the world accessible to every section of society. This conference proceeding is a hub of diverse ideas and arguments in the advance areas related to computational intelligence. In this COVID pandemic stage, even in an online mode, we believe that 'VIGYAAN' is a significant step in achieving our aims and principle.

Editor–VIGYAAN 2021-22

Contents

Big Data Security Issues and Challenges

Sandra Jose
Assistant Professor
Department of Computer Science
Carmel College, Mala, Kerala, India
sandra@carmelcollegemala.ac.in

Abstract: In view of the wide application and popularization of large data, more and more data security and privacy issues have brought great challenges to the development of large data. Starting from the characteristics of big data, this paper analyses various risks of information security, and puts forward the corresponding development strategy of big data security. The results show that the combination of technology and relevant policies and regulations can better solve the problem of big data security and privacy protection.

I. INTRODUCTION

The term big data refers to the massive amount of digital information companies and governments collect about us and our surroundings. This data is not only generated by traditional information exchange and software use via desktop computers, mobile phones and so on, but also from the myriads of sensors of various types embedded in various environments, whether in city streets (cameras, microphones) or jet engines (temperature sensors), and the soon- to proliferate Internet of Things, where virtually every electrical device will connect to the Internet and produce data. Every day, we create 2.5 quintillion bytes of data--so much that 90% of the data in the world today has been created in the last two years alone (as of 2011 [1]). The issues of storing, computing, security and privacy, and analytics are all magnified by the velocity, volume, and variety of big data, such as large-scale cloud infrastructures, diversity of data sources and formats, streaming nature of data acquisition and high volume inter-cloud migration.

II. BIG DATA, BIG SECURITY CHALLENGES

Big data presents a tremendous opportunity for Enterprises across industries. By tapping into new volumes and varieties of data, scientists, executives, product managers, marketers, and a range of others can start making more informed plans and decisions, discover new opportunities for optimization, and deliver breakthrough innovations. Without the right security and encryption solution in place, however, big data can mean big problems. Securing big data comes with its own unique challenges beyond being a high-value target. It's not that big data security is fundamentally different from traditional data security. Big data security challenges arise because of incremental differences, not fundamental ones. The differences between big data environments and traditional data environments include:

- The data collected, aggregated, and analyzed.
- The infrastructure used to store and house big data
- The technologies applied to analyze structured and unstructured big data

A. The Data

The variety, velocity and volume of big data amplifies security management challenges that are addressed

In traditional security management. Big data repositories will likely include information deposited by various sources across the enterprise. This variety of data makes secure access management a challenge. Each data source will likely have its own access restrictions and security policies, making it difficult to balance appropriate security for all data sources with the need to aggregate and extract meaning from the data. For

ISBN 978-81-949892-5-7

example, a big data environment may include a dataset with proprietary research information, a dataset requiring regulatory compliance, and a separate dataset with personally identifiable information (PII). A researcher might want to correlate their research with a dataset including PII data, but what restrictions should be in-place to ensure adequate security? Protecting big data requires balancing analysis like this with security requirements on a case-by-case basis.

In addition, many of the repositories collect data at high volumes and velocity from a number of different data sources, and they all might have their own data transfer workflows. These connections to multiple repositories can increase the attack surface for an adversary. A big data system receiving feeds from 20 different data sources may present an attacker with 20 viable vectors to attempt to gain access to a cluster.

B. The Infrastructure

Another big data challenge is the distributed nature of big data environments. Compared with a single high-end database server, distributed environments are more complicated and vulnerable to attack. When big data environments are distributed geographically, physical security controls need to be standardized across all accessible locations. When data scientists across the organization want access to information, perimeter protection becomes important and complicated to ensure access to users while protecting the system from a possible attack. With a large number of servers, there is an increased possibility that the configuration of servers may not be consistent – and that certain systems may remain vulnerable.

C. The Technology

An additional big data security challenge is that big data programming tools, including Hadoop and NoSQL databases, were not originally designed with security in mind. For example, Hadoop originally didn't authenticate services or users, and didn't encrypt data that's transmitted between nodes in the environment. This creates vulnerabilities for authentication and network security.

identifiable information, payment card data, intellectual property, health records, and much more. Consequently, the data sources being compiled need to be secured in order to address security policies and compliance mandates.

some of the security features provided by traditional databases, such as role-based access control. The advantage of NoSQL is that it allows for the flexibility to include new data types on the fly, but defining security policies for this new data is not straightforward with these technologies.

The biggest challenge for big data from a security point of view is the protection of user's privacy. Big data frequently contains huge amounts of personal identifiable information and therefore privacy of users is a huge concern.

Because of the big amount of data stored, breaches affecting big data can have more devastating consequences than the data breaches we normally see in the press. This is because a big data security breach will potentially affect a much larger number of people, with consequences not only from a reputational point of view, but with enormous legal repercussions.

Fig 1. Big Data

ISBN 978-81-949892-5-7

I. THE MASSIVE SCOPE OF BIG DATA SECURITY

To establish comprehensive big data security, executives and administrators have to address the following areas:

A. Data sources:

To most fully exploit the advantages of big data, organizations leverage various forms of data, including both structured data in a range of heterogeneous applications and databases and unstructured data that comes in a number of file types. Organizations may leverage data from enterprise resource planning systems, customer relationship management platforms, video files, spreadsheets, social media feeds, and many other sources. Further, more data sources are added all the time. Today, you don't know where new data sources may come from tomorrow, but you can have some certainty that there will be more to contend with and more diversity to accommodate. These big data sources can include personally

B. Big data frameworks:

Within the big data environment itself—whether it's powered by Hadoop, MongoDB, NoSQL, Teradata, or another system—massive amounts of sensitive data may be managed at any given time. Sensitive assets don't just reside on big data nodes, but they can come in the form of system logs, configuration files, error logs, and more.

C. Analytics:

The ultimate fruit of a big data initiative is the output, the analytics that help the business optimize and innovate. This information can be presented in dashboards and reports, and accessed via on-demand queries. In some businesses, big data analytics represent the most sensitive asset of all, intelligence that provides a critical competitive differentiator—and a huge competitive exposure if it falls into the wrong hands.

It is important to recognize that the attributes that make big data valuable to the business also make it valuable to others—whether they're hardened cyber criminals or a disgruntled system administrator looking to make a quick, illicit buck. Establishing effective security across the categories above—and the massive number of specific outputs, systems, and services that fall into each category—is both critical and challenging.

Further, given the massive, widely fluctuating processing demands associated with big data environments, many organizations are leveraging cloud-based services and platforms to support their big data initiatives. For those organizations running big data environments in the cloud, the task of managing security grows even more difficult. In the cloud, security teams have to contend with the threats of vendor's infrastructure administrators, potential exposure to other tenants, and a number of other additional risks.

II. SECURING BIG DATA ENVIRONMENTS WITH VORMETRIC

After the text edit has been completed, the paper is ready for the template. Duplicate the template file by using the Save As command, and use the naming convention prescribed by your conference for the name of your paper. In this newly created file, highlight all of the contents and import your prepared text file. You are now ready to style your paper; use the scroll down window on the left of the MS Word Formatting toolbar. Vormetric solutions for big data security enable organizations to maximize the benefits of big data analytics— while maximizing the security of their sensitive data and addressing the requirements of their compliance office. The Vormetric Data Security Platform offers the granular controls, robust encryption, and comprehensive coverage that organizations need to secure sensitive data across their big data environments—including big data sources, big data infrastructure, and big data analytic results. By delivering a single security solution that offers coverage of these areas,

Vormetric enables security teams to leverage centralized controls that optimize efficiency and compliance adherence. The Vormetric Data Security Platform offers capabilities for big data encryption, key management,

ISBN 978-81-949892-5-7

and access control— featuring several product offerings that share a common, extensible infrastructure. Further, the solution generates security intelligence on data access by users, processes, and applications.

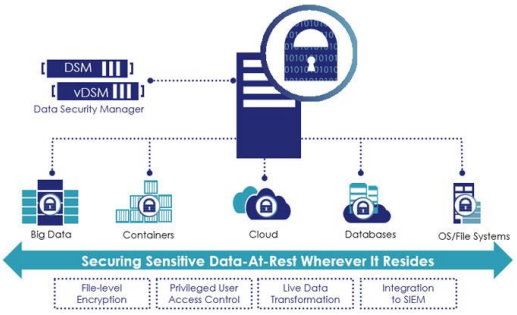


Fig 2. Big data Security

Vormetric Transparent Encryption efficiently protects data across all these areas, delivering encryption, privileged user access control, and security intelligence. In addition, with Vormetric Protection for Teradata Database, your organization can gain the comprehensive, granular controls required to secure the most sensitive assets across your Teradata environments, while enabling you to maximize the business benefits of your big data investments

Protecting Big Data Sources

As outlined earlier, organizations can leverage data from a broad array of sources, both structured and unstructured, for their big data initiatives. Data from databases, data warehouses, system logs, spreadsheets, and many other diverse systems may be fed into a big data environment.

To establish data security for these diverse data sources, organizations can use the following Vormetric solutions:

A.Vormetric Transparent Encryption:

This product encrypts and controls access at the file- system level. This encryption solution is easy to deploy because it doesn't require any changes to applications.

B. Vormetric Application Encryption:

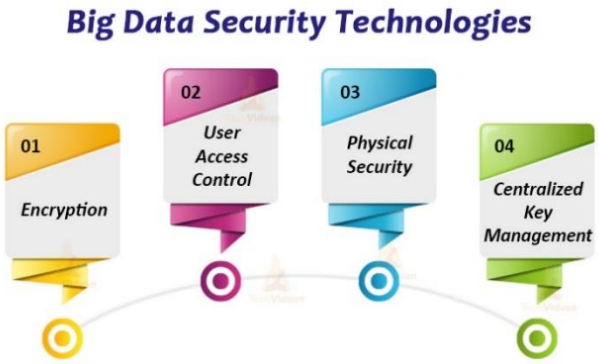


Fig 3. Big Data Security Technologies

With this encryption product, you can encrypt specific Columns in an application before it writes the field to a database. By encrypting a specific column, you can ensure a specific sensitive field will remain unreadable, even after it is imported into, and processed within, the big data environment.

ISBN 978-81-949892-5-7

III. SECURING BIG DATA FRAMEWORKS

In big data environments, data is routinely replicated and migrated among a large number of nodes. In addition, sensitive information can be stored in system logs, configuration files, disk caches, error logs, and so on..

IV. SAFEGUARDING BIG DATA ANALYTICS

Big data output comes in many forms, including on- demand dashboards, automated reports, and ad hoc queries. Very often, these outputs contain intellectual property that is very valuable to an organization—and a potential target of attack. To provide big data analytics security for these confidential assets, security teams can use the solutions Vormetric Transparent Encryption and Vormetric Application Encryption.

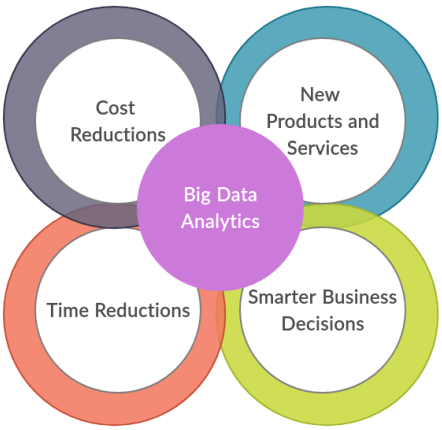


Fig 4. Big Data Analytics Techniques

V. CONCLUSION

This paper has highlighted the big data science, infrastructure related issues that have not been thoroughly vetted for security, the non-scalability real time monitoring techniques, etc. This clarifies attack surface of big data security. In the upcoming years, big data will have high impact on entire security spectrum which includes, anti- malware, data loss, network monitoring, user authentication & authorization, identity management, fraud detection, governance, risk and compliance. I hope that this paper will spur interest in R & D community to collaboratively focus on

REFERENCES

[1] Big Data Security Implications – http://www.academia.edu/3331728/ Big Data_-_Security_implications
[2] Top 10 Big Data Security and Privacy Challenge http://www.isaca.org/groups/professional-english/big- data/groupdocuments/big_data_top_ten_v1.pdf
[3] Big Data Security – Article in vormetric.com – http://www.vormetric.com/data-security-solutions/use-cases/big-data- security.

ISBN 978-81-949892-5-7

Fighting Against Covid 19: A Comparative Study of Deep Learning and Machine Learning Algorithms for Face Mask Detection

Soniya Davis
PG Student
Department of Computer Science
Sacred Heart College
Chalakudy
Kerala, India

Sindhu T
Assistant Professor
Department of Computer Science
Sacred Heart College
Chalakudy
Kerala, India

ABSTRACT:

The COVID-19 pandemic demands an efficient face mask detection application around the world. The number of cases of COVID-19 is still getting registered throughout the country. Wearing face masks in public areas clearly reduces the risk of coronavirus transmission. The introduced model of face mask detection is based on deep learning. The proposed system can be implemented with computer or laptop cameras allowing it to detect people who are wearing masks and not wearing masks. The introduced system includes a comparison between machine learning and deep learning algorithms to find the most suitable algorithm that yields the highest accuracy for face mask detection. In this project, certain classification algorithms such as SVM, Random Forest, and CNN for face mask detection are used. It is object detection and classification problem with two different classes (Mask and Without Mask). A dataset is used to build this face mask detector using Python, OpenCV, and Keras. Hence, this module will help to reduce the spreading rate of coronavirus by detecting the face masks.

KEYWORDS:Covid-19,MachineLearning, Deep Learning, CNN, OpenCV

1 INTRODUCTION

The COVID-19pandemic has had a lasting impact in many countries worldwide since December 2019. To reduce the spread of the disease, mandatory face mask rules are now becoming common in public areas around the world.Reports indicate that wearing facemasks in public areas will clearly reduce the risk of covid19. So, wearing a mask during this pandemic is a critical preventive measure and is the most vital step in times when social distancing is hard to maintain.

Face mask detection (FMD) refers to detecting whether a person is wearing a mask or not. This type of system is very helpful in many areas like Malls, Colleges, libraries, etc. The detection of face masks can be done with the images and also detect via live video. If people wear the mask, it will permit them else it will give a labelled image without a mask. Thus, this system will help to prevent people from Coronavirus transmission.

This research focuses on the real-time detection of face masks from videos using deep learning and it also focuses on the comparative study of deep learning and machine learning algorithms for face mask detection. The proposed framework uses a cascade classifier to detect the faces and their corresponding facial landmarks present in the video frame. These facial images and cues are then processed by a neural network to identify the face masks. The model is integration between deep learning and machine learning techniques with Open CV, Tensor flow,

ISBN 978-81-949892-5-7

and Keras. The goal is to identify whether the person in images/video streams is wearing a face mask or not with the help of deep learning and machine learning algorithms.

The algorithms such as SVM, Random Forest, and CNN are used for classification and detection. By this analysis, the accuracy of three algorithms can be determined and represented graphically.

II. PROPOSED SYSTEM

The research is based on the face mask detection of human beings to recognize whether a person is wearing a mask or not. To stop the spread of the virus governments are taking necessary steps. It is our duty to be responsible and help everyone to prevent the virus spread. So as a part of this, the proposed work is based on a face mask detection system that would might help in detecting whether the person is wearing a mask or not. Because it's a tedious and harmful job for any human to try do in this pandemic situation. Within the proposed system we present the application that might help to track the people that aren't wearing masks from moving objects in public areas to avoid the spreading of coronavirus during this pandemic time. The goal is to put in a security camera to spot who is with mask and without mask rather than a human doing that job, because any human is volatile to the virus. Here, the system detects the face mask of a gaggle of persons from videos. While entering the place everyone should scan their face, then enter ensuring they have a mask with them.

This research aims at the real time detection of face masks from moving objects using deep learning and it also focuses on the comparative study of deep learning and machine learning algorithms for face mask detection. For classification the algorithms such as SVM and Random Forest are used. By the classification process, the accuracy of these algorithms can be determined and can be represented graphically. The proposed model uses CNN which is an effective model to detect a face mask. The system provides better detection and classification technique.

III. METHODOLOGY

For classification and detection, the algorithms such as SVM, Random Forest, and CNN are used. The collected data is pre-processed for resizing the images which are then used for feature extraction. Using CNN, the features are classified and trained. The trained model is saved and we compare the three algorithms such as CNN, SVM, and Random Forest. After the comparison of algorithms, the best algorithm is found.

For the real-time detection of face masks, the collected real-time data is preprocessed and the matrix data is used for feature extraction. Deep learning architecture learns various features from the given data sets. The proposed framework uses a cascade classifier to detect the faces present in the video frame. These facial images and cues are then processed by a neural network (CNN) for detecting the face mask. The CNN model is constructed using the Keras library and OpenCV to detect whether a person is wearing a face mask or not.

A.Proposed workflow

The Convolutional Neural Network (CNN) model is developed using Tensor Flow with Keras library and OpenCV to detect if the person is wearing a face mask or not.

B. Deep learning architecture

The deep learning architecture learns several important nonlinear features from the given samples and this learned architecture is used to spot the face masks.

C.Image Processing

Haar Cascade classifier is used to spot the faces present in the video frame. The images captured by the system's webcam needed preprocessing before going to the next step. In the pre-processing step, the image is converted into a grayscale image because the RGB color image contains so much redundant information that is not necessary for face mask detection. Then, we resized the images into (150x150) size to maintain the uniformity of the input images to the architecture. Then, the images are normalized and after normalization, the value of a pixel resides in the range from 0 to 1. Normalization helped the learning algorithm to learn faster and capture necessary features from the images.

ISBN 978-81-949892-5-7

D.Dataset Collection

To train the deep learning architecture, we collected images. Data from the source is collected for training and testing the model. The dataset contains images of faces only. It consists of about 1,200 images of which 660 images containing people with face masks and 540 images containing people without face masks. For training purposes, 80% of images of each class are used and the rest of the images are utilized for testing purposes.

A. Architecture Development

The learning model is based on CNN and it is very helpful for pattern recognition from images. Neural networks need to see data from both classes. The network comprises an input layer, several hidden layers, and an output layer. The hidden layers contain multiple convolution layers.

B. Result Analysis

By giving a reasonable proportion of different classes, the dataset is partitioned into training and testing set. The dataset comprises 1200 samples in total where 80% is used in the training phase and 20% is used in the testing phase. The developed architecture is trained for 10 epochs since further training results cause over fitting on the training data. Over fitting generally occurs when a model learns the unwanted patterns of the training samples. Hence,training accuracy increases and test accuracy decreases.

IV.EXPERIMENTAL RESULTS

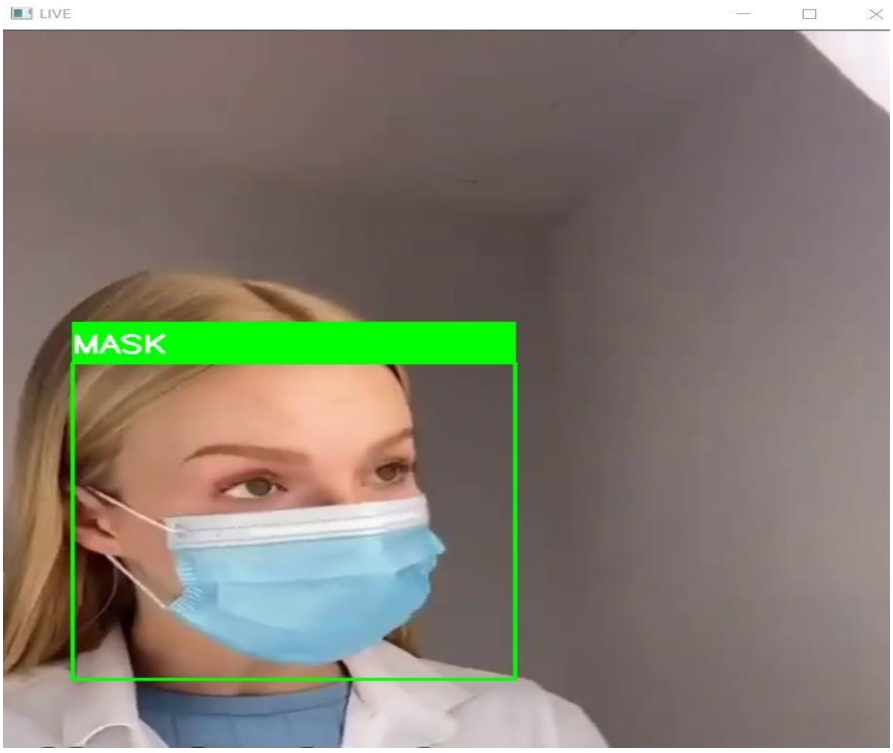


Fig 1: Face mask detection of single person

ISBN 978-81-949892-5-7

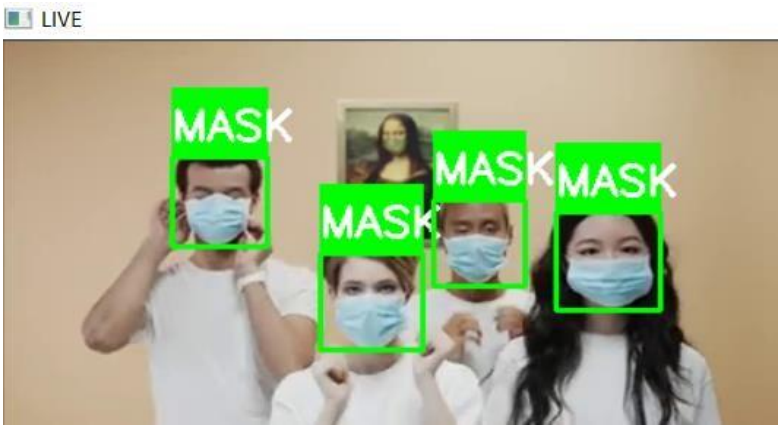


Fig 3: Face mask detection of a group of persons

Fig 2: Face mask detection of a person without mask

ISBN 978-81-949892-5-7

```
In [10]: svm_accuracy = model.score(x_test,y_test)*100
         print(svm_accuracy)
         86.5979381443299
```

```
rf_accuracy=clf.score(x_test,y_test)*100
print(rf_accuracy)
[1 1 1 0 0 0 0 1 0 0 0 1 0 0 1 0 1 0 1 1 1 0 0 0 0 0 0 0 1 1 0 0 0 1 1 1 0
 1 1 0 0 0 0 1 0 0 0 1 1 0 0 0 0 0 0 0 1 1 0 0 0 1 1 0 0 1 0 1 1 0 1 1 0 0 0
 1 0 1 1 1 0 0 1 0 0 1 0 0 1 1 1 1 0 0 0 0 0 1 0 1]
84.5360824742268
```

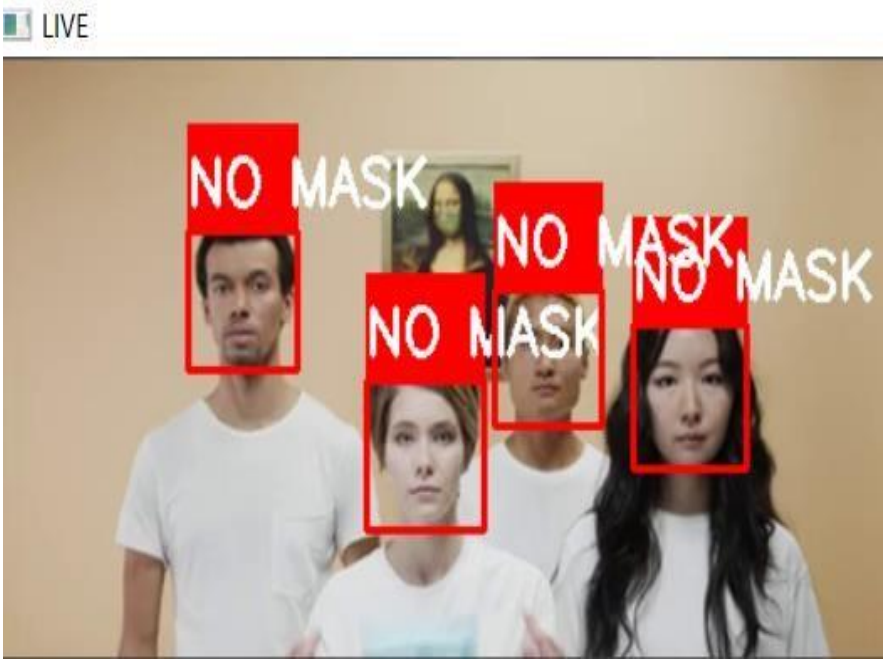


Fig 4: Face mask detection of group of persons without mask

Fig 5: Accuracy of CNN

V. CONCLUSION

```
INFO:tensorflow:Assets written to: model-010.model\assets
2/2 [==============================] - 1s 282ms/step - loss: 0.2091 - accuracy: 0.9516
[0.20905759930610657, 0.9516128897666931]
```

The present condition of Covid-19 demands an effective face mask detection application. The proposed model presents a real-time detection of face masks from videos using deep learning and the comparative study of machine learning & deep learning algorithms for face mask detection. It has been observed that each of the algorithms had an accuracy of more than 80%. CNN is the most effective in the detection of face masks as it had

ISBN 978-81-949892-5-7

the best accuracy (95%), over the other algorithms. So, the proposed model scheme of CNN is an efficient model to detect a face mask. The detector proposed, achieved higher accuracy and precision than the related work.

As per the suggestion of the government, wearing an N95 or double mask is good for public health care and hence the system can be improved to detect the face masks which are only N95 or double masks, and those who wear this will only allow their entry. Also, when a person is detected with no mask, the head of the organization can be notified via mail by keeping a record of live video streams and with more complex functions a screenshot of the person's face can also be attached to keep it as proof.

REFERENCES

1. Mohamed Loey, Gunasekaran Manogaran, Mohamed Hamed N. Taha “Fighting against COVID-19: A novel deep learning model based on YOLO-v2 with ResNet-50 for medical face mask detection” Elsevier 2021
2. W. H. Organization et al., “Coronavirus disease 2019 (covid-19): situation report, 96,” 2020.
3. Mingjie Jiang, Xinqi Fan and Hong Yan “Retina Face Mask: A face mask detector” 2020
4. R. Nageswara Rao, “Core Python Programming”, 2016.
5. Pankaj Jalote, “An Integrated Approach to Software Engineering”, 2005.
6. L. Liu, W. Ouyang, X. Wang, P. Fieguth, J. Chen, X. Liu, and M. Pietikainen, “Deep learning for generic object detection: A survey” International journal of computer vision, vol. 128, no. 2, pp. 261 – 318, 2020.
7. A. Kumar, A. Kaur, and M. Kumar, “Face detection techniques: a review,” Artificial Intelligence, “Deep learning and control algorithms of direct perception for autonomous driving,2019
8. Z. Wang, G. Wang, B. Huang, Z. Xiong, Q. Hong, H.Wu, K.Jiang, N. Wang, “Masked face recognition dataset and application,” and “Object detection with deep learning: A review, ”IEEE transactions on neural networks and learning systems, vol. 30, 2019.
9. https://scholar.google.co.in
10. https://www.researchgate.net

ISBN 978-81-949892-5-7

EMERGING TRENDS IN CLOUD COMPUTING

Muhammad Ameen
BSc. Electronics
Institute Of Human Resource Development
Nattika, Thrissur, Kerala, India ameensalim@gmail.com

Abstract
This study is a literature review on cloud computing trends as one of the fastest growing technologies in the computer industry and their benefits and opportunities for all types of organisations. It also addresses the challenges and problems that contribute to increasing the number of customers willing to adopt and use the technology. This study highlights the current and future trends of cloud computing and exposes the readers to the challenges associated with cloud computing. It also shows that the technology is promising and is expected to grow for there are many techniques to address the challenges of cloud computing such as security risks and privacy risks, through mobile cloud computing and cloud-computing governance.

Keywords – Cloud Computing, Cloud-Computing Governance, Mobile Cloud Computing, Privacy

I- INTRODUCTION

Cloud computing infrastructure is the backbone to the delivery pipeline of just about every digital service. Cloud computing allows companies to access computing services via the internet without needing to purchase infrastructure locally. Common business cloud services include data storage and analytics, development platforms and business software such as human resources (HR) and enterprise resource planning (ERP) systems. Consumer-facing cloud platforms range from media-streaming platforms to video conferencing software. The list is hardly exhaustive, with new cloud services constantly emerging to meet the changing demands of businesses and employees. Companies can choose to run their entire IT infrastructure in the cloud or implement cloud-based solutions for specific aspects of their operations. Typically, they pay for the services they need on a subscription basis, giving them the flexibility to pursue the precise technology capabilities they require.

II. CLOUD DELIVERY MODEL

Cloud services can be delivered in a variety of ways. The delivery model a company chooses to use varies based on its functionality requirements and the maturity of its IT and data governance needs. While public software-as-a-service (SaaS) solutions remain the largest market segment, vendors increasingly offer solutions that cater to a wide range of customers and requirements.

A. – Saas

Software-as-a-service (SaaS) applications deliver software over the internet that users access via a browser.The vendor manages the hardware, database, security and infrastructure, while users typically have some ability to configure the software to their needs. In the business context, these applications are often departmental. For example, customer relationship management for sales, service and marketing, HR software for HR.

ISBN 978-81-949892-5-7

B. Saas To Intelligent Saas

Software-as-a-Service (SaaS) will be infused with extra abilities and become intelligent SaaS. Artificial intelligence (AI) will lead the charge on this transformation. With many benefits like automation, analytical insights, and chatbots, AI will help expand what software can do. In the coming years, it will be hard for organizations not to embrace intelligent optimization within all tools.

AI, the Internet of Things, and blockchain are all technologies that will generate more data to add to the cloud. The real-time data generation from these solutions can help with insights into what's going on with customers and in the industry. Plus, these tools can run themselves in some instances, alleviating some monotonous administrative tasks.

C. - Paas

Platform-as-a-service (PaaS) cloud solutions provide developers with the software and operating systems they need to build cloud-based applications, be it a mobile app for better inventory tracking or a consumer-facing social media platform. Companies are also beginning to use PaaS cloud systems for their network security, since they can easily be customized to suit specific security requirements.

D. - Multicloud

Certain businesses want to distribute internal computer processing and storage requirements across multiple cloud platforms and applications, often from different vendors, based on their needs. It's common for them to choose different cloud providers for different functions, like ERP, security and marketing technology, for example. While an all-in-one business management platform that supports numerous functions is the best option for many companies, they may still require complementary solutions to help with other areas of the business. Businesses might also spread out their use of public clouds for compute resources to avoid lock in and gain leverage in negotiations.

E. - Private Cloud

A private cloud is a cloud computing model where services are provided over private infrastructure for the use of a single business, typically managed by that same business. Businesses choose private clouds to gain the benefits of cloud services through vendors without incurring the costs of building out and maintaining the cloud infrastructure themselves.

F. Hybrid Cloud

Many companies opt for a hybrid cloud model that combines public cloud services with the deployment of a private cloud, which is dedicated to a single business. This is especially true of organizations that collect sensitive data or operate in highly regulated industries like insurance, where data privacy is essential. A hybrid approach is attractive because it offers the necessary level of control without holding businesses back from innovation and scale as they roll out new services for their customers.

G. Serverless

Serverless cloud is a relatively new concept that's gaining traction in the market. Sometimes referred to as "functions-as-a-service," it means organisations aren't tied into leasing servers or paying for fixed amounts of storage or bandwidth. It promises a truly pay-as-you-go service where the infrastructure scales invisibly as an application requires it. Of course, it isn't really ***serverless*** – the servers are still there – but it adds another layer of abstraction between the user and the platform, meaning the user doesn't have to get involved with configurations and technicalities. ***Serverless*** within cloud computing will have a big part to play in the broader trend across the cloud and

ISBN 978-81-949892-5-7

the entire tech landscape of creating new user experiences that make innovation more accessible. ***Serverless*** computing is a form of cloud computing that lets businesses access IT infrastructure on-demand, without the capital investment and need to manage the infrastructure themselves. The difference between generic cloud computing and ***serverless*** is based on how resources are allocated — ***serverless*** is a subset of PaaS used by companies who need a lot of processing power, but only in short bursts.

Compiling software code is one example. ***Serverless*** models are gaining traction among companies big and small that want to build new applications quickly but lack the time, resources and/or budget to deal with the infrastructure. This let's growing businesses take advantage of greater computing power at a reasonable cost, while large organizations can roll out new digital services without adding to the burden of their already-stretched IT teams.

III. SMARTER WORKING WITH THE CLOUD

The cloud has emerged as more than a vehicle for computing power. Cloud storage and platforms also drive more efficient working practices, time and cost savings and innovation, helping companies modernize the way they work.

a. Machine Learning and Artificial Intelligence

Cloud-based artificial intelligence (AI) technologies, including machine learning, are helping businesses draw additional value from the ever-growing volumes of data they collect.

From logistics companies analyzing the efficiency of their transportation networks to e-commerce brands testing the performance of their websites in real time, AI algorithms empower businesses to glean new insights from their data and improve the way they work. Companies that don't have the budget or talent to build AI infrastructure of their own and many still don't take advantage of it by running systems from cloud service providers.

b. Automation

Automation is a key driver of cloud adoption, particularly when it comes to improving the efficiency of business operations. With their data and systems centralized on the cloud, companies can automate many of their internal processes, be it the consolidation of data from different locations or the creation of business intelligence dashboards. Today, many organizations are looking to tighten connections between different pieces of software with the aim of better managing their growing cloud footprints and ensuring that solutions from different vendors work together seamlessly. Cloud adoption is necessary and increasing quickly, which means that organisations have to deal with more computing; this will result in more data and application resources. This would require more admin jobs and time-consuming tasks. The automation of execution will reduce repetitive jobs, reduce errors and increase productivity. Therefore, companies of all sizes should aim to automate different processes. Automation will help simplify cloud administrators' jobs by saving cost and time.

c. Delegation Of It Ops

As more vendors launch solutions that can be hosted on external servers, some businesses will elect to outsource portions of their IT operations to third parties. Rather than hiring dedicated teams to build, manage and maintain their systems, companies can cut their operating costs and focus on the core product or service. That said, they must be

ISBN 978-81-949892-5-7

mindful of their sensitive data and technologies when deciding which operations to outsource so as not to compromise their governance or compliance practices.

IV. TRENDS IN CLOUD COMPUTING

Cloud computing trends demonstrate how this technology is changing the way businesses operate and how they allocate their IT budgets. Significantly, public cloud users (who share computing resources) no longer have to purchase and maintain hardware and other infrastructure or manage IT upgrades and software patches - that responsibility now falls on their cloud vendors. This leaves businesses and their IT teams able to focus on core business objectives like innovation, new product or service offerings and hiring new talent. It also helps to level the playing field for growing businesses that had been unable to afford the steep price tag of advanced technologies they now can access through a subscription. In recent years, hackers have found ways to compromise security in cloud computing, attacking computers through Wannacry and Ransomware and placing cloud-computing firms on guard. These continuing attacks alerted experts to increase their security and response time. Internet service providers work to enhance the quality of service on the Internet. Cloud-computing services require the ability to meet increasing demand for speed and storage globally.

a. Mobile Cloud Computing

Due to the wide availability and advances in smartphones, mobile cloud computing must be addressed in supporting applications and needed computational power. Therefore, mobile cloud computing can be thought of as combining mobile computing and cloud computing.

Mobile cloud computing is also defined as the integration of cloud computing with mobile devices to provide mobile devices with computational power, memory, and storage Mobile cloud computing may extend the hardware and battery life of the mobile device.

Issues and challenges of mobile cloud computing are performance, resources, and techniques. It is quite important for online social network services such as gaming, image handling, video processing, and general e-businesses.

b. Quantum Computing

Quantum is one of hottest topics in the cloud industry that challenges the present state of cloud computing and might transform it totally. Service providers are trying to cut-throat competition and in such a scenario Quantum Computing is heading to take over cloud computing in the near future.

c. Hybrid Cloud Solutions

Hybrid Cloud Solutions are expected to take its place very soon in the domain of cloud computing. Moreover, Hybrid Cloud Solutions are known for being dynamic, cost-effective, and also can adapt to the markets vibrant needs. With Hybrid Cloud Solutions, it is possible to attend to these market demands due to the rise in competition by large-scale enterprises.

d. AI In Cloud Computing

Cloud computing plays a key role in the deliverance of artificial intelligence (AI) services. Machine learning platforms require huge processing power and data bandwidth for training and processing data, and cloud data centers make this

ISBN 978-81-949892-5-7

available to anyone. Most of the "everyday" AI we see all around us – from Google Search to Instagram filters - lives in the cloud, and technology that routes traffic from data centers to our devices and manages storage infrastructure is built on machine learning. The development and evolution of cloud and AI are inextricably interwoven. Strong trends in AI will be "creative" algorithms – generative machine learning that can create anything from art to synthetic data to train more AIs – as well as language modeling – increasing the accuracy with which machines can understand human languages. Cloud computing will certainly play a key role in delivering these services to users as well as building the infrastructure to deliver them.

e. MULTI AND HYBRID CLOUD ENVIRONMENTS WILL CONTINUE TO GROW

Businesses recognize that cloud data management isn't about having one specific cloud platform or infrastructure, it's about choosing the solution that's right for the job at hand. In some cases, this isn't cloud at all, it might be on-premises or even legacy systems, especially where enterprises are locked into specific systems for business-critical solutions that are cumbersome or complex to adapt. The emerging cloud trend is that enterprises are becoming less worried about sticking with one vendor, and are embracing a multi cloud or hybrid cloud offering where they can get the best out of each solution.

V. ORGANIZATIONS WILL MAKE CLOUD SAVINGS A PRIORITY

The decentralized model of consumption has raised costs for organizations exponentially, and often without any control over the spiraling bottom line. Businesses will have to start to control these cloud costs as usage grows, streamlining the expenditure that they are not utilizing to full effect, and cutting out duplicate spending or unnecessary overheads.

As different cloud and hybrid services have different pricing and billing models, and costs can change from month to month, this could be a tall order. Innovative third-party solutions that can support organizations in getting granular insight into their hybrid network and provide unified management of costs will rise to the top.

A - SOLUTION-FOCUSED PARTNERS

As businesses continue to move their infrastructure using SaaS, PaaS and IaaS, one cloud trend is that there will be a continued need for third-party vendors who really specialize in meeting specific use cases and problem-solving new cloud challenges. While your cloud provider will be responsible for cloud infrastructure needs such as storage, outsources will be taking care of compute and networking, specific needs such as data, visibility, AI and ML technology, or IoT. These external partners will need to have a strong insight into how cloud computing and its associated technologies work, leaving the hardware element to the infrastructure leaders, but well-versed in how to get the best solutions on both the cloud and on-premises for specific organizational needs.

B. Integrating Blockchain Technology

Blockchain technology, which can track and record a product's journey, is helping with tracking through all stages of the product life cycle to improve delivery. When the cloud is connected to this technology, businesses will have an even better understanding of what happened during a product's delivery time. These insights, such as weather delays or tracking where a contaminated product originated, can save businesses millions of dollars. Although some companies are already executing this practice, it will continue to grow and mature in the coming years.

ISBN 978-81-949892-5-7

C. Increasing Security Capabilities

IT security and data compliance are major concerns for businesses and customers alike, and today's cloud solutions have evolved to address these concerns. Vendors have imbued their offerings with leading data controls and defenses that reduce the risk of human error when managing sensitive data.

With new data regulations, such as GDPR and CCPA, companies are no longer allowed to be careless with consumer data collection and storage. Especially with new security measures in place, businesses do not want to face monetary penalties or legal action. Because there are businesses unaware of the impact of these regulations, cloud companies will need to be well versed in possible weak points.

Security itself also needs to improve to face off with thieves using technology like AI or social engineering. Although incidents like this have been small, there may be an increase of new cyberattacks due to new laws. This is why cloud companies will need to maintain and improve security measures.

D. Hiring Digital Natives

The new employees joining companies today come highly knowledgeable in technology and the wide variety of tools available. Known as digital natives, these individuals may create a divide between employees who are technology literate and those who are not. Businesses will have to adopt new training techniques or bring on additional cloud solutions to assist new team members.

The cloud keeps growing and developing new abilities to assist businesses. New ways of collecting data, along with new technology capabilities, will lead to further cloud benefits for businesses.

VI. CLOUD-COMPUTING CHALLENGES

Adopting cloud computing has many challenges and problems

A. Cloud-Computing Governance

Due to the importance of cloud computing for improving organization performance, its governance plays an important role for decision-makers. Cloud-computing governance can be considered part of the general umbrella of IT governance. Drafting an international IT compliance regulations that can be adopted by cloud computing service providers all over the world. These compliance regulations should be applicable for both customers and the service providers in case of any dispute that might surface in the future between the two parties. This approach is expected to expel the customers' fear security analysis and performance evaluation from adopting the cloud computing technology.

B. Cloud-Computing Security

Computer security remains a critical and vital subject for scientists and practitioners. The problem is aggravated by the introduction of cloud computing because customers lack full control of the resources provided by cloud-computing service providers. Security in cloud computing is more challenging for customers and cloud-computing providers. IT governance offers visibility and IT control; therefore, efforts toward corporate governance can reduce operation risks, can establish compliance, and can protect the invested value to address cloud-computing security, mechanisms are needed to enhance and protect users from intrusion and attack. Researchers have long recognized the importance of protecting computer data from unauthorized access. Researchers discussed a three-way authentication approach that

ISBN 978-81-949892-5-7

helps in data security The approach proposed would help in effective three-factor security with low computational parameters that are effective while looking at security aspects compared to previously defined authentication techniques in cloud security. The protection of data from being tampered with by the cloud provider is a formidable task, especially if hackers are colluding with the cloud provider.

One method used to address the security and privacy issues of cloud computing is by introducing data encryption. However, the encryption of data increases the computational overhead and hence has a negative impact on the speed of data searches and retrieval. To alleviate such problems and gain the best results in obtaining encrypted data from outside sources and reducing the computational overhead of cloud servers, a cooperative method of preserving personalized searches that would maintain the privacy of the user.

Another interesting method used to address the problem of cloud-computing security is a lightweight-secure-conjunctive-keyword-search scheme in hybrid cloud environments (LCKS) based on a ciphertext-policy ABE algorithm, which supports file-owner authorized conjunctive keyword searches for multiple parties. The LCKS security analysis and performance evaluation indicate it is secure, highly efficient, and well suited for the hybrid cloud.

Cloud computing users must be aware of the vulnerabilities and the type of attack that might occur in cloud computing.

Cloud computing security is challenging and the problem becomes more complicated if we move to mobile cloud security.

The methods used for protecting computer data from unauthorized access and modifications can be classified into two main categories. The first type uses encryption and decryption. While methods using such approaches are very reliable in protecting the computer data from unauthorized access, these methods are not efficient especially for data that needs stringent requirements for storage and retrieval, because the encryption and decryption take considerable amount of time. The authors in this paper suggest that customers willing to adopt cloud computing have to conduct thorough investigations about their needs and whether efficiency is of paramount importance for them before choosing this method or the cloud computing service providers adopting such an approach. The second type of method that can be used for securing the data from unauthorized access and modifications, especially the internal threat, are based on Blockchains, a new technology that can be used to protect the customer from internal threats that might be committed by personnel of the cloud computing service providers.

C. Sase

With employees accessing more services and data from their own devices, which sit outside their companies' IT networks, businesses are rethinking their approach to security and risk management. Secure Access Service Edge (SASE), a term coined by Gartner, is a cloud-based approach to IT security that caters to the changing nature of work. Companies with a SASE architecture benefit from network security services such as secure gateways, firewalls, zero-trust network access and more, all based in the cloud. This is a robust approach to IT security that empowers businesses to deliver new cloud services quickly and with peace of mind that their systems are well-protected.

D. Cloud Based Disaster Recovery

Disaster recovery has been moving to the cloud, while similar to traditional disaster recovery, cloud-based disaster recovery backs up a company's data on an external cloud server and is generally more cost- and time-efficient, with the added bonus of being managed by an external provider. What's more, businesses can add, change and remove data from these external systems as they see fit without having to scale their own IT infrastructure. It's also common for

ISBN 978-81-949892-5-7

businesses to rely on cloud-based disaster recovery for critical servers and applications, such as large databases or ERP systems.

E. Data Privacy

One of the greatest challenges of cloud computing is the privacy issue. In the area of Internet of Things (IoT), devices are configured to access content or resources from multiple resources of variously integrated devices at the edge, which raises the issue of data privacy.

The healthcare sector has witnessed substantial increases in the use of cloud computing for the storage of patient medical information. Protecting patient medical data stored in the server from unauthorized personnel is very important for patients. The problem is exacerbated when the data is stored in the cloud-computing environment. Preserving privacy becomes a major concern for patients. Patients are concerned about the privacy of their data being stored on a cloud server. For this reason, healthcare providers must find solutions that guarantee the protection of patient data. The research that aims to address the data-privacy problem in cloud computing continues to increase. Researchers have proposed several new algorithms and techniques. Cryptosystem-based methods are an example of methods that employ high computing power of cloud servers that preserve data privacy. The main characteristic of these methods is to allow sharing and provide multi-user independent services. The encryption of data on the server side results in delays in the allocation and release of resource sharing by the cloud-computing server.

Another problem that faces cloud-computing clients is the choice of the cloud-computing cloud-computing service provider would better protect the confidentiality of data. Cloud-computing service providers and clients' practices should align, spelled out in security-level agreements. Among research areas that received considerable attention is mobile cloud computing. The proliferation of smartphones has encouraged the need to outsource data storage to the cloud server. Such requirements raise the issue of data privacy. Among the solutions used to protect data privacy is the use of encryption before sending the data to the cloud server. At the same time, users need to retrieve the data quite quickly and ensure the integrity of the queried data. To reduce the overhead of searching the data.

Encrypting and decrypting the data at the cloud server has proven its viability as an acceptable technique to preserve the privacy of data. In mobile computing, many users share information through the mobile cloud. To address the problem of data privacy stored in the cloud server, data owners usually encrypt their data before outsourcing them to the cloud server. Data encryption is a viable solution for preserving data privacy, but does not come without cost. The major problem of data encryption is the lack of efficiency. The encryption and decryption of the data takes considerable time and resources. The problem becomes more serious in the case of IoT devices, which have resource constraints. The issue of data privacy remains one of the hottest challenges in cloud computing due to the proliferation of IoT devices that have resource constraints and at the same time there is a compelling need to improve
the security to protect the customer's private data
.

VII. Innovation and Application Development

As companies look to differentiate themselves by quickly launching new products and services, cloud-based platforms are evolving to support their development needs at a record pace. From purpose-built coding environments to decentralized data storage, cloud computing has opened the door to new possibilities in application development.

ISBN 978-81-949892-5-7

A. Containers and Kubernetes Collaboration

Containers offer businesses a dedicated, cloud-based space where they can build, test and deploy new applications. This allows developers to focus on the details of their applications and IT teams to focus on deploying and managing solutions as they are developed, making the entire process quicker and more efficient.

Kubernetes is an open-source container orchestration solution that simplifies the process of deploying and managing applications developed in containers. In addition to automated scaling of applications based on customer demand, the software also monitors the performance of new services so businesses can address issues proactively.

B. Edge Computing

This form of cloud computing brings data processing collection, storage and analysis closer to the sources generating the data, rather than a centralized cloud. This reduces latency and powers the use of edge devices. Edge computing is the driving force behind smart devices, such as smartphones, smartwatches and smart cars, and the interconnection of all the data generated by these technologies.

C. Cloud Native

Cloud-native applications allow businesses to build and deliver new software to their customers more quickly than traditional monolithic cloud applications, which run on a single hardware or software hub. Instead, cloud-native applications are built as a network of distributed containers and microservices. That means multiple teams can work on new features at the same time, accelerating the pace of innovation.

D. Cloud Based Operating Models

In addition to supporting more efficient technology operations, cloud computing has led to be increased collaboration between employees across teams, departments and geographies.

E. - Collaboration

Remote work is not a new phenomenon, but it has become far more common recently and the rise of cloud collaboration platforms has made the approach more viable. Secure networks, conferencing and communication platforms have become must-haves for modern organizations. Collaboration platforms are also becoming more advanced, ranging from spreadsheets that update in real time and other types of work management software to advanced data warehouse modeling tools that keep teams across geographies up to date.

F. Virtual Cloud Desktop

A virtual cloud desktop, also known as desktop-as-a-service, delivers the entire desktop operating system and software applications as a cloud-based service directly to a laptop, desktop or other device. Companies only pay for the time their staff spend logged in to their devices, and they don't have to pay for hardware upgrades. Virtual cloud desktops can also be scaled instantly, which means companies always have the licenses and devices they need to support their growing workforce.

*G***.** Cloud ERP

Cloud ERP is Software as a Service that allows users to access Enterprise Resource Planning (ERP) software over the Internet. Cloud ERP generally has much lower upfront costs, because computing resources are leased by the month

ISBN 978-81-949892-5-7

rather than purchased outright and maintained on premises. Cloud ERP also gives companies access to their business-critical applications at any time from any location. While technically the only difference between Cloud ERP and on-premises ERP is where the software is physically located, there are other significant differences. Here we explain some of the key characteristics and advantages of Cloud ERP software. The Cloud is particularly valuable to small and medium-size businesses (SMB's) because it provides access to full-function applications at a reasonable price without a substantial upfront expenditure for hardware and software. Using the right cloud provider, a company can rapidly scale their business productivity software as their business grows or a new company is added.

VIII. Cloud Costs

The upfront costs of cloud computing are much lower than those that come with buying and setting up on-premises IT infrastructure and systems. The same applies to having to maintain and upgrade hardware and software, along with the cost of staff charged with those duties. However, that's not to imply migrating to the cloud is always an inexpensive proposition. If not managed properly, the variable cost nature of the cloud may potentially exceed budget and increase total costs over the long run.

IX. CONCLUSION

In summary, cloud-computing trends are poised to answer companies' current and future needs. Because technology is essential to firms, cloud computing allows companies to store and access their data at any time. This feature has caused cloud computing to become increasingly popular very quickly. Over time, service providers are working to increase the number of services they provide, which are likely to include enhanced analytics services. Various benefits arise from the use of cloud-computing and cloud-storage services. Foremost is security of data. Over time, more and more businesses will store their data in the cloud and will contract with service providers to perform data analytics using the cloud. Even more notable is that, in the future, companies will have no other option than to store their data in the cloud. Business competition will rest largely on data safety and the ability to share and access data. Organizations are likely to become increasingly interdependent. Companies require a reliable cloud-computing environment that meets their needs and desires. Optimally, global enterprises will develop a plan to improve their use of cloud computing. To align with those plans, Internet service providers will enhance Internet speeds and reduce times when computers are offline, enabling users to rely on the cloud to access data instantly. Companies that do not join in this effort are likely to be less competitive.

REFERENCES

- Zhang, L. et al., 2013. Moving big data to the cloud. INFOCOM, 2013 Proceedings IEEE
- Brij B. Gupta, Modern Principles, Practices, and Algorithms for Cloud Security
- Ronald L. Krutz, Russell Dean Vines, Cloud Security: A Comprehensive Guide to Secure Cloud Computing
- John M. Acken, Naresh K. Sehgal, and Pramod Chandra P. Bhatt, Cloud Computing with Security: Concepts and Practices
- Kris Jamsa, Cloud Computing: SaaS, PaaS, IaaS Virtualization, Business Models, Mobile, Security and More
- Mohammed M. Alani, Elements of Cloud Computing Security: A Survey of Key Practicalities
- Marcia R. T. Pistorious, The Quick Guide to Cloud Computing and Cyber Security
- Daniel Slamanig and Stefan Rass, Cryptography for Security and Privacy in Cloud Computing
- Dor Skuler, Hui-Lan Lu, and Igor Faynberg, Cloud Computing: Business Trends and Technologies
- Hunter Muller, On Top of the Cloud: How CIOs Leverage New Technologies to Drive Change and Build Value Across the Enterprise
- Theo Lynn, John G. Mooney, Brian Lee, Patricia Takako Endo, The Cloud-to-Thing Continuum: Opportunities and Challenges in Cloud, Fog and Edge Computing

ISBN 978-81-949892-5-7

● Michael H. Hugos, Derek Hulitzky, Business in the Cloud: What Every Business Needs to Know About Cloud Computing
●Vic (J.R.) Winkler, Securing the Cloud: Cloud Computer Security Techniques and Tactics
● Dinkar Sitaram, Geetha Manjunath, Moving To The Cloud: Developing Apps in the New World of Cloud Computing
● Kenneth P Birman, Guide to Reliable Distributed Systems: Building High-Assurance Applications and Cloud-Hosted Services
● Pankaj Arora, Raj Biyani, Salil Dave, To the Cloud: Cloud Powering an Enterprise ● IT Governance Publishing, Cloud Computing: Assessing The Risks
●Vaughn Stewart, Micheal Slisinger, Vytautas Malesh, Dr Stephen Herrod, Virtualization Changes Everything: Some Strategies for Cloud Computing
● Douglas K. Barry, Web Services, Service-Oriented Architectures, and Cloud Computing: The Savvy Manager's Guide

ISBN 978-81-949892-5-7

An Overview of Various Machine Learning Based Brain Cancer Detection Methods

Ms. Joicy Joy
Asst.professor in computer science department
NIMIT
joicy@naipunnya.ac.in

Ms. Nithya Paul
Asst.professor in computer science department
NIMIT
nithya@naipunnya.ac.in

Abstract

Cancer is a disease of the brain in which cancerous cells develop in the brain tissue. The cancerous cells grow to form a mass that interferes with everyday functions such as motor control, sensation, memory, and other bodily functions. In this Paper focus on four detection methods for brain cancer and done a comparative study on these detection methods.

Keywords - tensor flow, MRI, Machine learning, Helicoid demonstrator, hyper spectral images, HGG, LGG.

I.INTRODUCTION

A brain cancer is a collection, or mass, of abnormal cells in your brain. There are two main types of tumors: malignant tumors and benign (non-cancerous) tumors. Tumors that are malignant are called malignant tumors, and benign tumors are composed primarily of noncancerous cells. Most cancers cells that develop from brain tissue are called primary mind tumors while tumors that spread from different frame websites to the mind are termed metastatic or secondary brain tumors. Records endorse that brain most cancers happen every so often (1.4% of all new cancer patients in line with 12 months), so it isn't taken into consideration to be a commonplace infection and is possibly to increase in about 23,770 new humans according to 12 months with approximately 16,050 deaths as envisioned the country wide most cancers Institute (NCI) and the American Cancer Society.

Best about five% of brain tumors may be because of hereditary genetic conditions including neurofibromatosis, tuberous sclerosis, and some others. Brain most cancers is an extraordinary development of cells in the brain. Cancer cells multiply to create a tumor, which obstructs mind strategies, along with muscular control, sensation, memory, and other ordinary bodily features. Humans have a much less than 1% hazard of acquiring a malignant mind tumor in the course of their lifetime. Cancer is graded based on its pathologic characteristics or how the cells appear underneath a microscope. The grades/ stages of mind most cancers are as follows:

Grade 1: A stage 1-brain tumor is noncancerous or slow-developing. Its cells resemble healthful cells in appearance and are often cured with surgical operation.

Grade 2: The stage 2 mind tumors are malignant but slow growing. Its cells under a microscope appear quite uncommon. Those tumors have the capability to unfold to adjacent tissues or recur after preliminary remedy.

Grade 3: The stage 3 mind tumors are malignant and expand greater speedy than grade 1 and a couple of tumors. While viewed underneath a microscope, the malignant cells show extreme abnormalities. Level three mind most cancers can actively generate odd cells, which could unfold to other regions of the mind.

Grade 4: The stage 4 cancer mind tumors broaden rapidly and feature numerous peculiar features that may be visible underneath a microscope. Most cancers timeline is competitive, in which the tumors can spread to different regions of the brain and can even create their blood arteries to keep up with their fast boom. Now and again, in addition they characteristic tiny clusters of dead cells (necrosis).

ISBN 978-81-949892-5-7

In this Paper focus on four detection methods for brain cancer and done a comparative study on these detection methods. In section 2, different methods for cancer detection is explained. In section 3, describe the findings and conclusion on detection methods.

II.LITERATURE REVIEW ON DIFFERENT METHODS FOR CANCER DETECTION.

A. BRAIN CANCER DETECTION BY MRI USING TENSOR FLOW.

The common diagnosing method for brain cancer is MRI. MRI can be used to detect the brain cancer by analyzing the MRI but this procedure is vary time consuming for vast number of cases. By the use of convolutional neural network classifier, we implemented this by using tensor flow.

Tensor Flow is an open source software library released in 2015 by Google to make it easier for developers to design, build, and train deep learning models. The name Tensor Flow derives from the operations that such neural networks perform on multidimensional data arrays. These arrays are referred to as "tensors". In June 2016, Dean stated that 1,500 repositories on GitHub mentioned Tensor Flow, of which only five were from Google.
The Tensor flow can be used to tackle the problem of Classification. Tensor Flow is only one of several options available to developers; we choose to use it here because of its thoughtful design and ease of use. The data we acquired was from various online resources [3], [4], [5], [6]. The dataset was in dicom format. We used a tool called mango to obtain their equivalent JPG/PNG image.

The high-quality data is the key to great machine learning models. But good data doesn't grow on trees, and that scarcity can impede the development of a model. One way to get around a lack of data is to augment dataset. Smart approaches to programmatic data augmentation can increase the size of your training set. Here we flipped each image horizontally and rotation by 30° left as well as right by 30° in training set allowing us to work on wide variety of data. The convolutional neural network architecture used here is the most basic lenet architecture. This architecture contains five layers. Out of which two layer are convolutional layer, the other two layers are pooling layer and last layer is fully connecting layer. In convolutional layer are the current state-of-the-art model architecture for image classification tasks.

Convolutional layers, which apply a specified number of convolution filters to the image. For each sub region, the layer performs a set of mathematical operations to produce a single value in the output feature map.

Pooling layers, which down sample the image data extracted by the convolutional layers to reduce the dimensionality of the feature map in order to decrease processing time. A commonly used pooling algorithm is max pooling, keeps their maximum value, and discards all other values.

In a dense layer, every node in the layer is connected to every node in the preceding layer. The last convolutional module is followed by one or more dense layers that perform classification. The final dense layer in a CNN contains a single node for each target class in the model .The filter size and number of filters in first two convolutional layers are 3 and 32 respectively. While the third convolutional layer had same number of filter size but the number of filters were doubled to 64. The fully connected layer size used is 128.

One epoch is one forward pass and one backward pass of all the training examples through the network. The learning rate decided here is $\alpha=0.0001$ and this learning rate is utilized by Adam Optimizer.

ISBN 978-81-949892-5-7

B. HELICOID TOOL DEMONSTRATOR FOR REAL-TIME BRAIN CANCER DETECTION.

The goal of this demonstration is to show how the combination of hyper spectral imaging and machine learning can be a potential solution to precise real-time detection of tumor tissues during surgical operations. Hyper spectral Imaging (HI) collects a huge amount of spectral and spatial information from a scene so hundreds of bands covering a narrow and almost continuous portion of the spectrum conform each pixel.

The facts furnished isnormally based as recordscube wherein each plane includingan image capturedata sure wavel ength. Analyzing the recordscube theuseof machine learning strategies facilitates determining thekindof material p resent at each pixel.

The main goal of the Helicoid (Hyper spectral Imaging Cancer Detection) is to develop an experimental intraoperative setup based on hyper spectral cameras connected to a massively parallel processing platform running a set of algorithms capable of precise real-time discrimination between healthy and pathological tissues during neurosurgical operations. This information is provided to the surgeon through a dedicated display device, overlapping normal viewing images of the surgical scene with simulated color images that indicate the presence or absence of cancer in the exposed tissue.

The demonstrator shows the possibility to provide neurosurgeons with a tool to make precise real-time decisions during brain tumor resection. This tool will contribute to improve surgery results and, as a consequence, the quality of life of patients after surgery

C.DETECTION AND CLASSIFICATION OF HGG AND LGG BRAIN TUMOR USING MACHINE LEARNING.

The proposed computerized system uses k-means as the segmentation technique for clustering whilst Discrete Wavelet Transform (DWT) and Principal Component Analysis (PCA) are the main parts of the feature extraction and feature reduction mechanisms, respectively.

A proposed system is designed for the precise detection and classification of normal and abnormal brain MRI's and then the classification of the abnormal MRI's into HGG or LGG glioma tumor. Brain MRI is read by the system, and then Otsu binarization is applied to convert the image into a binary image. After that k-means, clustering is applied for segmentation. Later, DWT and PCA are applied. Finally, SVM is used for classification. In stage 1, the images are classified into normal or abnormal MRI's. In stage 2, the abnormal MRI images are classified into HGG or LGG glioma tumor MRI's.

The system can be made more reliable by using a larger number of data. More relevant features for accurate classification can be found out. This computerized system could be further used for the classification of other brain diseases and for other alternative medical images of different pathological condition, types, and disease status.

D.SPATIO-SPECTRALCLASSIFICATION OF HYPER SPECTRAL IMAGES FOR BRAIN CANCER DETECTION.

Surgery for brain cancer is a major problem in neurosurgery. The diffuse infiltration into the surrounding normal brain by these tumors makes their accurate identification by the naked eye difficult. Since surgery is the common

ISBN 978-81-949892-5-7

treatment for brain cancer, an accurate radical resection of the tumor leads to improved survival rates for patients. However, the identification of the tumor boundaries during surgery is challenging.

Hyper spectral imaging is a non-contact, non-ionizing and non-invasive technique suitable for medical diagnosis. This study presents the development of a novel classification method taking into account the spatial and spectral characteristics of the hyper spectral images to help neurosurgeons to accurately determine the tumor boundaries in surgical-time during the resection, avoiding excessive excision of normal tissue or unintentionally leaving residual tumor. The algorithm proposed in this study to approach an efficient solution consists of a hybrid framework that combines both supervised and unsupervised machine learning methods. Firstly, a super-vised pixel-wise classification using a Support Vector Machine classifier is performed. The generated classification map is spatially homogenized using a one-band representation of the HS cube, employing the Fixed Reference t-Stochastic Neighbors embedding dimensional reduction algorithm, and performing a K-Nearest Neighbors filtering.

The information generated by the supervised stage is combined with a segmentation map obtained via unsupervised clustering employing a Hierarchical K-Means algorithm. The fusion is performed using a majority voting approach that associates each cluster with a certain class. To evaluate the proposed approach, five hyper spectral images of surface of the brain affected by glioblastoma tumor in vivo from five different patients have been used. The final classification maps obtained have been analyzed and validated by specialists. These preliminary results are promising, obtaining an accurate delineation of the tumor area.

III FINDINGS AND CONCLUSIONS

Brain tumor detection is done using MRI and analyzing it. The machine learning is very powerful strategy for the detection of the cancer tumor from MRI. Here we achieved the training accuracy of 99% and validation accuracy of 98.6%, with validation loss from 0.704 to 0.000 over 35 epochs. This is very basic image classification method of lenet architecture. The more powerful approaches are available. When the model was created on the CPU based tensor flow and GPU version of tensor flow is much faster to train, which will result in much faster model creation. The more sophisticated system should take MRI images in dicom format directly and operate on them.

The demonstrator shows the possibility to provide neurosurgeons with a tool to make precise real-time decisions during brain tumor resection. This tool will contribute to improve surgery results and, as a consequence, the quality of life of patients after surgery. Classification maps from the different machine learning algorithms evaluated, as well as implementation results on a many core platform, when in-vivo human brain hyper spectral images are employed as inputs, are shown as evidences of the tool potential.

Automatic segmentation and classification of the MRI images is the only way by which brain tumor can be detected and classified in early stages with a high percentage of accuracy. The total number of images used in the system is 440. Among them, there are 100 normal images and 340 abnormal images. The computerized system has successfully classified HGG and LGG with an accuracy, sensitivity and specificity of 99%, 100% and 98.03% respectively. The system can be made more reliable by using a larger number of data. More relevant features for accurate classification can be found out. This computerized system could be further used for the classification of other brain diseases and for other alternative medical images of different pathological condition, types, and disease status.

Brain cancer detection algorithm to classify HS images of brain tumor in surgical-time during neurosurgical operations. It has been demonstrated that the use of HSI as a new non-invasive surgical-time visualization tool can improve the outcomes of the undergoing patient, assisting neurosurgeons in the resection of the brain tumor. The identification of the tumor boundaries and the tumor infiltration into normal brain is highly relevant in order to avoid excessive resection of normal brain and to avoid unintentionally leaving residual tumor. Currently, further

ISBN 978-81-949892-5-7

investigations are being carried out by the research team in order to generalize the results obtained, to optimize the algorithms and validate their findings, as well as to increase the image database and optimize the acquisition system. Furthermore, the use of other hardware acceleration platforms (such as GPUs or FPGAs) are currently under consideration to implement the full brain cancer detection algorithm. Such implementation must explore the design space to achieve the best tradeoff between real-time execution, memory usage and power dissipation using heterogeneous platforms. This next generation of medical HSI systems could offer neurosurgeons a real-time visualization tool to assist them during the entire process of the tumor resection providing several TMD maps per second.

REFERENCES

[1] Prastawa M, Reyes M, Van Leemput K. "The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS)", IEEE Transactions on Medical Imaging 34(10), 1993-2024 (2015) DOI: 10.1109/TMI.2014.2377694

[2] Bakas S, Akbari H, Sotiras A, Bilello M, Rozycki M, Kirby JS, Freymann JB, Farahani K, Davatzikos Advancing the Cancer Genome Atlas glioma MRIcollections with expert segmentation labels and radiomic features", Nature Scientific Data, 4:170117 (2017) DOI: 10.1038/sdata.2017.117

[3] Spyridon Bakas, Hamed Akbari, Aristeidis Sotiras, Michel Bilello, Martin Rozycki, Justin Kirby, John Freymann, Keyvan Farahani, and Christos Davatzikos. (2017) Segmentation Labels and Radiomic Features for the Pre-Operative Scans of the TCGA-GBM collection. The Cancer Imaging Archive.

[4] Spyridon Bakas, Hamed Akbari, Aristeidis Sotiras, Michel Bilello, Michel Rozycki, Justin S Kirby, John B Freymann, Keyvan Farahani, Christos Davatzikos. "Advancing The Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features", Nature Scientific Data,

[5] Aaswad Sawant, Mayur Bhandari, Ravi Kumar Yadav, Rohan Yele, Mrs. Sneha Bendale “BRAIN CANCER DETECTION FROM MRI: A MACHINE LEARNING APPROACH (TENSORFLOW)” International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056 Volume: 05 Issue: 04 | Apr-2018.

[6] R. Salvador1, H. Fabelo, R. Lazcano, S. Ortega, D. Madroñal, G. M. Callicó, E. Juárez, C. Sanz Demo: HELICoiD Tool Demonstrator for Real-Time Brain Cancer Detection

[7] F. P. Polly, S. K. Shil, M. A. Hossain, A. Ayman1, and Y. M. Jang “Detection and Classification of HGG and LGG Brain Tumor Using Machine Learning”.

[8] Himar Fabelo, Samuel Ortega, Daniele Ravi, B. Ravi Kiran, Coralia Sosa, Diederik Bulters, Gustavo M. Callico´, Harry Bulstrode, Adam Szolna, Juan F. Piñeiro, Silvester Kabwama “Spatio-spectral classification of hyperspectral images for brain cancer detection during surgical operations”

ISBN 978-81-949892-5-7

Bitcoin-Applications In Modern Era And Its Methodology-
An Overview

Siji Jose, Fredy Varghese
Assistant Professor, Naipunnya institute of information Technology, Pongam, Kerala
fredy@naipunnya.ac.in , siji@naipunnya.ac.in

Abstract

Bitcoin is a digital currency that is not backed by any government or central bank. It can be exchanged for goods or services from merchants who accept bitcoins. These bitcoins are the secure data blocks. BitCoin is a cutting-edge technology that provides a number of advantages, including quick transaction times, cheap transaction fees, and the elimination of the requirement for a third-party intermediary to complete transactions. Regulators have been wary about BitCoin because the technology has been exploited for illicit purposes, such as online drug purchases and Ponzi scams. This article gives a fundamental overview of BitCoin's uses and implementation methods.
Keywords: Cryptocurrency, Blockchain, Mining, Investment

1. Introduction

People have been transacting with fiat money since its inception. It has become more easy to trade and transact. Following the worldwide financial crisis of 2008, the first cryptocurrency, Bitcoin, was launched in 2009. It was created by Nakamoto, an unidentified group or individual who created Bitcoin as the first digital currency enabling easier day-to-day transactions between individuals. Bitcoin works without the use of a middleman like banks or monetary institutions. It is a peer-to-peer transaction that does not need the disclosure of one's identify. Unlike present practise, the bank acts as a middleman or go-between, knowing the identities of both the buyer and the seller, raising concerns about personal data protection. The Bitcoin platform has made cryptocurrency trading and transactions much easier and more autonomous, all while protecting personal information and details. Choosing this type of transaction has given some people the ability to transact freely and anonymously.

Bitcoin was the world's first digital coin to leverage the blockchain network. It's made up of computers from all around the network who participate in a transaction log. By preventing fraudsters from using the currency more than once, this blockchain offers one of the most secure security mechanisms. Proof of work is used in the blockchain protocol to ensure that miners converge on this structure[7]. Hashing is the computational operation, and hashing power refers to the computational power required to mine the currencies . Even for industry players and experts conducting study in this subject, the bitcoin market's system is complex and difficult to comprehend. Many studies have revealed the benefits of Bitcoin, including security , low transaction, high returns, and use as an alternative tool for a country's bailout process. Despite this, some researchers have pointed out the risks and drawbacks of using this digital coin, such as a lack of regulation , a high electricity bill due to energy consumption a lack of security, and other issues such as anonymity .

1.1 Background and History

Nakamoto first presented Bitcoin in 2009, with 50 coins in circulation at the time. Only computer geeks around the world were sceptical in the early stages of the excitement. Mt Gox, a Japanese business, launched a website in 2010 that used Bitcoin as a mercantilism mechanism,with twenty coins ever-changing hands for four.951 cents. The

ISBN 978-81-949892-5-7

overall value of the transaction was around one dollar. The price of Bitcoin has risen dramatically as the number of people using it has grown, and at the time this paper was prepared, the price had risen to $6,777 USD.

According to Bohme et al, the value of bitcoin is based on scarcity (2015). It is the foundation for determining the monetary value of any object. Under the current practise of using fiat currency, the monetary authority or central bank holds and reserves the money. The central bank of a country has the power to regulate money circulation and absolute quantity[2]. To govern a country's fiscal economy, the bank can only produce a limited amount of this paper money, resulting in scarcity. This scarcity will be recorded in the bank's books and legally safeguarded.
Is Bitcoin regarded genuine money? That is the fundamental topic that has arisen since Bitcoin's introduction. History has shown, that money must meet the following criteria: (1) A value store. It is a form of purchasing power that consumers can use to purchase items from the present to the future. (2) An exchange medium. (3) A unit of account and the ability to make payments. The value of any merchandise purchasable which will be measured. Money should theoretically meet all of these requirements, but this is not always the case[1]. All three criteria square measure controversial once watching Bitcoin and different cryptocurrencies in their current state. One may argue that it has a store value because of its purchasing power, but owing to the uncertainty, it is impossible to predict if Bitcoin will be utilised in the same way in the future as it is now. Some argue that bitcoin can be used as a medium of commerce, but others argue that the products that may be exchanged are limited.If these three conditions are established as pre-requisites for any commodity to be granted the status of money, it should be accepted in the context of its usage and application. During the difficult times of World War II, when captives in war camps used cigarettes for transaction, Radford (1945) claimed that cigarettes met all of these criteria. Cooking salt was once considered valuable during the Roman Empire, when soldiers' wages were paid in salt[8]. To anyone who have access to a computer and the internet, bitcoin can be considered money. The issue stems from the fact that only a small percentage of the world's population has access to internet gadgets. As a result, Bitcoin is merely on the market to people who have access to the net,cherish the captive within the war camp and also the Roman troops. Thanks to its period convertible to a traditional currency with mounted worth, per Jakob Bohme et al.cryptocurrency, particularly Bitcoin, is more of a platform for payment than currency. In terms of portfolio analysis, risk management, and sentiment analysis, this cryptocurrency is unlike any other asset (Dyhrberg, 2016). In terms of having a certain value, cryptocurrency has a similar portfolio to other assets such as gold, property, equities, and equity. Cryptocurrency, on the other hand, reflects public sentiment when its value rises and more individuals are prepared to accept it as payment. These distinctions give the market a variety of options, which investors and stakeholders alike can take advantage of. As a result, accepting bitcoin as a replacement for fiat money in today's economics is still premature and requires additional theoretical and practical understanding.

2.Hopes and Benefits

Cryptocurrency's prospects are optimistic, despite the fact that it is a relatively new commodity. Despite the fact that its price and value have increased, the fruits and future potential are still sought after. The following section examines the real-world applications of cryptocurrencies for users, investors, and the government.

2.1.1. Secure

Since the internet's inception, the blockchain has been regarded as one of the most advanced platforms and technologies. It ensures the security and privacy of internet transactions. In their case study, Ying et al. (2018) found that, in addition to allowing the usage of cryptocurrencies, the blockchain can preserve secret information and eliminate the need for any intermediaries[2].

Blockchain technology is extremely safe. Fraudsters will be unable to commit such a crime since many ledgers cannot be changed or validated at the same time). Hash power is the ability to regulate computational power. According to Khatwani (2018), hash power is the amount of computing power required for a cryptocurrency network

ISBN 978-81-949892-5-7

to run continually. Using a cryptocurrency algorithm is safer and more secure than using a credit card. Cryptocurrency boasts substantially lower processing fees and provides secure transactions, despite the fact that it is currently understudied.

2.1.2. Cost of Deals

People have used some kind of monetary form for day-to-day transactions throughout history. The barter system had started the business as early as a trade system, when people exchanged or bartered their things with mutual agreement. As time passes, fiat money was created to allow individuals to trade without needing to bring large items to exchange. As the world enters the twenty-first century, cryptocurrency has swept the market. Bitcoin has been used by large international corporations as a type of cash, and it has even been used to pay monthly payments to staff. In terms of transaction costs, cryptocurrency and Bitcoin transactions are less expensive than traditional currencies[1]. Due to its low transaction cost, cryptocurrency has key qualities such as decentralisation and deregulation (Kim, 2017). There have been numerous problems with the present payment system, which is based on credit and payroll cards. The amount of interest imposed to users who default on their payments is far too excessive, putting them in a financial bind[5]. This has not been the case with cryptocurrencies, where trading takes place only once end-to-end users have consented, and only then will money be sent.

Furthermore, bitcoin can be used at any time of the year, 24 hours a day, 7 days a week. Data pricing is available immediately, allowing everyone in the world to trade for free as long as they have access to the internet (Pieter & Vivanco, 2017). With the recent rise of the internet of things (IoT) and reliance on big data, users will appreciate the opportunity to trade without time constraints[3]. This kind of remuneration would benefit the younger generation, who are expected to become business owners in the future and work on their own schedules rather than being bound by traditional working hours. This kind of trading is also excellent for computer knowledgeable individuals who do not wish to incur the additional costs associated with using other payment systems.

2.1.3. High Performance

The specific characteristics of cryptocurrencies, as well as its flexibility to fit into many economic functions, make it a one-of-a-kind asset (Briere et al., 2015). Bitcoin has a history of being a highly volatile currency with a high rate of return for investors[10]. Apart from that, Bitcoin's risk is low due to its inclusion in a variety of portfolios. Investors are well aware that the best way to benefit from their investments is to buy low and sell high. Those who held Bitcoin in the early days of its launch may have made 1000-10000 percent returns on their investment (Bohme et al., 2015).

Ciaian et al. (2016) looked at both classic and current factors of currency supply and demand, such as currency attractiveness. He also looked into the interactions between pricing determinants. It is predicted that, as a result of market demand, the price of Bitcoin would rise, particularly when the supply of Bitcoin in circulation is greater than it is presently. When the amount of Bitcoin is half every four years, it means that fewer new Bitcoins will be created, making the currency more stable. The scarcity of Bitcoin will only drive up the price, despite the fact that it is the most widely traded currency in the planet. Investors will desire Bitcoin during this stable moment.and users alike. According to Kristoufek (2013), the price of Bitcoin has climbed in lockstep with the number of searches on Wikipedia and Google Trends. It demonstrates that the relationship between search queries and the price of Bitcoin is intertwined. This means that as more people become aware of the existence of cryptocurrencies and how it may benefit them, the price will likely rise[7]. The price of cryptocurrencies will stabilise in the near future, as more people become computer and internet savvy, and those who have hung on to their coins will reap the benefits of their investment.

Cryptocurrency may be used in the same way that fiat money or credit cards can be used to buy actual items from retailers. Apart from that, Wingfield (2013) claims that Bitcoin can be utilised for a broader range of applications. The rise in popularity of cryptocurrency, particularly Bitcoin, was aided by a number of specific events that impacted its utility, such as the banking crisis in Cyprus from 2012 to 2013 and the European sovereign debt crisis (ESDC)

ISBN 978-81-949892-5-7

from 2010 to 2013. Cyprus had taken the first moves toward obtaining a bailout by imposing a charge on bank deposits. This is due to the risk associated with traditional deposits (Luther & Salter, 2017). Ha and Moon (2018) used genetic programming to evaluate the profit pattern in cryptocurrency investing. The result demonstrates that the pattern from the study had frequent and profitable signals.

It was also replicated in trading with the pattern and the signals, demonstrating that it can be beneficial for any cryptocurrency portfolio. Hong (2017) found that Bitcoin return was large utilising time series momentum. It was discovered that Bitcoin had a high return during a period of 8 weeks. The prediction was supported by evidence of empirical return continuance and reversal, as well as the predictability of the time series, based on well-known asset theories. As previously said, due to the volatility of cryptocurrencies, the time length of Bitcoin's continuation and reversal was less than that of another asset. Institutional investors can make money by investing in Bitcoin instead of shares in their portfolio..

3.Challenges

Despite the benefits in cryptocurrency, the cryptocurrency nevertheless faces numerous hurdles. Because of the risk and obstacles that trading and investing in cryptocurrencies poses, onlookers and new investors have undoubtedly taken a cautious approach to deciding whether to invest heavily or not.

3.2.1. Law

Fiat money, on the other hand, is safe to use because it is regulated by a country's central bank. The central bank has complete control over all policies and the outcome of a country's monetary policy. When it comes to bitcoin, everyone can create many accounts at no expense. There are no centralised verification procedures in place, and it is not mandatory for people to use their true names (Böhme et al., 2015). This procedure is a little hazy, and the idea of illicit activity underlying all of the bitcoin registration and trading could be a scam in some way. Being anonymous on the internet provides the ideal environment for crooks and fraudsters to carry out their schemes. This trade platform would be used by cybercriminals to carry out their illegal activities, including scamming and cheating. According to Kethineni et al. (2017), cryptocurrency is more likely to be utilised by criminals in money laundering and drug trafficking schemes. Despite the fact that blockchain technology was created to make life easier for users all around the world, criminals will always find a way to profit.

Previously, several regulatory agencies, such as those in China, had refused to endorse Bitcoin as a money (Cheung et al., 2015). China has outlawed the use of Bitcoin and other digital currencies in financial institutions and other forms of commerce. Because bitcoin trading and commercial activities cannot be traced on its trading platform, as well as the anonymity of the persons involved, the authority's move is understandable. Despite the fact that some countries favour the usage of digital currency, China may have outlawed it due to its growing economy and status as one of the world's economic superpowers.

3.2.2. Bills

Aside from the initial cost of purchasing the hardware, a miner's other major expense is energy consumption (Hayes, 2017). It has been revealed that mining the digital currency has eaten up greater electrical bills compared to the rewards offered by solving a block (O'Dwyer & Malone, 2014). Cryptocurrency mining has consumed a lot of electricity. The cost of mining is different from the performance of the hardware. It is reported that the output of electricity from mining cryptocurrencies ranges from 10MW \s(equivalent to a small power plant) to 3-6 GW (the estimated energy consumed by small to medium size country such as Bangladesh and Denmark) (Vranken, 2017). (Vranken, 2017). Vranken (2017) emphasised the long-term viability of cryptocurrencies[5]. He discovered that mining these digital currencies via proof of work consumes a lot of energy and takes a lot of computing power. Nonetheless, these advanced processors, which include CPUs and GPUs, are required in blockchain mining to

ISBN 978-81-949892-5-7

prevent double spending, which is a security concern. Within the next decade, mining activities are likely to decline, and only those with a significant amount of up-to-date hardware and the ability to reduce electricity consumption costs would be able to survive.

The cost of mining cryptocurrency was summarised by Becker et al. (2013). The great majority of these currencies use proof of work, which necessitates a considerable level of power consumption due to the mathematical work performed by the hardware involved[6]. This is particularly dangerous in large-scale mining operations. As a result, mining cryptocurrency will become the villain, contributing to carbon dioxide emissions and threatening to destroy the planet through global warming. More research into the impact of cryptocurrency on the environment is needed. It's not worth sacrificing the planet for a quick buck. If it is demonstrated that mining will cause more harm than good, nations and even the United Nations should interfere to ensure that the environment is not harmed.

3.2.3. Crash and Bubble

An efficient market, according to Fama (1970), is one in which prior information is accessible that can fully reflect the prices of the market's history. Cryptocurrency is considered a weak kind of commodity because investors are unable to foresee future prospects due to the lack of historical data (Urquhart, 2016). This is especially true given that cryptocurrency was first introduced in 2009, over a decade ago. An investment made in such a short period of time is unlikely to have a track record, and investors cannot rely on it to be lucrative[11]. According to Fry and Cheah (2015) and Urquhart (2016), cryptocurrency would be less volatile if it had a real form of account and a way to store value. For example, there is a possibility of crashes and bubbles. In the not-too-distant future, cryptocurrency is expected to reach its peak. Despite this, there had been no significant bubble that would eventually devalue Bitcoin or any other cryptocurrency.

The monthly average volatility returns for cryptocurrency, such as Bitcoin, are substantially higher than gold. The highest monthly volatilities for gold and other currencies, on the other hand, are higher than the lowest monthly volatilities for Bitcoin (Dwyer, 2015). This volatility in Bitcoin suggests that cryptocurrency would be a risky long-term investment. According to Dwyer, this history of volatility gives the chance for a bubble and crash to occur (2015). Cheung et al. (2015) found that between 2011 and 2013, Bitcoin saw three major bubble bursts, each lasting 66 to 106 days. The Mt Gox exchange has been hit by the largest scandal in this bubble disaster (Yermack, 2013).

Previous study has demonstrated that speculation can destabilise assets (Blau, 2018). Bitcoin's price volatility indicates that it is fueled by trade that is tainted by speculation. Its viability as a currency may be jeopardised by speculation. Bitcoin's price has climbed from a few pennies in its early trading days to $1,132.26 by the end of 2013. After a few months, the price had dropped by about 60%. (Blau, 2018). This was a telltale symptom of a bursting asset bubble. Due to the small number of people who use Bitcoin as their primary cryptocurrency today, determining its fair value is challenging . To trade Bitcoin without paying interest, you don't need an account. There are only about 9000 retailers worldwide who accept Bitcoin as a form of payment (Yermack, 2013)[12]. This shaky implementation of Bitcoin could lead to scams and other schemes that result in monetary losses. Investors are looking for ways to benefit from cryptocurrencies, hoping that it will protect them from the perils of speculation (Li et al., 2018). As evidenced by the modest Bitcoin bubble burst in multiple situations listed above, bubbles are expected to form when authorities and economic policy intervene by not promoting cryptocurrency.

3.2.4. Attack on network

According to Kshetri (2017), blockchain technology has low vulnerability and security due to its decentralised nature. Manipulation and forgery are now possible. In the context of the Internet of Things, blockchain technology faces numerous hurdles in terms of its identity and access control system (IoT). Pool formation mining activities are exposed to two types of attacks. Malicious pool members or pool operators are both to blame. By combining the resources in their pool, hostile pool operators can launch a Sybil attack on the network. While malicious pool members can theoretically boost the processing power in a mining pool, they can also destabilise it in the future16].

ISBN 978-81-949892-5-7

These individuals jump from one pool to the next in order to disrupt the pools mining returns and prevent the mined block from being effective (Conte De Leon et al., 2017).

Another flaw of cryptocurrencies is that it is vulnerable to code-based attacks. The network's coding, which was created by the originator, Nakamoto, is vulnerable to bug attacks. This network is now managed by a core group on Github as an open source project. In June 2013, an unknown attacker assaulted Bitcoin nodes on their path, relaying information on the network that was not involved in mining activity (Bradbury, 2013). As history has shown, future attacks on the blockchain network are unavoidable. Despite their success thus far, fraudsters will eventually find a way to attack the blockchain's cryptography network if this vulnerability is not addressed seriously.

4.Discussion

Since the beginning of cryptocurrency's existence, detractors have been raising their voices. The FBI shut down the Silk Road affair, which led to the accusation of cryptocurrencies being used for unlawful purposes. Silk Road was a popular Bitcoin-based trading platform. Silk Road was accused of being a marketplace for drug and other criminal operations. However, according to Alstyne (2014), the Silk Road's closure was not due to Blockchain or cryptocurrencies. Instead, like with any other platform that might benefit them, it was thieves and fraudsters who exploited this technology for their own economic motives. The second example is Mt Gox, which has lost money totaling 350 million dollars (McMillan, 2014).

One would think that time is a major element before a system is fully established and becomes a reliable and durable technology. Cryptocurrency will take some time to be adopted and used by the general public around the world. Even Paypal, the electronic payment system that was created in 1999, has been a target for scammers on multiple occasions. It had been attacked at least five times after becoming completely resilient and establishing itself as the world's most efficient and trustworthy electronic payment system (Jackson & Grey, 2014).In the history of Bitcoin, bubbles have been caused by speculation, while price drops have been caused by government and central monetary agencies intervening. As a result, in the public's interest in expecting the benefits of cryptocurrencies, the government should establish policies and regulations that protect the public's interest as well as the interests of major economic players. A stable market would ensure that a country's fiscal policy can be balanced without relying on central bank intervention. All of this is contingent on how the government responds to the current cryptocurrency market, whether it favours or demonises it for good.

5.Applications of Bitcoin

As more and more establishments around the world accept bitcoin payments, bitcoins are being used to buy products and services. Bitcoin transactions offer a customizable amount of secrecy, and their trail is relatively tough to follow. As a result, bitcoins are being utilised to conduct anonymous transactions. Because bitcoins are not tied to any country or governed by any government, international transfers are simple and inexpensive. There is the liberty of not needing authorization from any government in order to carry out your transactions. They utilise very strong cryptographic techniques to give a secure way to transact online. Bitcoin payments are popular among users and businesses since there are no credit card fees to pay. It might be as an investment, with the expectation of a considerable increase in value in the future. It can be used to gamble on websites such as SatoshiDice, RoyalBitcoin, Bitzino, and Peerbet, among others. Bitcoins are being used to make online purchases as a growing number of merchants accept bitcoin transactions. Bitcoin wallet apps allow users to make payments in bitcoins using their smartphones. Unlike credit card or bank payments, no personal information is required to complete the purchase. As a result, the inconvenience of proving one's identification can be eliminated.

6.Improvement and Future Work on Crypto currencies

The advent of cryptocurrencies will undoubtedly play a big part in the global economic fabric. It is a fact that every economist, researcher, and investor must act and take significant steps to improve their understanding of blockchain technology in general (Fauzi et al., 2019; Fauzi et al., 2018a). Because bitcoin has not yet reached maturity in terms of time, more research into its technology, potential, and risks should be conducted to assure that the opportunities

ISBN 978-81-949892-5-7

are not only coincidental. Furthermore, the impending difficulties do not protect stakeholders from financial failures[17].

Research into decreasing the 51 percent attack on the blockchain mining network should be expanded in the future (Shi, 2016). In terms of securing the customer's monetary assets, the security protocol should be better, if not identical, to that of a traditional centralised banking system. Users' security merits a ground-breaking testimonial from stakeholders in this new business, so that the blockchain technology's confidence and trust can become the norm for users when conducting daily transactions over the internet.

A proof of stake system would use less energy to mine these digital currencies, lowering the cost of mining the money (Vranken, 2017). A person must validate the coins they own and the amount they have under proof of stake process[20]. The individual must construct a transaction involving their coins, which they must send to their account as a reward, along with information about a certain proportion. Proof of stake is similar to a raffle system in that all miners have the same chance. Furthermore, a hybrid technique combining both proof of work and proof of stake has been proposed, in which a percentage of the proof of work is awarded to all active nodes, while the stake controls the number of lottery tickets obtained. Proof of Activity (PoA) is a concept proposed by Bentove et al. (2014) that combines proof of work and proof of stake. The term "activity" in PoA refers to active users who maintain the entire online node and who should be rewarded. Offline users, on the other hand, can still acquire coins in proof of stake, which can lead to double spending of the same block. PoA provides significantly stronger protection in the face of potential cryptocurrency threats. It has more storage space and allows for low-cost network communication. Furthermore, PoA offers minimal transaction fees, uses less energy, and the network topology can be improved. Thus, compared to proof of work, the PoA alternative serves as a better platform for cryptocurrencies due to its capacity to avoid double spending and, most crucially, the cost of purchasing the coin.

Many new cryptocurrencies have already entered the market, and there are many more waiting to be released. In terms of price and market capitalization, various emerging currencies have challenged and competed with Bitcoin. Even though only Ethereum and Ripple had achieved three-figure prices at the time of writing. On the emergent cryptocurrency, Bornholdt and Sneppen (2014) created a model dubbed the Moran process. The model simulates the market in which currencies are traded. The model can mimic the constant rate of new coins being mined, the trading operations of the agents, and the communication among market users (Cocco et al., 2017)[6]. With the acknowledgement of the account holders, all currencies are interchangeable among themselves. It was also discovered that Bitcoin may be traded with other money, and the widely regarded Bitcoin may be replaced in the future by other exciting coins with better qualities. As a result, examining attributes such as security, return on investment, and cheap mining costs can help identify which of the new and upcoming digital coins will eventually replace Bitcoin.

Becker et al. discussed future work on leveraging the proof of work (2013). One of them is to reuse the byproduct of the proof of work to get the most out of it. This byproduct can be reused in the sense that a resource used to solve a mathematical puzzle by an already awarded user can be used to reward another user from the formulated solved problem. Another option is to convert the electrical energy created by the mining operation into heat energy[9]. This can be accomplished in cold climate countries, where the significant amount of heat energy created during the computation of solving the mathematical puzzle can be used to heat residential houses and other heat-intensive household tasks.

According to Dyhrberg (2016), Bitcoin can be used to hedge against the financial stock market and US currencies. It was proposed that, similar to gold, which does not rely on a central authority, Bitcoin can reduce market risk. Bitcoin's active trading frequency among users supports it in the short run by demonstrating hedging capabilities. This suggests that Bitcoin has a clear and promising future in the world stock market portfolio research, as well as low risk management. Dyhrberg (2016) goes on to say that Bitcoin and gold are effective tools for mitigating investment risk[10]. Although this statement is perhaps a little premature in pricing Bitcoin on par with gold, the potential is there, and it opens the door to further investigation.

ISBN 978-81-949892-5-7

It has been acknowledged that having a thorough understanding of blockchain technology will be critical in limiting the negative effects of using cryptocurrencies in everyday activities (Fauzi, 2019). As a result, experts in this subject should work with policymakers and government agencies to develop legislation and policies governing a country's use of cryptocurrencies. To help individuals comprehend the potential and risks of utilising cryptocurrency, knowledge management among industry actors and researchers should be improved. Even professionals from higher education institutions should communicate with the public since they have knowledge resources that can help the community have a better understanding of particular situations (Fauzi et al., 2018b).

7.Conclusion

Cryptocurrencies aren't going away anytime soon. The future of trade is bright, thanks to new emerging technologies that can help humanity. Users and industry participants can, of course, assess whether Bitcoin can help or hurt them based on their goals and expectations for possessing it. This article looked at the advantages of cryptocurrencies in terms of technology security, low transaction costs, and high investment returns. Law and regulation, excessive energy usage, the likelihood of a crash and bubble, and network attacks were all discussed as challenges. Improving the security protocol, working on proof of activity, employing the byproduct of proof of work, and implementing the knowledge management system are among the improvements and future work on cryptocurrencies. More in-depth studies on numerous elements of cryptocurrencies should be conducted, given the favourable outlook of blockchain technology and the likelihood of government regulation. Taking advantage of opportunities in cryptocurrencies and blockchain technology might be advantageous for researchers. The implementation of cryptocurrency to the best of its capacity would then become one of the most significant discoveries of the twenty-first century.

References

[1] Ali, R., Barrdear, J., Clews, R., & Southgate, J. (2014). The economics of digital currencies. *Bank of England Quarterly Bulletin*, 276-286. https://papers.ssrn.com/sol3/ papers. cfm?abstract_id=2499418

[2] Androulaki, E., Karame, G. O., Roeschlin, M., Scherer, T., & Capkun, S. (2013, April). Evaluating user privacy in bitcoin. In: *International Conference on Financial Cryptography and Data Security* (pp. 34-51). Berlin/Heidelberg, Germany: Springer.

[3] Angel, J. J., & McCabe, D. (2015). The ethics of payments: Paper, plastic, or Bitcoin? *Journal of Business Ethics*, *132*(3), 603-611 . DOI 10.1007/s10551-014-2354-x

[4] Bariviera, A. F., Basgall, M. J., Hasperué, W., & Naiouf, M. (2017). Some stylized facts of the Bitcoin market. Physica A: *Statistical Mechanics and its Applications*, *484*, 82-90. https:// doi.org/10.1016/j.physa.2017.04.159

[5] Becker, J., Breuker, D., Heide, T., Holler, J., Rauer, H. P., & Böhme, R. (2013). Can we afford integrity by proof-of-work? Scenarios inspired by the Bitcoin currency. In: *The Economics of Information Security and Privacy* (pp. 135-156). Berlin/ Heidelberg, Germany: Springer.

[6] Bentov, I., Lee, C., Mizrahi, A., & Rosenfeld, M. (2014). Proof of Activity: Extending Bitcoin's Proof of Work via Proof of Stake. *ACM SIGMETRICS Performance Evaluation Review*, *42*(3), 34-37.

[7] Bitcoin Chart (2018). Bitcoin chart. Retrieved April 5, 2018 from: https://charts.bitcoin.com/

[8] Blau, B. M. (2018). Price dynamics and speculative trading in Bitcoin. *Research in International Business and Finance*, *43*, 15-21 . http://dx.doi.org/10.1016/j.ribaf. 2017.07.183

[9] Böhme, R., Christin, N., Edelman, B., & Moore, T. (2015). Bitcoin: Economics, technology, and governance. *The Journal*
of Economic Perspectives, *29*(2), 213-238. http://dx.doi. org/10.1257/jep.29.2.213

[10] Bornholdt, S., & Sneppen, K. (2014). Do Bitcoins make the world go round? On the dynamics of competing crypto-currencies. arXiv preprint arXiv:1403.6378. https://arxiv.org/abs/1403.6378

ISBN 978-81-949892-5-7

[11] Bouri, E., Gupta, R., Tiwari, A. K., & Roubaud, D. (2017). Does Bitcoin hedge global uncertainty? Evidence from waveletbased quantile-in-quantile regressions. *Finance ResearchLetters*,*23*,87-95.

[12] Bradbury, D. (2013). The problem with Bitcoin. *Computer Fraud & Security*, *2013*(11), 5-8.

[13] Brière, M., Oosterlinck, K., & Szafarz, A. (2015). Virtual currency, tangible return: Portfolio diversification with bitcoin. *Journal of Asset Management*, *16*(6), 365-373.

[14] Cocco, L., Concas, G., & Marchesi, M. (2017). Using an artificial financial market for studying a cryptocurrency market. *Journal of Economic Interaction and Coordination*, *12*(2), 345-365. DOI 10.1007/s11403-015-0168-2

[15] Cheung, A., Roca, E., & Su, J. J. (2015). Crypto-currency bubbles: an application of the Phillips–Shi–Yu (2013) methodology on Mt. Gox bitcoin prices. *Applied Economics*, *47*(23), 2348-2358.

[16] Ciaian, P., Rajcaniova, M., & Kancs, D. A. (2016). The economics of Bitcoin price formation. *Applied Economics*, *48*(19), 1799-1815.

[17] Conte de Leon, D., Stalick, A. Q., Jillepalli, A. A., Haney, M. A., & Sheldon, F. T. (2017). Blockchain: properties and misconceptions. *Asia Pacific Journal of Innovation and Entrepreneurship*, *11*(3), 286-300

[18] Dos Santos, R. P. (2017). On the Philosophy of Bitcoin/Blockchain Technology: Is it a Chaotic, Complex System? *Metaphilosophy*, *48*(5), 620-633.

[19] Dyhrberg, A. H. (2016). Hedging capabilities of bitcoin. Is it the virtual gold? *Finance Research Letters*, *16*, 139-144. http:// dx.doi.org/10.1016/j.frl.2015.10.025

[20] Dwyer, G. P. (2015). The economics of Bitcoin and similar private digital currencies. *Journal of Financial Stability*, *17*, 81-91.

[21] Eyal, I., & Sirer, E. G. (2014, March). Majority is not enough: Bitcoin mining is vulnerable. In: *International conference on financial cryptography and data security* (pp. 436-454). Berlin/ Heidelberg, Germany: Springer.

[22] Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, *25*(2), 383-417 [23] Fauzi, M. A., Tan, C. N. L., & Ramayah, T. (2018a). Knowledge sharing intention at Malaysian higher learning institutions: The academics' viewpoint. *Knowledge Management & E-Learning: An International Journal*, *10*(2), 163-176.

[24] Fauzi, M. A., Nya-Ling, C. T., Thurasamy, R., & Ojo, A. O. (2018b). An integrative model of knowledge sharing in Malaysian higher learning institute. *Kybernetes. 47*(5), 1031-1052.

[25] Fauzi, M. A., Tan, C. N. L., Thurasamy, R., & Ojo, A. O. (2019). Evaluating academics' knowledge sharing intentions in Malaysian public universities. *Malaysian Journal of Library & Information Science*, *24*(1) , 123-143. https://doi.org/10.22452/ mjlis. vol24no1.7

[26] Fauzi, M. A. (2019). Knowledge sharing in Asia Pacific via virtual community platform: a systematic review. *International Journal of Web Based Communities*, *15*(4), 368-394. DOI: 10.1504 /IJWBC. 2019.10023490

[27] Fry, J., & Cheah, E. T. (2016). Negative bubbles and shocks in cryptocurrency markets. *International Review of Financial Analysis*, *47*, 343-352. http://dx.doi.org /10.1016/j. irfa.2016.02.008

[28] Ha, S., & Moon, B. R. (2018). Finding attractive technical patterns in cryptocurrency markets. *Memetic Computing*, *10*(3), 301-306 .

ISBN 978-81-949892-5-7

The Nature and Application of Problem Based Learning in Academic Environments

Sarithadevi S
Asst Professor, Department of Computer Science
Naipunnya Institute of Management and Information Technology, Kerala,India
*

Abstract-
Problem based learning (PBL) is an approach to education in which students are presented with problems to solve, which acts as a stimulus for learning. It is considered an alternative to conventional mode of instruction and is suited for those domains which value applied knowledge rather than theoretical knowledge. In fields like medicine and engineering, a professional must be able to apply the academic knowledge he/she has acquired to solve real world problems. PBL is not exclusive to tertiary education and can be great tool to engage the intellectual competencies of students at the primary and secondary level. In today's skill-driven job market, problem based learning deserves greater attention and adoption.

Index Terms- problem based learning, self-directed learning, group learning, education

I. INTRODUCTION

Problem Based Learning (PBL) is originated in 1960s in the medical school of McMaster University. Problem based learning is a term used in teaching and learning environment. Instead of traditional chalk and board method of education system, problem based learning focus on sharing complex problem to students and can ask them to solve the problem. By sharing a problem with students, it becomes a problem of the entire class including teacher and students. This particular scenario can be compared with real life situation. For example, personal problem of a husband/wife can be shared with all other members of the family. Likewise in research, it is good to share the problem with research guide or to the peer groups. By sharing the problem to a group of people, everyone can analyse and try to find out a solution to the problem together.

By sharing the problem with others, you may get different ideas to solve the problem. Knowledge about the problem is very important to solve the particular problem. If you have good knowledge/understanding about the problem, then it is said to be half solved. In PBL method, students can be divided into small groups and the problem can be assigned to each group for finding the solution. The problems are never subject specific, they can from the areas of science, mathematics, social science, etc. The real life problems may be associated with topics like weather, poverty, unemployment, health, education etc. Since our student community lives along with all these problems in the present situation, they may able to find their own solutions to these real life problems.

Given below are examples of problem based learning assignments in science -

1) Some researchers found that we can use certain types of bacteria to clean up radioactive pollution in water. But not every scientist agrees with that. So, some methodology can be used by students to analyse the data on bacteria, and decide whether bacteria is acting as pollutant eater or not. Finally, they can present their work on this particular problem.

ISBN 978-81-949892-5-7

2) Little children can also be tiny problem solvers. At Two Rivers Public School, Washington DC, the first grade students are given tiny task to find about spider expedition. First, the teachers asked the students to draw the diagram of spider. Then, each student wrote a story of their own spider. They analysed the parts of their own spider. The little children were trying to find out whether the spiders are dangerous or not.

II. Characteristics of Problem Based Learning

It's the foremost preliminary step for proceeding with any research work writing. While doing this go through a complete thought process of your Journal subject and research for its viability by following means:

A. PBL is student centric and experimental

In PBL, teachers act as facilitators or mentors. Students have to involve in the learning process by themselves with the intension of problem solving.

B. Solutions to many authentic problems

By using the concept of PBL, many real world, authentic problems can be solved in the areas of medicine, engineering, administration, etc..

C. Student's pre existing knowledge can be taken into consideration

The facilitator can form student teams in such way that whoever is good at certain area, they can be assigned to solve problems of that specific area.

D. Active involvement of students are needed in all PBL stages

Active participation of students is needed in all stages of PBL. Students should involve in all stages like planning, organising and evaluate their work.

E. New information can be collected from peer groups

Students can share their learning and work experiences with their peer groups. This will lead to self-directed learning among student groups.

F. New knowledge can be accomplished

In PBL method, the problems can be considered as a tool to attain the essential knowledge. The analytical and research skill of students increases by solving the real life problems. Their communicative as well as decision making skills improve.

III. Advantages of PBL

A. Interdisciplinary

There may be students from different subject domain working on a problem. Therefore, knowledge of different disciplines can be blended together for learning new facts.

B. Keen learning

Students can go ahead as life long learners. They don't need to memorise things and cognitive skills of the problem solvers increases because of continuous brain activity.

ISBN 978-81-949892-5-7

C. Transferrable skills of students increases

The transmissible skill of students increases by solving real life problems. The transferrable skills acquired by them becomes important in their career as well since employers always selects those people who can survive in new environments.

D. Enhancing teamwork

Working as a group helps the students to communicate and collaborate well with others. Their communication and listening skill improve, responsibility and honesty becomes their strength during the journey of their life.

IV. Disadvantages of PBL

A. Poor performance on tests

Dedicating too much time in problem based learning, causes poor performance among students in systematic or standard tests. They may not be able to achieve high score for multiple choice and short answer questions.

B. Students unpreparedness

Even if most of the students will be involved in problem based learning, some of them may not be ready to involve in such kind of practices. Their participation may be thwarted due to immaturity, unfamiliarity and lack of prerequisite knowledge.

C. Teacher unpreparedness

Teachers may have to change their teaching styles to implement PBL in curriculum. Teachers should not give false or incorrect statement about the cognitive thinking capability of students as it may demotivate students in finding solutions to real life problems.

D. Time consuming Evaluation

The teacher has to choose a correct evaluation method for the entire process. She/He may have to assess each student in the group by closely recording their individual contribution towards problem solving.

E. Varying degrees of relevancy and applicability

It is tough to find out a problem that students can solve with their content and skills. Sometimes it may feel that PBL is too hard or difficult.

V. Problem Based Learning As an approach to teach a topic in Software Engineering

“Principles of Software Engineering” is an important course that is part of the syllabi of undergraduate programs in Computer Applications or Computer Science. This course mainly focuses on the theoretical aspects of software development. According to the curriculum provided by the University of Calicut, this particular subject consists of five modules, in which the learning outcome of the first module is to understand different software life cycle models, i.e., what are the different approaches for software development. The theory parts of many software development methods like waterfall model, spiral model, increment model, agile model, extreme programming model, etc., are taught in this module.

An important topic in ‘Software Engineering’ is SDLC (Software Development Life Cycle). This SDLC consists of five different stages. They are –

1) Planning
2) Define requirements
3) Development
4) Testing

ISBN 978-81-949892-5-7

5) Deployment and maintenance

Module 2 to Module 5 consists of the theory part of the above five steps. The teaching style can be changed by implementing PBL on the subject wherever possible.

The adoption of problem based learning can be better understood with an example, detailed in subsections A and B.

A. Real life problem

Open course selection by students in college is a time consuming and very hectic process. Open course is a subject offered by one individual department to other department students. For example, department students from other departments can join for the courses offered by the Computer Science department. Computer Science department can provide four different courses to other department students. Likewise, each department can offer their courses to students from other departments. So, keeping track of the list of students and assigning the students to the particular teacher is always found confusing. The problem can be stated as - How a computerized information system can help to solve the problem of open course selection in the college? The concept of the topic SDLC (Software Development Life Cycle) can be used to solve the above problem.

B. Method

In class of 35 students, they can be divided into 7 groups of 5 each. Each group can come up with their own solutions to the above problem. Firstly, students may need a deeper understanding of the SDLC concept as well as the given problem. They may go to each department in college, and have to note down the requirements of the department. They can draw data flow diagram in the designing phase. Development stage is done via programming the software. After that they can test the software they developed, once it is tested against all bugs, software can be implemented.

A mentor can be assigned to each group for monitoring their work. They are free to choose any SDLC model for their work. Finally, each group can present their solutions to the open course selection problem. Assessment of all the 35 students can be done by setting evaluation criteria.

Each team can be assessed based on the following criteria -

Specification Check	Method Used	Effectiveness	Delivery Time	Documentation	Quality of the software

Each student can be assessed based on the following criteria –

Role in the team (E.g., Planner, designer, developer, tester)	Involvement in the problem	Response from other students about his/her contribution	Understandability of the problem by the student	% of contribution to the team	Any innovative idea from the student to solve the problem

VI. PBL - FUTURE SCOPE

The universities in India should make changes in their curriculum structure so that PBL method can be incorporated into teaching and learning process. This will require reduction of content in the syllabus of each programme. By reducing the content of syllabus, students will able to get enough time to solve the real life problems associated with the topics they learn. Training should be given to teachers to meet the requirements of PBL method of teaching. Financial level services should also be provided from the administrative level to colleges. Teachers, students and

ISBN 978-81-949892-5-7

management should work together to implement the concept of problem based learning in colleges.

REFERENCES

[1] S. Kurt, “Problem-based learning (PBL),” Educational Technology, 08-Jan-2020. [Online]. Available: https://educationaltechnology.net/problem-based-learning-pbl/.

[2] “Solving real-world issues through problem-based learning,” Edutopia, 01-Nov-2016. [Online]. Available: https://www.edutopia.org/video/solving-real-world-issues-through-problem-based-learning.

[3] “Defining characteristics of PBL - University of Iowa.” [Online]. Available: https://teach.its.uiowa.edu/sites/teach.its.uiowa.edu/files/docs/docs/Defining_Characteristics_of_PBL_ed.pdf.

[4] “Transferable skills: Definitions and examples,” Indeed Career Guide. [Online]. Available: https://www.indeed.com/career-advice/resumes-cover-letters/transferable-skills.

[5] “12 top job skills for any industry: Transferable skills you need,” Indeed Career Guide. [Online]. Available: https://www.indeed.com/career-advice/career-development/top-job-skills.

[6] P. Game, “5 advantages and disadvantages of problem-based learning [+ activity design steps],” Prodigy Education. [Online]. Available: https://www.prodigygame.com/main-en/blog/advantages-disadvantages-problem-based-learning/.

[7] R. Abdo, “10+ project based learning examples for educators,” Venngage, 17-Nov-2021. [Online]. Available: https://venngage.com/blog/project-based-learning-examples/.

[8] “Problem based learning in the software engineering classroom.” [Online]. Available: https://www.researchgate.net/publication/224400629_Problem_Based_Learning_in_the_Software_Engineering_Classroom.

[9] Google search. [Online]. Available: https://www.google.com/search?q=assessment%2Bof%2Bstudents%2Bperformance&source=lnms&.

ISBN 978-81-949892-5-7

Digital Signature in ITR filing

Kesiya Johnson,
Student of MSC
Department of Computer science science
Naipunnya College, Thrissur, India ,India

Dr.Sarika S,
Assistant Professor
Department of Computer

Naipunnya College ,Thrissur

Abstract-

Everyday new technologies are brought and progressed very rapid in all fields. Now new generation gifted to tax payers for filing their income tax go back via on-line is E-submitting. The E-filing is the new effective method of filing income tax return through online and make E-payment tax with digital signature. It saves our golden time,strength,fee and also reduces our anxiety. So, the tax payers are required to use E-filing centres. This present look at examines that the existing customers are satisfied with the E-filing facilities but most of the people tax payers are not aware to the E-submitting procedures so sufficient steps are required for create greater awareness within the minds of tax payers regarding Esubmitting of earnings tax with the aid of using digital signature.

Keywords—taxpayers, E-filing, digital signature, tax

I. INTRODUCTION

The advanced technology is seen everywhere, from e - ticking to efiling the tax return, everything can be done easily at the comfort of your home. While filing an income tax return online a requirement that you have to furnish is to affix your digital signature with your tax return documents to authenticate these docs. In the IT Act 2000, a digital signature enjoys the identical status as a normal signature. It attests and verifies that the taxpayer has authenticated the tax return documents in secure surroundings, without fraud. Virtual or Digital signatures, that are issued by Certification authorities, contain particulars just like the taxpayer's name, public key, name of issuing Certification Authority, expiration date of public key(12years), the digital signature and its serial wide variety. Tampering with digitally signed files and claiming forgery over digital signatures isn't a viable option, especially since some assessments are nearing completion to confirm the same. Changes and additions to digitally signed files are also included in the signing process.

II. CRYPTOGRAPHY

Cryptography is a method of defensive facts and communications through the usage of codes, so that only those for whom the information is intended can read and process it. In computer science, cryptography refers to comfy data and conversation techniques derived from mathematical concepts and a fixed of rule-based calculations called algorithms, to transform messages in ways which might be difficult to decipher. these deterministic algorithms are used for cryptographic key technology, digital signing, verification to defend data privateness, web surfing at the internet and private communications which includes credit card transactions and email.

A. Objectives of cryptography

1.Confidentiality:The information cannot be understood by anyone for whom it was accidental.

ISBN 978-81-949892-5-7

2.Integrity: The statistics can't be altered in storage or transit between sender and supposed receiver without the alteration being detected.
3.Non-repudiation: The creator/sender of the information cannot deny at a later degree their intentions in the advent or transmission of the facts.
4.Authentication: The sender and receiver can affirm every other's identity and the starting place/vacation spot of the statistics

.

B. Types of cryptography

i. Symmetric-key encryption algorithms create a fixed length of bits known as a block cipher with a secret key that the sender uses to encrypt data (encryption) and the receiver uses to decrypt it(decryption). This only required a single key for both encryption and decryption process. Block and Stream algorithms comprise symmetric key cryptography, which is widely used on the Internet today. Two popular encryption algorithms are the Advanced Encryption Standard (AES) and the Data Encryption Standard (DES). This method of encryption is often faster than symmetric encryption, but it allows both the sender and the data receiver to have access to the secret key.

ii Asymmetric key algorithms: It use a pair of keys, a public key associated with the sender for encrypting messages and a private key that only the receiver knows for decrypting that information. This required two key one to encryption and the other one to decryption. When someone wants to send an encrypted message, they will retrieve the recipient's public key from a shared directory and use it to encrypt the message until it is sent. The receiver will next use their associated private key to decrypt the message. When the sender encrypts a message with their private key, the message can only be decrypted with the sender's public key, allowing the sender to be authenticated. These encryption and decryption operations are fully automated, so users do not have to manually lock and unlock messages.

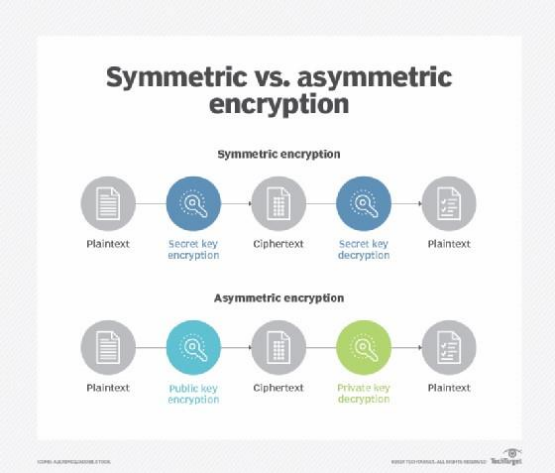


Fig.1 Types of cryptography

III. DIGITAL SIGNATURE

A mathematical algorithm is routinely used to validate the authenticity and integrity of a message using a digital signature, which is a type of electronic signature (e.g., an email, a credit card transaction, or a digital

ISBN 978-81-949892-5-7

document). Digital signatures are used to identify users and protect information in digital messages and documents by creating a virtual fingerprint that is unique to them. The email content becomes part of the digital signature in emails. Electronic signatures such as digital signatures are far more secure than other types of electronic signatures. Digital signatures can offer proof of origin, identity and standing of digital files, transactions or digital messages. Signers also can use them to renowned knowledgeable consent. In many countries, including the United States, digital signatures are considered legally binding inside the identical manner as conventional handwritten record signatures.

A. How DSC works

Digital signatures are based on asymmetric key cryptography or public key cryptography. Using a public key algorithm, such as RSA algorithm, two keys are generated, creating a linked pair of keys ,one private key and one public key. Digital signatures work through asymmetric key cryptography's two mutually authenticating cryptographic keys. The individual who creates the DS uses a private key to encrypt signature-related data, the only way to decrypt that data is with the signer's public key. If the recipient cannot open the document with the signer's public key, that's a sign there's a problem with the document or the signature. This is how digital signatures are authenticated or valid.

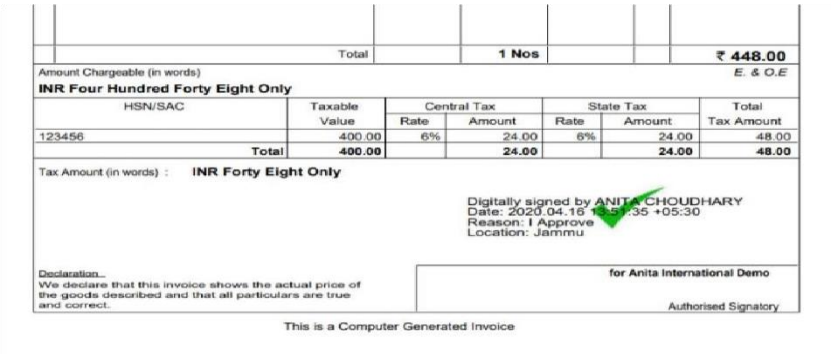


Fig 2 DSC verification

B. Different classes of Digital signature

- Class 1 Certificate: These are issued to individuals or private users. This Certificate confirms that the user's name and e-mail ID are valid and approved by the Certifying Authorities on their database.
- Class 2 Certificate: These are issued only to business personnel and individuals. They confirmed that the information in the application provided by the subscriber is the same as the information in popular consumer databases.
- Class 3 Certificate: These are issued only to individuals and organizations. They are very high assurance certificates, mainly for the purpose of e-commerce applications. It is issued when the individual appears in-person before the certifying authorities.

C. Benefits of digital signature

1.A digital signature can't be edited or tampered with.
2. It is secure to track a digitally signed document.
3. It brings down the wastage of paper and is an eco- friendly.
4. Helps the efficiency of the entire e-filing process.
5. Reduces cost

ISBN 978-81-949892-5-7

TABLE 1 DSC COMPARISON

	Digital Signature
Visible	No
Unobtrusive	Yes
File changes	Not Allowed
Virtually attached to file	Yes
Physically embedded in file	No
Data authenticity	Yes
Copyright protection	Yes
Global identification	partial

D. Certifying authorities for Digital signature

The licensed certifying authorities who authorized by government appointed Controller of Certifying Authority:

- 1. Safe script
- 2. Capricorn CA
- 3. IDRBT
- 4. GNFC
- 5. e Mudra CA
- 6. NSDL e-Gov CA
- 7. Indian Air Force
- 8. Verasys CA
- 9. CDAC CA

.

E. How to get digital signature

The purpose of obtaining a digital certificate, the user will have to submit certain documents to the certifying authority (CA). It includes an application form that has been duly signed, a passport size photo an identification proof, Aadhaar card number, PAN card verification etc. The applicant may be asked to provide the mobile number, email address and home or organization address of the user. The different countries will have different requirements from the applicants for the issuance of digital signature certificate. The process of obtaining digital signature certificates varies depending on the certifying authorities.

ISBN 978-81-949892-5-7

F. Mandatory taxpayers for ITR filing using DSC

Digital signature certificates are mandatory for some services / user categories such as e-Verification of returns filed by political parties and companies as well as other persons whose accounts are required to be audited under Section 44AB of the Income Tax Act. In other case, it is optional.

.

G. Steps to create DSC in ITR filing

1: Fill up the Income Tax Return form, generate the file as an XML (Extensible Mark-up Language) file and save it.
2: Step in to the Income Tax India website. Log in to your account using your user password and ID.
3: After login, click on the tab "Submit Return" and then select the assessment year.
4: Select the Income Tax Return Form Name from the drop-down menu list.
5: The next field will be "Do You Want to Digitally Sign the File?"
Then select the "Yes" button.
6: Select the type of digital signature you want to use, it can be
"Sign with USB Token" or "Sign With .PFX file"
7: Upload the ITR with the help of digital signature certificate and verify it.

Fig 3 USB token

.

H. Current problem in taking DSC

As previously stated, the signature is signed with the USB token after the digital signature certificate has been verified and approved by certified authorities. The password will be included in the DSc. There's a chance you'll lose your USB token. If it is lost, other people or hackers can quickly track down the clients DSc by targeting all of their authentication information. This is a fairly rare problem in this field.

IV. PROPOSED SYSTEM

The problem mentioned above is an example of a threat. We can use an OTP password with the USB token to get around this situation. If someone tried to access the USB Token, they could easily access the password that has been attached to it. It may also request an OTP verification in addition to the password. Only the client's mobile number will receive the OTP password. As a result, this will provide a security mechanism which useful in future attacks.

ISBN 978-81-949892-5-7

V. CONCLUSION

From the above study, the usage of ITR filing using DSC must be focused to make a better way of using online method in a developing country like India. Still the usage of DSC is increasing day by day among the citizens for its secure techniques. This study focuses on the digital signature and its authentication process. The basic objective of research is to provide an awareness about DSC and its basic attacks. The majority of individuals are unaware of this problem. The paper aids them in raising awareness about the problem. This case study will be implemented in my future research

REFERENCES

[1] B. A. Fourazan, DebdeepMukhopadhyay, "Cryptography and Network Security", Tata McGraw Hill, 2nd edition,pp.15,210234,2010Katz, Jonathan, and Yehuda Lindell. Introduction to modern cryptography. CRC press, 2020.

[2] Alqad, Ziad, et al. "A New Approach for Data Cryptography." International Journal of Computer Science and Mobile Computing 8.9 (2019): 30-48.

[3] Curry, Ian. "An Introduction to Cryptography and Digital Signatures." Entrust Securing Digital Identities and Information (2001).

[4] Aysu, Aydin, Bilgiday Yuce, and Patrick Schaumont. "The future of real-time security: Latency-optimized lattice-based digital signatures." ACM Transactions on Embedded Computing Systems (TECS) 14.3 (2015): 1-18.

[5]Alfred M., Oorschot P., and Vanstone S., Handbook of Applied Cryptography, CRC press, 1997

ISBN 978-81-949892-5-7

A SOLUTION FOR CREDIT CARD FRAUD USING MACHINE LEARNING

Anna Helna
Student, M.Sc. computer science
Department of Computer Science
Naipunnya Institute of Management and Information Technology,
Pangram, Koratty East, Thrissur
helnasteephan.19@gmail.com

Laiby Thomas
Assistant professor,
Department of Computer Science
Naipunnya Institute of Management and Information Technology,
Pangram, Koratty East, Thrissur
laiby@naipunnya.ac.in

Abstract

This Credit card fraud is becoming a common problem today. Digital world made people to relay on credit cards and debit cards and it in turn became a threat. Machine learning is a new approach to find a solution for this problem. It can automatically detect fraud characteristics. By using a person's historical data, the user's pattern and behavior can be analyzed and can determine whether the transaction is fraudulent or not. There are many approaches are there for detecting the credit card fraud. Random Forest Algorithm, K-Nearest Neighbored and K-Means Clustering are among them. In this paper, various techniques analyzed and compared for finding the best solution.

Keywords—Machine Learning, Credit Card, Fraud Detection

I. INTRODUCTION

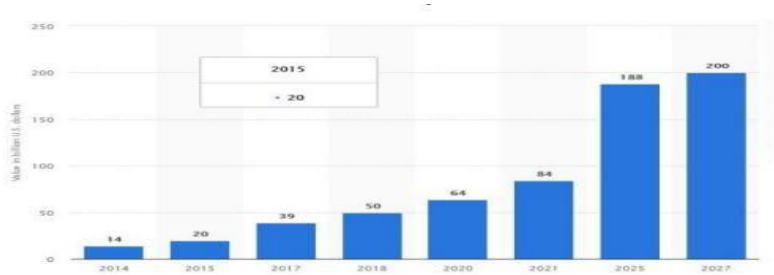
The people in the twenty first century live in a digitalized world that comprises with the knowledge of computer for the ease of living. The money transactions became digitalized and there came the use of credit cards. Credit card is a electronic payment system that are used for the non-cash transactions. The bank or financial institution issue credit card to a consumer to facilitate payment to a merchant of goods and services [1]. Credit score card typically refers to a card this is assigned to the purchaser (cardholder), commonly permitting them to purchase items and services inside credit restrict or withdraw coins in advance. Credit card presents the cardholder and gain of time, i. e., it presents time for their clients to pay off later in a prescribed time, by using sporting it to the subsequent billing cycle. [2]

A credit card is considered fraudulent when another person uses your credit card for you without your authorization. Fraudulent steal the credit card PIN code or account details to perform one of the unauthorized transactions without robbery of the original physical card. Using credit card fraud detection, we could find out if new transactions are fraudulent or bona fide [3].

In 2017, there had been 1,579 facts breaches and nearly 179 million statistics amongst which credit card frauds have been the most common shape with 133,1/2reviews, then employment or tax-related fraud with 82,051 reports, telephone frauds with 55,1/2reports followed through bank frauds with 50,517 reviews from the statics released via FTC. [2]

The fraud this is devoted may involve the card which include a credit score card or debit card. In this, the cardboard itself acts as a fraudulent source within the transaction. The reason of committing the crime may be to achieve the goods without paying cash or to gain the unauthorized fund. Credit score cards are a nice goal for fraud. The motive is that in a completely quick time a variety of money can be earned without taking many risks or even the crime will take many weeks to be detected [3].

ISBN 978-81-949892-5-7



Growth of E-Commerce sites [3]

As the usage of the internet these days [fig.1] may be very a good deal growing there may be many possibilities for the fraudsters to commit the fraudsters to commit the credit card frauds. The principles fraud instances which are in those of the e-commerce websites. In the gift generation, human beings are displaying a good deal hobby in getting matters on-line instead of going and purchasing them, and due to this, the growth of the e-trade websites is growing and thereby there's a huge hazard of credit card fraud. With the intention to keep away from such credit score card frauds, we want to discover the first-rate set of rules that reduces credit card frauds [3].

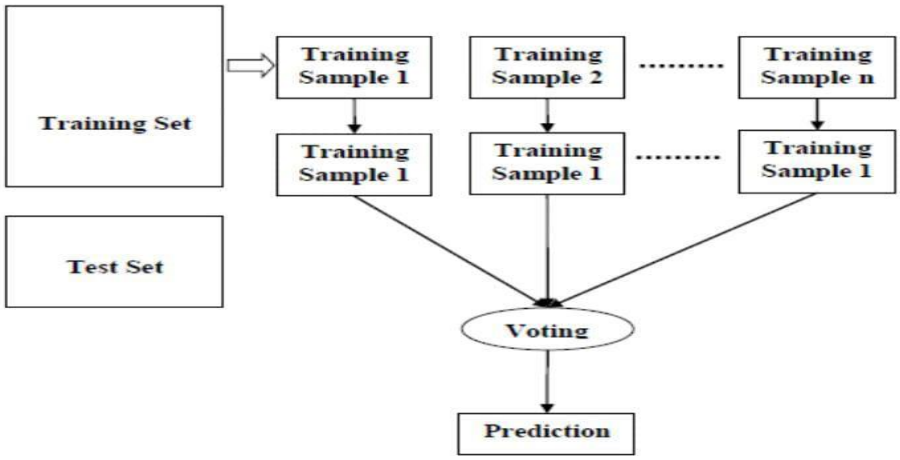
II. METHODS FOR DETECTING FRAUD

Methods for credit card fraud detection with quite a few studies strategies and numerous fraud detection strategies with a special interest in the neural networks, facts mining, and allocated statistics mining. Many different techniques are used to locate such credit score card fraud. Whilst accomplished the literature survey on various strategies of credit card fraud detection, we able to conclude that to hit upon credit score card fraud there are numerous other methods in machine getting to know itself.

Random Forest Algorithm

Random forest algorithm is one of the broadly used supervised learning algorithms. This could be used for each regression and class purposes. However, this set of rules is especially used for classification problems. Normally, a forest is made from bushes and in addition, the Random forest algorithm creates the selection trees on the pattern facts and receives the prediction from every of the pattern facts. Then Random forest of rules is better than decision trees due to the fact it reduces the over-becoming by using averaging the end result. [3]

ISBN 978-81-949892-5-7



Random Forest Algorithm

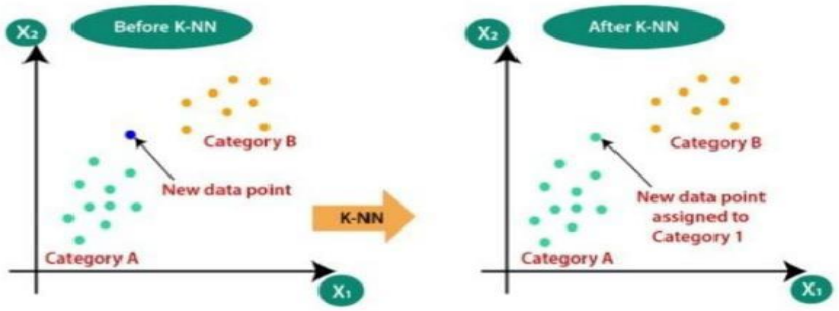
Steps of Random Forest Algorithm

- Assume the Kaggle praise card extortion dataset that is prepared and arbitrarily select a portion of the example information.
- Utilizing the arbitrarily made example information presently makes the Decision trees that are utilized to characterize the cases into the misrepresentation and non-extortion cases.
- The Decision trees are framed by dividing the hubs; the hubs which have the most noteworthy data gain make it as the root hub and characterize the misrepresentation and non-extortion cases.
- Presently the larger part vote is performed and the Decision trees might bring about 0 as result which incorporates that these are the non-extortion cases.
- At last, we track down the exactness, accuracy, review, and F1 – score for both the extortion and non-misrepresentation cases.

K-Nearest Neighbor

A basic, simple to-carry out managed AI strategy that utilizations ordered information to create a work that gives a reasonable result when given extra unlabeled information. Both characterization and relapse issues can be addressed with the K-Nearest Neighbors' (KNN) calculation, which is fast and clear to apply. Utilizes marked information to show a capacity that produces an OK presentation for new information. In the KNN calculation, the similarity between the new case and the case and the cases that are currently arranged is determined. When the new case is set in a class that is most similar to the gathering. In a comparable to design, KNN arranges every single open datum and orders new focuses relying upon how comparative they are. This depicts whenever new information arises, it is simply a question of fitting a k-N characterization plan to it. The calculation gets altogether more slow as indicators/autonomous factors increment. As displayed in the figure underneath: [1]

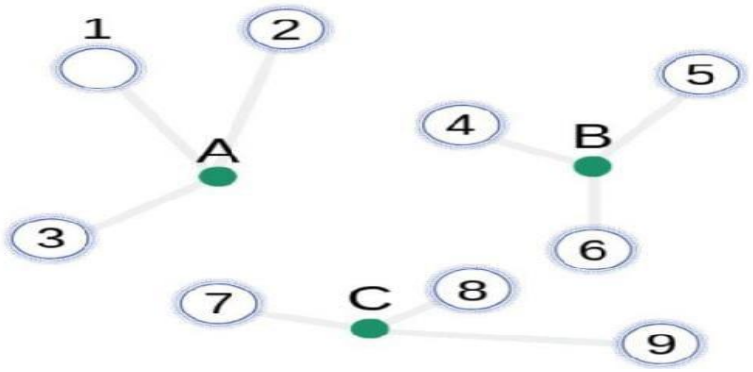
ISBN 978-81-949892-5-7



General structure working of the KNN[1]

K-Means Clusterings

As a result of its effortlessness and adequacy, it is the most broadly utilized unaided learning technique. By computing the mean distance between main informative elements, this strategy assigns focuses to gatherings. It then, at that point, rehashes this cycle to work on the exactness of its classifies over the long haul. The K-Means in the figure beneath are clarified through the following strides: to decide the quantity of groups, pick k. then at that point, pick k areas or centroids at irregular. (it could something else from dataset.) in the following advance: Assign every information highlight the centroid that is nearest to it, shaping the preset K groups. Then, at that point, ascertain the change and reposition each group's centroid. Rehash the third step, reassigning each highlight the group's cutting edge closest centroid. Steps to wrap up: if there is a reassignment, go to stage4: any other way, move to FINISH. The model is done. [1]



General structure working of the K-MC [1]

III. TECHNIQUES COMPARISON

A. Random Forest Algorithm

i. Advantages

- Decreases over fitting in choice trees and assists with working on the exactness
- It is adaptable to both order and relapse issues

ISBN 978-81-949892-5-7

- It functions admirably with both all out and nonstop qualities
- It mechanizes missing qualities present in the information
- Normalizing of information isn't needed as it utilizes a standard based methodology

ii. Disadvantages

- It requires a lot of computational power as well as assets as it fabricates various trees to consolidate their results.
- Additionally, requires a lot of time for preparing as it consolidates a ton of choice trees to decide the class.
- Because of the group of choice trees, it additionally endures interpretability and neglects to decide the meaning of every factor.

B. K-Nearest Neighbor

iii. Advantages

- KNN is called Lazy Learner (instance based learnering). It advances nothing in the preparation period. There is no preparation period. It stores the preparation dataset and gains from it just at the hour of making constant forecasts.
- New information can be added without affecting the calculation execution or precision
- KNN Algorithm is extremely simple to carry out. You really want just two information

iv. Disadvantages

- Execution issue with huge informational collection: the time expected to compute the distance between the new point and each current focuses is immense. This then, at that point, corrupts the presentation of the calculation.
- Doesn't function admirably with high aspects/Features: the KNN calculation doesn't function admirably with huge no of features layered information in light of the fact that with enormous number of features, it becomes challenging for the calculation to ascertain the distance in each features.
- Worth of K: it is truly critical to figure out what worth to appoint to K. with various K you come by various outcomes

C. K-Means Clustering

v. Advantages

- Rehashed method.
- Effective and quick
- Chips away at ordered advanced information.

vi. Disadvantages

- Load of repeats.
- Need to choose you have K worth.
- Should comprehend the instance of your information well.

IV. DISCUSSION OF COMPARISON

Through this paper find the advantages and disadvantages of each algorithm. Each algorithms advantage and disadvantages are different each other. That shows the one is more effective to detect fraudulent transaction.

V. CONCLUSION

Credit card fraud is become a threat to the people in this digitalized world. For detecting credit card fraud, we, Machine Learning Algorithm is used. Random Forest and K-Nearest Neighbor and K-Means Cluster are different approaches that are used for this. These algorithms help the credit card companies to identify the fraudulent transactions more

ISBN 978-81-949892-5-7

accurately within a short time and low cost. All this technique has its own advantages and disadvantages. Based on the company need one can choose any one of these approach. This kind of comparative study will allow the people to build a hybrid approach most accurate for fraudulent credit card transaction detection.

REFERENCES

[1] *Survey Paper On Credit Card Fraud Detection Techniques.* Fayyomi, Aisha Mohammad, Eleyan, Derar and Eleyan, Amina. 09, SEPTEMBER 2021, INTERNATIONAL JOURNAL OF SCIENTIFIC&TECHNOLOGY RESEARCH, Vol. 10, pp. 72-79.

[2] Credit Card Fraud Detection Using Machine Learning Alorithms. Dornadula, Vaishnavi Nath and S, Geetha. Chennai,India : s.n., 2019, INTERNATIONAL CONFERENCE ONRECENT IN ADVANCED COMPUTING, pp. 632-641. *(references)*

[3] Credit Card Fraud Detection Using Machine Learning. Sailusha, Ruttala, et al. 2020, International Conference on Intelligent Computer and Computer System, pp. 1264-1270.

[4] "K-Nearest Neighbor(KNN) Algorithm for Machine Learning – Javatpoint." https://www.javatpoint.com/k-nearest-neighbor-algorithm-for-machine-learning

[5] . https://www.javatpoint.com/k-means-clustering-algorithm-in-machine-learning

[6] . https://www.yadsmic.com/post/k-nearest-neighbors-advantages-and-disadvantages

[7] .https://www.mygreatlearning.com/blog/random-forest-algorithm

[8]https://www.tutorialspoint.com/machine_learning_with_python/machine_learning_with_python_classification_algorithms_random_forest.html

ISBN 978-81-949892-5-7

Exploration of Image Processing Tools in Machine Learning - A Review.

Ms. Nintu Varughese	Ms. Anuja Shaju	Mr.Fredy Varghese
Scholar,	*Scholar,*	*Assistant Professor*
Department of Computer Science	*Department of Computer Science*	*Department of Computer Science*
NIMIT,Pongam,Thrissur,	*NIMIT,Pongam,Thrissur,*	*NIMIT,Pongam,Thrissur,*
nintuvarughese02@gmail.com	*anujamariam6@gmail.com*	*fredy@naipunnya.ac.in*

Abstract:

Image processing is a very important technology, and industrial need appears to be increasing every year. Machine learning image processing first arose in the 1960s as a way to replicate the human vision system and automate the picture analysis process. Solutions for certain activities began to emerge as technology progressed and improved. Many modern image processing approaches use Machine Learning Models like Deep Neural Networks to alter images for a range of objectives, such as applying creative filters, tweaking an image for optimal quality, or improving certain image details for computer vision applications.

I. INTRODUCTION

Image processing is the process of converting an image to a digital format and applying various functions to it in order to create a better image or extract additional information from it. When the input is an image, such as a video frame or image, and the output is an image or attributes associated with that image, it is referred to as a signal time. In most cases, the Image Processing system treats images as two equal symbols while employing set approaches. It is currently one of the fastest expanding technologies, with applications in a variety of industries[1]. Within the engineering and computer science industries, graphic design is at the heart of research..Image processing is a technique for applying operations on an image in order to improve it or extract relevant information from it. It's a sort of signal processing in which the input is an image and the output is either that image or its characteristics/features.

Keywords: Digital image processing (DIP), Geographic Information System (GIS), Biometric Verification, face recognition, image enhancement, signature recognition.

II.IMAGE PROCESSING TOOLS

Basically image processing include different stages. The fundamental tools that are used for digital image processing techniques are briefly discussed as following:

ImageSensors: Image sensors senses the intensity, amplitude, co-ordinates and other features of the images and passes the result to the image processing hardware. It includes the problem domain.

Image Processing Hardware: Image processing hardware is the dedicated hardware that is used to process the instructions obtained from the image sensors[1]. It passes the result to general purpose computer.

Computer: Computer used in the image processing system is the general-purpose computer that is used by us in our daily life.

Image Processing Software: Image processing software is the software that includes all the mechanisms and algorithms that are used in image processing system[3].

ISBN 978-81-949892-5-7

Mass Storage: Mass storage stores the pixels of the images during the processing.
Hard Copy Device: Once the image is processed then it is stored in the hard copy device. It can be a pen drive or any external ROM device.
Image Display: It includes the monitor or display screen that displays the processed images.
Network: Network is the connection of all the above elements of the image processing system.

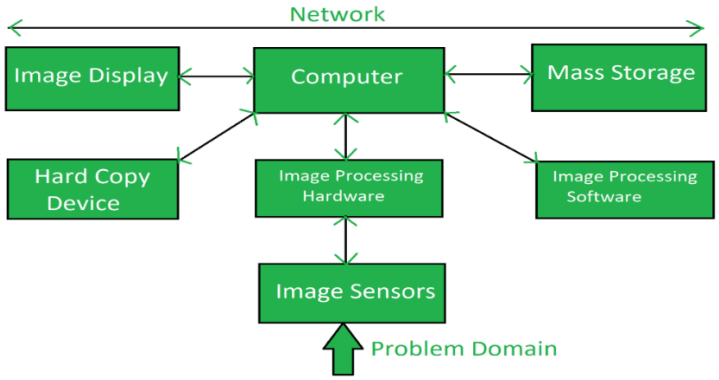
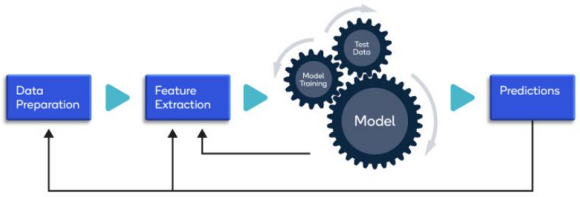


Fig-1 Components of image processing tools:

III. LITERATURE SURVEY

Machine learning algorithms usually have a set of steps or pipelines for learning from data. To learn and predict extremely accurate results, machine learning algorithms require a large volume of high-quality data. As a result, we'll need to make sure the images are well-processed, annotated, and generic enough to be used in machine learning image processing[2]. This is where Computer Vision (CV) comes in; it's a field concerned with machines' ability to comprehend image data. We can analyse, load, transform, and modify photos with CV to create a perfect. dataset for the machine learning algorithm

Fig-2 Methodology of image processing in machine learning



1. Phases of image processing

.**Acquisition**: It could be as simple as being given an image which is in digital form. The main work involves Scaling and Color conversion (RGB to Gray or vice-versa)

Image Enhancement: It is amongst the simplest and most appealing in areas of Image Processing it is also used to extract some hidden details from an image and is subjective.

ISBN 978-81-949892-5-7

Image Restoration: It also deals with appealing of an image but it is objective (Restoration is based on mathematical or probabilistic model or image degradation).

Color Image Processing: It deals with pseudocolor and full color image processing color models are applicable to digital image processing.

Wavelets And Multi-Resolution Processing: It is foundation of representing images in various degrees.

Image Compression: It involves in developing some functions to perform this operation. It mainly deals with image size or resolution.

Morphological Processing: It deals with tools for extracting image components that are useful in the representation & description of shape.

Segmentation Procedure: It includes partitioning an image into its constituent parts or objects. Autonomous segmentation is the most difficult task in Image Processing.

IV. IMAGE PROCESSING PLATFORMS FOR MACHINE LEARNING

There are several types of image processing tools are available. This paper comparing some of the tools of image processing in machine learning in order to discover the best tool which allows to significantly save time and resources and get the best results.

2.1 Matlab: Matlab is an extraordinary tool for making image processing applications and is generally utilized in research as it permits quick prototyping. Millions of engineers and scientists worldwide use MATLAB for a range of applications, in industry and academia, including deep learning and machine learning, signal processing and communications, image and video processing, control systems, test and measurement, computational finance, and computational biology.The basic data element of MATLAB as the name suggests is the Matrix or an Array[3]. MATLAB toolboxes are professionally built and enable you to turn your imaginations into reality. MATLAB programming is quite similar to C programming and just requires a little brush up of your basic programming skills to start working with. Some applications of MATLAB are Curve fitting, Control systems, Signal Processing, Mapping, Deep learning.

2.2 Scikit-image: scikit-image (formerly scikits. image) is an open-source image processing library for the Python programming language. It includes algorithms for segmentation, geometric transformations, color space manipulation, analysis, filtering, morphology, feature detection, and more. Some applications are Financial cybersecurity analytics, Product development, Neuroimaging, Barcode scanner development.

2.3 Caffe: Caffe (Convolutional Architecture for Fast Feature Embedding) is a deep learning framework focused on solving the problem of image classification and segmentation. Some applications are Caffe is being used in academic research projects, Start-up prototypes, Large-scale industrial applications in vision, Industrial applications in speech, Industrial applications in multimedia[8].

2. 4 Mahotas: Mahotas is a computer vision and image processing library for Python. It includes many algorithms implemented in C++ for speed while operating in numpy arrays and with a very clean Python interface. Mahotas currently has over 100 functions for image processing and computer vision and it keeps growing.Applications of Mahotas: Watershed, Convex points calculations , Zernike & Haralick, local binary patterns and TAS features.

2.5 OpenCV: OpenCV is a great tool for image processing and performing computer vision tasks. It is an open-source library that can be used to perform tasks like face detection, objection tracking, landmark detection, and much more.

ISBN 978-81-949892-5-7

It supports multiple languages including python, java C++. Applications of OpenCV: Automated inspection and surveillance, Robot and driver-less car navigation and control, Medical image analysis[5].

2.6 Simple ITK: Simple ITK is an open-source, cross-platform system that provides developers with an extensive suite of software tools for image analysis through which we can get the characteristic of the image. It is intended to facilitate its use in rapid prototyping, education, interpreted languages. It is one of the image processing libraries which is available in multiple programming languages including C++, Python, R, Java, c#, Lua, Ruby, and Tcl. As it is available for multiple programming.

2.7 Pgmagick: Pgmagick is an open-source python library and a wrapper for GraphicsMagick, which is a robust collection of tools and libraries to read, write, and manipulate an image. It supports around 88 formats of image. Image Processing is used to extract useful information from an Image or an Image dataset.

Comparison table

Basis	Matlab	Sci-kit	Caffe	Mahotas	OpenCV	SimpleITK	Pgmagick
Accessibility	Not open source, not free	Open source	Open source	Open Source	Open source	Open source	Open source.
Execution Speed	3 times faster as compared to Mahotas.	Lesser compared to caffe.	Faster than sci-kit	Lesser than sci-kit.	Faster compared to Matlab.	Lesser than OpenCV	Lesser than OpenCV
Quality	Quality of on-going product is lesser than sci-kit.	Quality of on-going product is better than Matlab.	Quality of on-going product is moderate	Quality of on-going product is better than SimpleITK	Quality of on-going product is better.	Quality of on-going product is lesser than OpenCV	Quality of an on-going product is lesser than OpenCV.

Table-1 Machine Learning Tools

V.CONCLUSION

The success, efficiency of execution, and quality of projects may depend on many factors, but choosing the right tools is one of the most important – it allows to significantly save time and resources and get the best results. OpenCV and Matlab are discovered as the best image processing tools during the comparison done with other major tools in the basis of Accessibility, Execution speed and Product quality.MATLAB is more convenient in developing and data presentation, however, OpenCV is much faster in execution. In the case of OpenCV, the speed ratio reaches more than 80 in some cases.However, OpenCV is comparatively harder to learn due to a lack of documentation and error handling.OpenCV is a free and open-source software, while on the other hand MATLAB is a licensed software which is pretty costly. According to the findings of this research OpenCV is better than all the other tools mentioned in this paper.

ISBN 978-81-949892-5-7

VI. REFERENCES

[1] A.A. Thakre, A. V. Lad, and K. Mala, "Measurements of tool wear parameters using machine vision system," Modelling and Simulation in Engineering, vol. 2019, 2019.
[2] R. G. Lins and S. N. Givigi, "Automatic crack detection and measurement based on image analysis," IEEE Transactions on Instrumentation and Measurement, vol. 65, no. 3, pp. 583–590, 2016
[3] J. Taylor, "Carbide cutting tool variance and breakage: unknown factors in machining economics," Advances in Machine Tool Design and Research, pp. 487–504, 2016.
[4] Y.-t. Kao, S.-c. Chang, S. Dauzere-Peres, and J. Blue, "Opportunity for improving fab effectiveness by predictive overall equipment effectiveness (poee)," in e-Manufacturing and Design Collaboration Symposium (eMDC), 2016. IEEE, 2016, pp. 1–4
[5] L. Cerce, F. Pusavec, and J. Kopac, "3d cutting tool-wear monitoring in ˇ the process," Journal of Mechanical Science and Technology, vol. 29, no. 9, pp. 3885–3895, 2015
[6] L. Bao, Y. Song, Q. Yang, H. Yuan, and G. Wang, "Tree filter: efficient structure-preserving smoothing with a minimum spanning tree," IEEE T. on Image Processing, vol. 23, no. 2, pp. 555-569, 2014.
[7] J. Kowalczuk, E. T. Psota, and L. C. Perez, "Real-time stereo matching on CUDA using an iterative refinement method for adaptive supportweight correspondences," IEEE Trans. Circuits and Systems for Video Technology, vol. 23, no. 1, pp. 94-104, 2013
[8] C. Cho and S. Lee, "Low-Complexity Topological Derivative-Based Segmentation," IEEE Trans. on Image Processing, vol. 24, no. 2, pp. 734-741, 2015
[9] Y. Duan, H. Chang, W. Huang, and J. Zhou, "Simultaneous bias correction and image segmentation via L0 regularized Mumford-Shar model," in Proc. IEEE ICIP, pp. 6-10, 2014.
[10] Guo songlu Research on thermal analysis and thermal error compensation of high-speed machining center[D]. Zhengzhou university,2014.
[11] Guan bo. Research on tool geometric parameter image measurement technology [D]. Taiyuan: north university of China,2015.
[12] Zhang chuanchuan. Canny operator edge extraction research [D]. Taiyuan: north university of China,2015
[13] Jung, Franck, Detecting building changes from multitemporal aerial stereopairs. ISPRS Journal of Photogrammetry and Remote Sensing, 2004. 58(3-4): p. 187-201.
[14] Patel, Krishna Kumar, Abhijit Kar, and MA Khan, Development and an Application of Computer Vision System for Nondestructive Physical Characterization of Mangoes. Agricultural Research: p. 1-16.
[15] Esposito, Floriana, Donato Malerba, and Francesca A Lisi, Machine learning for intelligent processing of printed documents. Journal of Intelligent Information Systems, 2000. 14(2-3): p. 175-198.
[16] White, David, James D Dunn, Alexandra C Schmid, and Richard I Kemp, Error rates in users of automatic face recognition software. PloS one, 2015. 10(10): p. e0139827.
[17] Steger, Carsten, Markus Ulrich, and Christian Wiedemann, Machine vision algorithms and applications. 2018: John Wiley & Sons.
[18] Dash, Tirtharaj and Tanistha Nayak, English Character Recognition using Artificial Neural Network. arXiv preprint arXiv:1306.4621, 2013.
[19] Huval, Brody, Tao Wang, Sameep Tandon, Jeff Kiske, Will Song, Joel Pazhayampallil, Mykhaylo Andriluka, Pranav Rajpurkar, Toki Migimatsu, and Royce Cheng-Yue, An empirical evaluation of deep learning on highway driving. arXiv preprint arXiv:1504.01716, 2015.
[20] Garg, Shree, Ankush K Singh, Anil K Sarje, and Sateesh K Peddoju. Behaviour analysis of machine learning algorithms for detecting P2P botnets. in 2013 15th International Conference on Advanced Computing Technologies (ICACT). 2013: IEEE.

ISBN 978-81-949892-5-7

Artificial Neural Networks in Agriculture

Deepak K.V , Ms.Laiby Thomas
Assistant Professor, Naipunnya institute of information Technology, Pongam, Kerala
*d*eepak@naipunnya.ac.in *, laiby@naipunnya.ac.in*

Abstract
The application of Artificial Neural Network (ANN) has been evident in the agricultural sector recently. The sector faces numerous challenges in order to maximize its yield including improper soil treatment, diseases and pest infestation, big data requirements, low output, and knowledge gap between farmers and technology. The main concept of ANN in agriculture is its ability to be flexible, accurate, cost-effective and efficient. Crop yield prediction (CYP) is vital in role to address the ever growing demand for food, requirements of burgeoning world population and to prevent world starvation . ANN can offer an effective and practical solution for the problem. The area of Machine Learning (ML) and Deep learning (DL) have been evaluated. This paper presents a review of the application of AI in Species management, Field condition management, crop management and livestock management.

Index Terms— Species management, Field conditions management, crop management and livestock management, artificial neural network, Crop yield prediction, Machine Learning, Deep learning

I. INTRODUCTION

Agriculture is the bedrock of sustainability of any economy [1]. It plays a key part in long term economic growth and structural transformation [2], [3], [4], though, it may vary by countries [5]. In the past, agricultural activities were limited to food and crop production [6]. Agriculture is the mainstay of all economic activity and stability. Self-sufficiency in agricultural production gives a nation the pride of being able to feed its population. Unfortunately, agriculture is the most vulnerable profession, since it is affected by unforeseen factors like heavy rain, lack of rain, insect attacks [3], and inherent qualities of the soil and policies of the Government. Hence, prediction on agricultural production on an annual basis is tough, but worthy of effort since it will give the farmers a foresight of what they could expect in the next year.

An Artificial Neural Network is an information processing model that is inspired by the way biological nervous systems, such as the brain, process information. They are loosely modeled after the neuronal structure of the mammalian cerebral cortex but on much smaller scales. In simpler terms, it is a simple mathematical model of the brain which is used to process nonlinear relationships between inputs and outputs in parallel like a human brain does every second.ANN is the result of extensive research carried out in the field of Artificial Intelligence (AI) and machine learning problems around the world [2]. The tool easily adopts itself to the prediction of mathematical relations, recognition of particular patterns, error correction and noise removal in signals. Hence, the application of ANN for predicting agricultural production in Tamilnadu might enable estimation of the yield levels of crops as a function the annual rainfall levels in the previous years [6].Data mining is a process of extracting/equating the meaningful data from the large database by using different tools and techniques. Data mining

(sometimes called information or knowledge discovery) [3] is the process of evaluating data from different outlooks and summarizing it into useful information - information that can be used to increase revenue, cuts costs, or both. Data mining software is one of a number of diagnostic tools for analyzing data. It allows users to investigate data from many different dimensions or angles, categorize it, and summarize the relationships identified [4] [5]. Technically, data mining is the process of finding correlations or forms among dozens of fields in large relational databases. Data mining or knowledge discovery in databases (KDD) is an interdisciplinary field where we integrate techniques from different fields including data base systems, statistics, mathematics, high performance computing, artificial intelligence and machine learning.

ISBN 978-81-949892-5-7

ANN in agriculture is a very recent research topic. It consists in the application of data mining techniques to agriculture. The ANN techniques used in agriculture for prediction of problem, disease detection, optimizing the pesticide and so on [7] .Recent technologies are nowadays able to provide a lot of information on agricultural-related activities, which can then be analyzed in order to find important information and to collect relevant information. This ANN techniques are used for disease detection, pattern recognition by using multiple application. ANN is about to identify the similarities between searching the valuable business information from the large database systems such as finding linked products in gigabytes of store scanner data or the mining a mountain for a vein of valuable dataset. Both kind of processes required either shifting through an immense amount of material, or to perform the search intelligently so that exactly match will be performed. ANN can be done on a database whose size and quality are sufficient [2] [5].

II. LITERATURE REVIEW

After a through background work, some of the most valuable recent documents and papers are,ANNs are computer programs designed to simulate just the way the human brain processes information. In other words, they are the digitized models of the human brain [8]. a. According to [9], the ANN are considered to be the best procedures for extracting information from imprecise and non-linear data. ANN techniques have turned out to be a very vital tool for a wide variety of applications across many disciplines, including crop production prediction.

B. J I ET AL [44] developed agricultural management need simple and accurate estimation techniques to predict rice yields in the planning process.

B.A. Smith et al [45] discuss year-round air temperature prediction models were developed for prediction horizons of 1 to 12 h using Ward-style ANNs. These models were intended for use in general decision support.

D.L. Ehret et al [46] introduce all crop attributes responded in much the same way to individual climatic factors.

The authors in their paper [47] predicted crop yield on the basis of fertilizer. By using the artificial neural network, they have predicted the best fertilizer and suggested best fitted crop according to soil features for crop yield purposes.

In paper [48], authors emphasis on color images of plant leaf to detect early stage of the disease. For this purpose, the authors used Support Vector Machine (SVM) on the visual symptoms of plants where plant diseased areas are showed as spots, stains or strikes.

In this research paper [49], the authors worked on plant and weed distinguishing technique via Machine Learning and image processing technique to successfully manage weed separation from the crops.

III. ANN IN AGRICULTURE

A. Introduction to Artificial Neural Network (ANN)

Ever Artificial Neural Network is a replica of the biological neural network, which helps in the several fields like function approximation, pattern recognition (finger print identification, face recognition etc.) and error correction (removal of noise from incoming signal)[3]. ANN is organized in the form of neurons grouped into layers. Usually, there are two layers reserved for taking input data and sending the computed results. There may be some hidden layers, lying between the input layer and the output layer. Some of the essential terms related to ANN are presented here.

ANNs are inspired by the human brain functionality and represent a simplified model of the structure of the biological neural network emulating complex functions such as pattern generation, cognition, learning, and decision making. Such models are typically used for regression and classification tasks which prove their usefulness in crop management and detection of weeds, diseases, or specific characteristics. The recent development of ANNs into deep learning that has expanded the scope of ANN application in all domains, including agriculture.

The Artificial Neural Networks ability to learn so quickly is what makes them so powerful and useful for a variety of tasks. But how do they learn? Information flows through a neural network in two different ways. When the model is learning (being trained) or operating normally (after being trained either being used or tested), patterns

ISBN 978-81-949892-5-7

of information from the dataset are being fed into the network via the input neurons, which trigger the layers of hidden neurons, and these in turn arrive at the output neurons. This is called a feed forward network. Not all neurons "fire" all the time. Each neuron receives inputs from the neurons to its left, and the inputs are multiplied by the weights of the connections they travel along. Every neuron adds up all the inputs it receives in this way and (this is the simplest neural network) if the sum is more than a certain threshold value, the neuron "fires" and triggers the neurons it's connected to (the neurons on its right).

An artificial neural network takes its structure and working concept from human nervous system. Artificial Neural Networks (ANN) are multi-layer fully-connected neural nets that look like the figure below. They consist of an input layer, multiple hidden layers, and an output layer. Every node in one layer is connected to every other node in the next layer. The earliest artificial neural network called as perceptron is based on biological neuron. It has an input layer and an output layer, the input layer is directly connected to the output layer. It can be used for very simple tasks like classifying linearly separable patterns. In order to solve problems having higher complexity layers of neurons between input and output layers were introduced called as hidden layer. So, the basic work of a neural network is to take the input, process it and forward it as output to the following layer. Fig 1 gives the architecture of a neural network. Each neuron in the network first takes the input, perform summation of the weighted input, apply the activation function, calculate the output and forward it to the following layer.

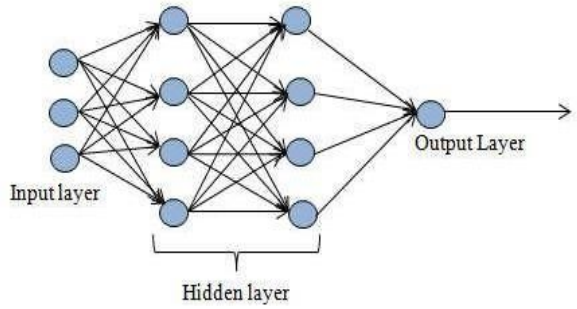


Fig 1. Neural Network Architecture

B. Applications of ANN in Agriculture

Although agricultural practice is broad, this research considers soil, crop, disease and weeds as major contributors to agricultural production. It is paramount to review the application of ANN to agriculture in respect to soil, crop, diseases and pest management. Let's discover how agriculture can benefit from Machine Learning at every stage:

i) Species management

•**Species Breeding:** Our favorite, this application is so logical and yet so unexpected, because mostly you read about harvest prediction or ambient conditions management at later stages.

•**Species selection** :It is a tedious process of searching for specific genes that determine the effectiveness of water and nutrients use, adaptation to climate change, disease resistance, as well as nutrients content or a better taste. Machine learning, in particular, deep learning algorithms, take decades of field data to analyze crops performance in various climates and new characteristics developed in the process. Based on this data they can build a probability model that would predict which genes will most likely contribute a beneficial trait to a plant.

•**Species Recognition**: While the traditional human approach for plant classification would be to compare color and shape of leaves, machine learning can provide more accurate and faster results analyzing the leaf vein morphology which carries more information about the leaf properties.

ii) Field conditions management

•**Soil management**: For specialists involved in agriculture, soil is a heterogeneous natural resource, with complex processes and vague mechanisms. Its temperature alone can give insights into the climate change effects on the

ISBN 978-81-949892-5-7

regional yield. Machine learning algorithms study evaporation processes, soil moisture and temperature to understand the dynamics of ecosystems and the impingement in agriculture. The application of organic materials is essential to improve soil quality [12]. Production of vegetables and other edible crops is often significantly affected by several soil-borne pathogens that require control through soil management [13]. Sensitivity to soil degradation is implicit in the assessment of the sustainability of land management practices, with recognition of the fact that soils vary in their ability to resist change and recover [14]. An artificial neural network (ANN) model predicts soil texture (sand, clay and silt contents) based on attributes obtained from existing coarse resolution soil maps combined with hydrographic parameters derived from a digital elevation model (DEM) [15]. The dynamics of soil moisture are characterized and estimated by a remote sensing device embedded in a higher- order neural network (HONN) [16].

•**Water Management:** Water management in agriculture impacts hydrological, climatological, and agronomical balance. So far, the most developed ML-based applications are connected with estimation of daily, weekly, or monthly evapotranspiration allowing for a more effective use of irrigation systems and prediction of daily dew point temperature, which helps identify expected weather phenomena and estimate evapotranspiration and evaporation.

•**Yield Prediction**:Yield prediction is one of the most important and popular topics in precision agriculture as it defines yield mapping and estimation, matching of crop supply with demand, and crop management. State-of the art approaches have gone far beyond simple prediction based on the historical data, but incorporate computer vision technologies to provide data on the go and comprehensive multidimensional analysis of crops, weather, and economic conditions to make the most of the yield for farmers and population.

•**Crop Quality**: The accurate detection and classification of crop quality characteristics can increase product price and reduce waste. In comparison with the human experts, machines can make use of seemingly meaningless data and interconnections to reveal new qualities playing role in the overall quality of the crops and to detect them. Crop prediction methodology is used to predict the suitable crop by sensing various soil parameters and parameter related to the atmosphere. Parameters like soil type, PH, nitrogen, phosphate, potassium, organic carbon, calcium, magnesium, sulfur, manganese, copper, iron, depth, temperature, rainfall, humidity [17].

•**Disease Detection:** Both in open-air and greenhouse conditions, the most widely used practice in pest and disease control is to uniformly spray pesticides over the cropping area. To be effective, this approach requires significant amounts of pesticides which results in a high financial and significant environmental cost. ML is used as a part of the general precision agriculture management, where agro-chemicals input is targeted in terms of time, place and affected plants. To have an optimal yield in agricultural harvest, disease control is necessary. Plant and animal diseases are a major limiting factor regarding the increase of yield. Several factors play role in the incubation of these diseases which attack plants and animals, which include genetic, soil type, rain, dry weather, wind, temperature, etc. Due to these factors and the unsteady nature of some diseases causative influence, managing the effects is a big challenge, especially in large scale farming. Table III lists the AI applications in disease management available in the literature. To effectively control diseases and minimize losses, a farmer should adopt an integrated disease control and management model that includes physical, chemical and biological measure [18]. To achieve these is time consuming and not at all that cost effective [19], hence the need for application of AI approach for disease control and management. Explanation block (EB) gives a clear view of the logic followed by the kernel of the expert system [20].

•**Weed Detection:** Apart from diseases, weeds are the most important threats to crop production. The biggest problem in weeds fighting is that they are difficult to detect and discriminate from crops. Computer vision and ML algorithms can improve detection and discrimination of weeds at low cost and with no environmental issues and side effects. In future, these technologies will drive robots that will destroy weeds, minimizing the need for herbicides. Weed consistently reduces the farmers' expected profit and yield [21]. A report confirms a 50% reduction in yield for dried beans and corn crops if weed infestations are not controlled [51]. There is about 48% loss in wheat yield due to weed competition [22, 23].

ANN IN AGRICULTURE – SUMMARY

ISBN 978-81-949892-5-7

Sl no	Paper	Application and Strength
1	[44]	Can predict soil enzyme activity. Accurately predicts and classifies soil structure.
2	[30]	Can predict monthly mean soil temperature
3	[31]	It predicts soil texture
4	[15]	Able to predict soil moisture.
5	[16]	Successfully reports soil texture.
6	[32]	Cost-effective, saves time, has 92% accuracy
7	[33]	Can estimate soil nutrients after erosion.
8	[17]	Predicts crop yeild.
9	[35]	Above 90% success rate in detecting crop nutrition disorder.
10	[36]	Can predict the response of crops to soil moisture and salinity
11	[37]	Can accurately predict rice yield.
12	[20]	Works at a high speed. Can multitask.
13	[38]	95% accuracy
14	[39]	Has above than 90% prediction rate.
15	[40]	High performance. Reduces trial and error.
16	[41]	Cost effective, enhanced performance.
17	[42]	Quickly detects stress in crop that will prompt timely site– specific remedies.
18	[43]	High weed recognition rate with short processing time
19	[50]	Machine vision and artificial neural Network(ANN)procedures were used to estimate live body weight of broiler chickens
20	[51]	The greenhouse conditions were controlled by using artificial neural network (ANN).
21	[52]	A current bottleneck of state-of- the-art machine learning methods for image segmentation in agriculture, e.g. convolutional neural networks (CNNs), is the requirement of large manually annotated datasets on a per-pixel level.
22	[53]	Artificial neural networks and principal components were used to detect surface defects on apples in near infrared images.

ISBN 978-81-949892-5-7

23	[54]	Information and Communication Technologies (ICT) has provided access to data and information, specially with the advent of the Internet, and this led to changes in society. In this context, this paper aims to demonstrate how data can assist agriculture in control of production, using the data available in government databases.
24	[55]	Estimation of fuel consumption in agricultural mechanized operations using artificial neural Networks
25	[56]	Develop artificial neural networks for the estimation of tractor fuel consumption during soil preparation, according to the adopted system.
26	[57]	Identification of the state of maturity of fruits with artificial neural networks
27	[58]	Agricultural Crop Yield Prediction Using Artificial
		Neural Network Approach
28	[59]	Classification of Agricultural Pests Using DWT and Back Propagation Neural Networks
29	[60]	Forecasting of reference evapotranspiration by artificial neural networks
30	[61]	Soil prediction using artificial neural networks and topographic attributes
31	[62]	Work models based on neural networks of the backpropagation type were developed in order to predict the occurrence of frosts from meteorological data such as temperature, relative humidity, cloudiness and wind direction and speed.

IV. CONCLUSION

ANN in agriculture is a very recent research topic. It consists in the application of data mining techniques to agriculture. The ANN techniques used in agriculture for prediction of problem, disease detection, optimizing the pesticide and so on. Recent technologies are nowadays able to provide a lot of information on agricultural-related activities, which can then be analyzed in order to find important information and to collect relevant information. This ANN technique are used for disease detection, pattern recognition by using multiple application. ANN is about to identify the similarities between searching the valuable business information from the large database systems such as finding linked products in gigabytes of store scanner data or the mining a mountain for a vein of valuable dataset. Both kind of processes required either shifting through an immense amount of material, or to perform the search intelligently so that exactly match will be performed. ANN can be done on a database whose size and quality are sufficient.

ISBN 978-81-949892-5-7

V. REFERENCES

[1] Agriculture and Food Science Technology, Vol. 4, No. 4, pp. 343-346, 2013
[2] B. F. Johnston, P. Kilby, Agriculture and Structural Transformation: Economic Strategies in Late-Developing Countries, Oxford University Press, 1975
[3] S. Kuznets, "Modern economic growth: Findings and reflections", American Economic Association, Vol. 63, No. 3, pp. 247–258, 1973
[4] M. Syrquin, "Patterns on Structural Change", in: Handbook of Development Economics, Vol. 1, Elsevier, 1988
[5] R. Dekle, G. Vandenbroucke, "A quantitative analysis of China's structural transformation", Journal of Economic Dynamics and Control, Vol. 36, No. 1, pp. 119-135, 2012
[6] M. Fan, J. Shen, L. Yuan, R. Jiang, X. Chen, W. J. Davies, F. Zhang, "Improving crop productivity and resource use efficiency to ensure food security and environmental quality in China", Journal of Experimental Botany, Vol. 63, No. 1, pp. 13-24, 2012
[7] Chen, H.M., Tsai, K.H., Qi, G.Z., Yang, J.C.S. and Amini, F. (1995), Neural Network for Structural Control, ASCE Jl. of Computing in Civil Engineering, 9(2), 168- 176.
[8] Chao, L.C. and Skibniewski, M.J. (1994), Estimating Construction Productivity: Neural Network Based Approach, ASCE Jl. of Computing in Civil Engineering, 8(2), 234-251
[9] Coulibaly, P., Anctil,F. and Bobee, B. (2000), Daily Reservoir Inflow Forecasting using Artficial Neural Networks with Stopped Training Approach, Elsevier Jl. of Hydrology, 230, 244-257.
[10] Demuth, H., Beale, M. and Hagan, M., (2006), Neural Network Toolbox for use with MATLAB®, The MathWorks, Inc., 3, Apple Hill Drive, Natick, MA 01760-2098, USA
[11] Elazouni, A.M., Nosair, I.A., Mohieldin, Y.A. and Mohamed, A.G. (1995), Estimating Resource Requirements at Conceptual Design Stage Using Neural Networks, ASCE Jl. of Computing in Civil Engineering, 11(4), 217-223
[12] M. Pagliai, N. Vignozzi, S. Pellegrini, "Soil structure and the effect of management practices", Soil and Tillage Research, Vol. 79, No. 2, pp. 131-143, 2004
[13] G. S. Abawi, T. L. Widmer, "Impact of soil health management practices on soilborne pathogens, nematodes and root diseases of vegetable crops",
Applied Soil Ecology, Vol. 15, No. 1, pp. 37-47, 2000
[14] J. K. Syers, Managing Soil for Long-Term Productivity, The Royal Society, 1997
[15] Z. Zhao, T. L. Chow, H. W. Rees, Q. Yang, Z. Xing,
F. R. Meng, "Predict soil texture distributions using an artificial neural network model", Computers and Electronics in Agriculture, Vol. 65, No. 1, pp. 36-48, 2009
[16] Elshorbagy, K. Parasuraman, "On the relevance of using artificial neural networks for estimating soil moisture content", Journal of Hydrology, Vol. 362, No. 1-2, pp. 1-18, 2008
[17] S. S. Snehal, S. V. Sandeep, "Agricultural crop yield prediction using artificial neural network approach", International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering, Vol. 2, No. 1, pp. 683-686, 2014
[18] BEA, Value Added by Industry as a Percentage of Gross Domestic Product, available at: https://apps.bea.gov/iTable/iTable.cfm?ReqID=51&step=1#reqid=51&step=51&isuri=1&5114=a&5102=5, 2018
[19] Weed Science Society of America, Facts About Weeds, available at: http://wssa.net/wp-content/uploads/WSSA-Fact-SheetFinal.pdf
[20] K. Balleda, D. Satyanvesh, N. V. S. S. P. Sampath,
K. T. N. Varma, P. K. Baruah, "Agpest: An Efficient Rule-Based Expert System to Prevent Pest Diseases of Rice & Wheat Crops", 8th International Conference on Intelligent Systems and Control, Coimbatore, India, January 10–11, 2014

ISBN 978-81-949892-5-7

[21] Canadian Journal of Plant Science, Vol. 81, No. 2, pp. 339–342, 2001
[22] M. Khan, N. Haq, "Wheat crop yield loss assessment due to weeds", National Agricultural Research Cen intensification tre, Vol. 18, No. 4, pp. 449–453, 2002
[23] S. Fahad, S. Hussain, B. S. Chauhan, S. Saud, C. Wu, S. Hassan, M. Tanveer, A. Jan, J. Huang, "Weed growth and crop yield loss in wheat as influenced by row spacing and weed emergence times", Crop Protection, Vol. 71, pp. 101–108, 2015
[24] A. N. Rao, S. P. Wani, J. K. Ladha, Weed Management Research in India-An Analysis of the Past and Outlook for Future, ICAR, 2014
[25] A. Datta, H. Ullah, N. Tursun, T. Pornprom, S. Z. Knezevic, B. S. Chauhan, "Managing weeds using crop competition in soybean [Glycine max(L.) Merr.]", Crop Protection, Vol. 95, pp. 60–68, 2017
[26] T. Mruthul, Chemical weed management in sesame (Sesamumindicum L.), MSc Thesis, College of Agriculture, Raichur, University of Agricultural Sciences, 2015
[27] C. J. Swanton, R. Nkoa, R. E. Blackshaw, "Experimental methods for crop-weed competition studies", Weed Science Society of America, Vol. 63, No. 1, pp. 2–11, 2015
[28] P. Jha, V. Kumar, R. K. Godara, B. S. Chauhan, "Weed management using crop competition in the United States: A Review", Crop Protection, Vol. 95, pp. 31–37, 2017
[29] P. Milberg, E. Hallgren, "Yield loss due to weeds in cereals and its large-scale variability in Sweden", Field Crops Research, Vol. 86, No. 2–3, pp. 199–209, 2004
[30] E. R. Levine, D. S. Kimes, V. G. Sigillito, "Classifying soil structure using neural networks", Ecological Modelling, Vol. 92, No. 1, pp. 101-1081996
[31] M. Bilgili, "The use of artificial neural network for forecasting the monthly mean soil temperature in Adana, Turkey", Turkish Journal of Agriculture and Forestry, Vol. 35, No. 1, pp. 83-93, 2011
[32] D. H. Chang, S. Islam, "Estimation of soil physical properties using remote sensing and artificial neural network", Remote Sensing of Enviroment, Vol. 74, No. 3, pp. 534-544, 2000
[33] T. Behrens, H. Forster, T. Scholten, U. Steinrucken, E. D. Spies, M. Goldschmitt, "Digital soil mapping using artificial neural networks", Journal of Plant Nutrition and Soil Science, Vol. 168, No. 1, pp. 21-33, 2005
[34] M. Kim, J. E. Gilley, "Artificial neural network estimation of soil erosion and nutrient concentrations in runoff from land application areas", Computers and Electronics in Agriculture, Vol. 64, No. 2, pp. 268-275, 2008
[35] H. Song, Y. He, "Crop Nutrition Diagnosis Expert System Based on Artificial Neural Networks", 3rd International Conference on Information Technology and Applications, Sydney, Australia, July 4–7, 2005
[36] X. Dai, Z. Huo, H. Wang, "Simulation of response of crop yield to soil moisture and salinity with artificial neural network", Field Crops Research, Vol. 121, No. 3, pp. 441-449, 2011
[37] B. Ji, Y. Sun, S. Yang, J. Wan, "Artificial neural networks for rice yield prediction in mountainous regions", Journal of Agricultural Science, Vol. 145, No. 3, pp. 249-261, 2007
G. Liu, X. Yang, Y. Ge, Y. Miao, "An Artificial Neural Network–Based Expert System for Fruit Tree Disease and Insect Pest Diagnosis", International Conference on Networking, Sensing and Control, Lauderdale, USA, April 23–25, 2006
[38] X. Wang, M. Zhang, J. Zhu, S. Geng, "Spectral prediction of phytophthorainfestans infection on tomatoes using artificial neural network", International Journal of Remote Sensing, Vol. 29, No. 6, pp. 1693- 1706, 2006
[39] A. M. Tobal, S. A. Mokhtar, "Weeds identification using evolutionary artificial intelligence algorithm", Journal of Computer Science, Vol. 10, No. 8, pp. 1355- 1361, 2014
[40] P. Moallem, N. Razmjooy, "A multi-layer perception neural network trained by invasive weed optimization for potato color image segmentation", Trends in Applied Sciences Research, Vol. 7, No. 6, pp. 445-455, 2012
[41] Y. Karimi, S. O. Prasher, R. M. Patel, S. H. Kim, "Application of support vector machine technology for weed and nitrogen stress detection in corn", Computers and Electronics in Agriculture, Vol. 51, No. 1-2, pp. 99-

ISBN 978-81-949892-5-7

109, 2006

[42] C. C. Yang, S. O. Prasher, J. Laundry, H. S. Ramaswamy, "Development of neural networks for weed recognition in corn fields", American Society of Agricultural and Biological Engineers, Vol. 45, No. 3, pp. 859-864, 2002

[43] S. Kolhe, R. Kamal, H. S. Saini, G. K. Gupta, "A web-based intelligent disease-diagnosis system using a new fuzzy-logic based approach for drawing the interferences in crops", Computers and Electronics in Agriculture, Vol. 76, No. 1, pp. 16-27, 2011

[44] B. J I ET AL Artificial neural networks for rice yield prediction in mountainous regions. Journal of Agricultural Science (2007), 145, 249–261. [46] B.A. Smith et al,.Artificial Neural Networks for Automated Yearround Temperature Prediction. Computers and Electronics in Agriculture 68 (2009) 52–6.

[45] D.L. Ehret et al, Neural network modeling of greenhouse tomato yield, growth and water use from automated crop monitoring data. Computers and Electronics in Agriculture 79 (2011) 82–89.

[46] Snehal S. Dahikar ,An Artificial Neural Network Approach for Agricultural Crop Yield Prediction Based on Various Parameters International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE) Volume 4, Issue 1, January 2015

[47] A.Camargo, Image pattern classification for the identification of disease causing agents in plants. Computers and Electronics in AgricultureVolume 66,
Issue 2, May 2009, Pages 121-125

[48] ViNguyenThanh Le1Beniamin Apopei1KamalAlameh1, Effective plant discrimination based on the combination of local binary pattern operators and multiclass support vector machine methods, Information Processing in Agriculture

[49] AmraeiS., Mehdizadeh S. A., Salari S. (2017) Broiler weight estimation based on machine vision and artificial neural network

[50] Alipour M., Mohammad Loghavi(2013)Development and Evaluation of A Comprehensive Greenhouse Climate Control System Using Artificial Neural Network.Universal Journal of Control and Automation 1(1): 10-14, 2013 DOI: 10.13189/ujca.2013.010102

[51] Barth, R., IJsselmuiden, J., Hemming, J., Henten, E.V., 2017b. Synthetic bootstrapping of convolutional neural networks for semantic plantpart segmentation. Computers and Electronics in Agriculture doi:https:
//doi.org/10.1016/j.compag.2017.11.040.

[52] Bennedsen B. S., PetersonD. L., TabbA. (2007)Identiying Apple surface Defects Using PrincipalComponents Analysis and Artificial Neural Networks. Transactions of the ASABE Vol. 50(6): 2257- 2265
:Ciencia y Agricultura (Rev CienAgri), 2016, RevistaCiencia y Agricultura, Vol. 23, págs. 117-132. 0122- 8420

[53] Snehal M. S.Dahikar1, Dr.SandeepV.Rode (2014)Agricultural Crop Yield Prediction Using Artificial Neural Network Approach International Journal of Innovative Research in Electrical, Electronics, Instrumentation and Control Engineering.Vol. 2, Issue 1,January 2014

[54] Gaurav Kandalkar et al, .(2014)Classification of Agricultural Pests Using DWT and Back Propagation Neural Networks.International Journal of Computer Science and Information Technologies, Vol. 5 (3) , 4034- 4037.

ISBN 978-81-949892-5-7

A NOVEL MODEL FOR DETECTION OF OBJECTS IN VARIED AND COMPLEX IMAGES

Livin P Wilson
Assistant Professor
Dept. of Computer Science
Naipunnya Institute of Management and Information Technology, Pongam, Thrissur
livin@naipunnya.ac.in

Abstract: An effective and accurate object detection is very important topic in the advancement of computer vision systems.Development in the area of deep learning technique helps to increase accuracy of object detection.This project is completely based on deep learning.To achieve highly efficient object detection the help of open image data set is used.Open image data set is a advanced type of data set which have many features such as large scale,hierarchical tag system etc. Many strategies such as larger backbone,expert models of heavier classifiers are employed. Tensorflow which is a deep learning framework is used to build a Faster R-CNN architecture for automatically recognizing objects in images.For obtaining high level features Inception Resnet V2 convolutional feature extractor is used.

Keywords— Object detection, Deep learning, Tensorflow, Faster R-CNN

INTRODUCTION

To understand more about a image the location and the concepts of each objects with in a image is more impotent than classifying different images. Now a days there is a great development in the field of deep learning and computer vision this help to make great improvement in object detection. Object detection means it is the recognition, detection and localisation of different objects with in a complex image it makes more helpful for understanding more about the minute things with in a image. Object detection is the finding of instances of real world objects such as faces vehicles or buildings in images or video. Object detection has a wide range of application the simplest example that experience in our daily life is face detection while we are uploading or taking a photos in our mobile phone or SLR cameras.To better understand the visual content we should not only know what is the object so called classification task, however additionally grasp wherever is that the terribly object the so-called location task. The object detection task is to simultaneously provide these two information for a given image.Depending on the pipeline most of the object detection techniques could be divided into two categories one-stage method and two-stage method.Generally the one-stage methods focus on the speed performance while the dominant benefit of the two-stage strategies is that the precision performance.

Here the concentration is on the two-stage methods considering the outstanding precision performance.CNN is the most popular model of deep learning. A CNN architecture is referred to as VGG16. Different layers of CNN is known as a feature map. The feature map of the input layer is a 3D matrix of pixel intensities for different color channels such as RGB.The feature map of internal layer is an induced multi-channel image its pixel can be viewed as a specific feature. Every neuron is connected with a small portion of adjacent neurons from the previous layer. Different types of transformations can be done on feature map such as filtering and pooling. Convolution operation convolutes a filter matrix with the values of a field of neurons and takes a nonlinear function such as sigmoid ReLU

to obtain final responses. Pooling operation, such as max pooling, average pooling, L2- pooling and local contrast normalization, summaries the responses of a receptive field into one value. With an interleave between convolution and pooling, an initial hierarchy is constructed which can be fine-tuned in a supervised manner by adding several layers for visual tasks.According to the tasks the final layer with different activation functions is added to get a conditional probability for each output neuron. And the whole network can be optimized on a function called mean squared error or cross-entropy loss via the stochastic gradient descent method. The typical VGG16 has totally 13 convolutional layers, 3 fully connected layers,3 max-pooling layers and a softmax classification layer.The conv feature maps are produced by convoluting 3*3 filter windows, and feature map resolutions are reduced with 2 stride max-pooling layers. An discretionary check image of identical size as coaching samples may be processed with the trained network. Rescaling or cropping operations could also be required if completely different sizes square measure provided.Hierarchical feature representation which is the multilevel representations from pixel to high-level semantic feature can be learned from data automatically and hidden factors of input data can be disentangled through multi-level nonlinear mappings. Compared with traditional shallow models, a deeper architecture provides an exponentially increased expressive capability. The architecture of CNN provides an opportunity to jointly optimize several related tasks together such as Fast RCNN combines classification and bounding box regression into a multi-task leaning manner.Benefitting from the large learning capacity of deep CNNs due to these benefits CNN has been wide applied into many research fields such as image super-resolution reconstruction, image classification, image retrieval, face recognition, pedestrian detection and video analysis.

Open Images Dataset which contains images and ground-truth annotations for the three tasks is used in this project. Open Images V4 has several attractive characteristics, compared to previously available datasets . The images were collected from Flickr without a predefined list of class names or tags, leading to natural class statistics and avoiding the initial design bias on what should be in the dataset. They were released by the authors under a Creative Commons Attribution (CC-BY) license that allows to share and adapt the material, even commercially; particularly so for models trained on these data, since it makes them more easily usable in any context. Also, we removed those images that appear elsewhere in the internet to reduce bias towards web imagery, favoring complex images containing several objects.Open Images V4 is large scale in terms of images (9;178;275),annotations (30;113;078 image-level labels, 15;440;132 bounding boxes, 374;768 visual relationship triplets) and the number of visual concepts (classes) (19;794 for image-level labels and 600 for bounding boxes).This makes it ideal for pushing the limits of the data-hungry methods that dominate the state of the art. For object detection in particular, the scale of the annotations is unprecedented (15:4 million bounding boxes for 600 categories on 1:9 million images).The number of bounding boxes we provide is more and is greater than the next largest datasets (COCO and ImageNet). Also there are 8 annotated bounding boxes per image on average demonstrating the complexity of the images and the richness of our annotations. We hope this will stimulate research into more sophisticated detection models that will exceed current state-of-the-art performance and will enable assessing more precisely in which situations different detectors work best. Open Images V4 goes beyond previous datasets also in that it unified with the annotations for image classification, object detection in the same set of images. This allows for cross-task training and analysis, potentially supporting deeper insights about each of the three tasks.

I. Literature Survey

1. Sliding window-based object detection

Bergboer (2007) *et al* [2] have created several studies for learning strategies that are use for context based object detection that are tired paintings for gradient kind and context detection technique. The gradient technique is used for ever-changing spatial context into a gradient.In context detection technique it uses the window to look the image regions within which the objects contained.Gradient technique works with assumptions so it takes to a lot of time constraints despite the fact that a single object is to be detected in a picture.And the alternative introduced context detection works supported window uses the temporal arrangement constraint as a result of it got to search every window for the presence of Associate in Nursing object. Inaccurate localisation may be a terribly huge downside

whereas victimisation the slippy window-based object detection. Segvic (2011) have decide a way to increase the localisation accuracy and was achieved by removing spatial cluster of close detection responses. It helps to introduce 3 main goals they're high recall,high preciseness and correct localisation. spatial cluster will be wont to cut back the range of false positives.Sliding window technique at the start fixes the size of the window in that it can be looking for the object. For increasing the speed of detections Comaschi (2013) have planned a window used for real time applications here a method that the step size is deciding at the run time.He incontestible that however this system improves the performance of Viola Jones object detection. It conjointly helps to come through a quickening of two.03x in frames per second while not decreasing the accuracy.

2. Contour-based object detection

Stiene (2006) *et al* [3] have projected associate OD approach that is predicated on vary pictures as vary pictures ar well suited to contour extraction.They have used a 3D optical device scanner and reliable contour extraction with floor interpretation for the method of OD though this method has high performance whereas functioning on vary pictures changing the natural pictures into vary pictures is the overhead method that desires to be performed when associate object is to be detected. Contour-based OD will be well developed as a matching drawback between model contour elements and image edge fragments and therefore principle (2012) have used this drawback and have treated it as a tangle of finding dominant sets in weighted graphs wherever the nodes of the graph ar pairs composed of contour elements and edge fragments and therefore the weights between nodes ar supported form similarity the most advantage of this method is that it will discover multiple objects gift in a picture in one pass.Still the question arises which will this method discover objects in associate occluded image or alternative varieties of pictures primarily the objects in a picture areoften defined on the premise of their look and by the form of their contours.Schlecht and Ommer (2005) have investigated a native illustration of contours for OD that enhances appearance-based data they need combined contour and look data into a general optionbased detection formula and have claimed that the mixture has considerably improved the performance compared to the opposite voting ways.

3. Graph-based object detectio

Model-based ways play a central role to unravel totally different issues in laptop vision a specific necessary category of such ways depends on graph models wherever AN object is rotten into a variety of elements, every one being diagrammatic by a graph vertex. He (2004) have given a skeleton-based graph matching methodology for beholding and object localisation that makes use of the skeleton model and contour phase model for this purpose. the employment of those models helps scale back the matching area relatively.This methodology has worked with satisfaction in case of medicine pictures however still there remains a scope to implement this methodology on varied different forms of pictures and see if we tend to still get satisfactory results. Felzenszwalb ANd Huttenlocher (2004) have addressed the matter of segmenting a picture into regions; this is often achieved by process a predicate in order to live AN proof for a boundary between 2 regions by creating use of a graph-based illustration of the image and by developing an economical segmentation formula based mostly on the predicate outlined earlier. but finding a segmentation that is neither too coarse nor too fine is AN NP-hard downside, therefore there remains a large scope in redesigning this methodology of image segmentation and to urge smart results.

4. Fuzzy-based object detection

Reyes and Dadios (2004) have developed a logit-logistic fuzzy color constancy (LLFCC) formula for dynamic color object recognition. This approach focuses on manipulating a color locus that depicts the colors of associate degree object. A set of adaptative distinction manipulation operators is introduced associate degreed utilized in conjunction with a fuzzy illation system and a brand new perspective in extracting color descriptors of an object ar conferred. once more the question here arises regarding what color ranges will be detected feasibly. Munoz-Salinas (2004) create use of the info provided by the camera of a automaton so as to assign a belief degree on the existence of a door in it this can be done by analysing the segments of the image. Many fuzzy ideas ar outlined to guide the search method and realize totally different cases during which doors will be seen. Options of the segments like size, direction or the distance between them ar measured and analysed exploitation fuzzy logic in order to establish a

membership degree of the segments on the outlined fuzzy ideas.This work is solely restricted to indoor environments.Bernardin (2007) have conferred associate degree automatic system for the observation of the indoor environments exploitation the pan-tilt-zoomable cameras. The system makes use of Haar like feature classifier and color bar graph filtering in order to attain reliable formatting of person tracks. The system uses a combination of adaptative color and KLT feature trackers for face and higher body that permits for sturdy pursuit and track recovery within the presence of occlusion.

5. Context-based object detection

Wolf and Bileschi (2006) have developed a detector for object context with the help of victimization context, authors have incontestable detection of locations that area unit seemingly to contain the item of interest. chiefly they have shown that context could be determined from basic visual options like color and texture. Occlusion remains some extent. Perko and Leonardis (2010) have conferred a framework for visual-context aware OD authors have tried to extract visual discourse info from pictures which might be used before the method of OD. additionally, bottom-up strikingness and object cooccurrences area unit utilized in order to outline auxiliary visual context. Finally all the individual discourse cues area unit integrated with native appearance-based object detector by victimization a absolutely probabilistic framework. This system is tested on still pictures will it work on different sorts of pictures remains a difficulty.

II. Background

In olden days the first application of image processing was in the field of paper printing. A picture transmission technique called Bartlane cable picture transmission system was introduced in 1920s this is shown in fig 1.In this system pictures are transmitted through the cable by coding the picture using a special printing equipment at the transmission point and this will be reconstructed at the receiving end.Later at the end of 1921 this technique was used in photographic reproduction method. In Baslane system coding of images was done using 5 distinct gray level. This was further developed in 1929 by increasing it in to 15 gray levels.It is obtained by 15 different tone equipments.After the invention of computer there was a major progress in image processing during the year of 1960. In 1960s Jet Propulsion Laboratory, Bell Laboratory and other facilities well developed first digital image processing for satellite imagery, wirephoto standards.After this development the Jet Propulsion Laboratory in 1964 transmitted the pictures of moon by the help of Ranger7.Early in 1970's digital image processing was started to use in medical field and astronomy. Computerized tomography (CT) is a image processing application introduced first in medical field for this X-rays were used. X-rays are passed through the object in a perticular manner from a starting point and are collected at its opposite end.For the invension of tomography in the year 1979 Sir Godfrey N Hounsfield and professor Allan M Cormark received Nobel price for medicine. The aim of early image process was to enhance the standard of the image. It was geared toward folks or individuals or personalities to enhance the visual impact of people. In image processing, we assume that the input is a low-quality image, and the output is an image with much improved quality. Common image process embrace image improvement, restoration, encoding, and compression. The first made application was the Jet Propulsion Laboratory (JPL).

For the development of JPL image process techniques was used. In 1964, taking into account of the position of the sun and the environment of the moon.The cost of process was fairly high with the computing equipment of that era. This semiconductor diode to pictures being processed in period, for some dedicated problems such as television standards conversion.

Fig 1: Early digital image in 1920

The aim of early image process was to enhance the standard of the image. It was geared toward folks or individuals or personalities to enhance the visual impact of people. In image processing, we assume that the input is a low-quality image, and the output is an image with much improved quality. Common image process embrace image improvement, restoration, encoding, and compression. The first made application was the Jet Propulsion Laboratory (JPL). For the development of JPL image process techniques was used. In 1964, taking into account of the position of the sun and the environment of the moon.The cost of process was fairly high with the computing equipment of that era. This semiconductor diode to pictures being processed in period, for some dedicated problems such as television standards conversion. As general purpose computers became quicker and they began to take over the role of dedicated hardware for about the foremost specialised and computer-intensive operations. It is generally usedbecause it is not solely the foremost versatile methodology conjointly the most affordable.Digital image processing uses computer algorithms for performing image processing on digital images.Image processing perform some operations in a image for extracting information from it or to get an enhanced image.

Object detection

Object detection is used with the help of computer vision or image processing.Object detection detects instances of objects in a perticular class such as humans, buildings or vehicles in images and videos.Very good researched domains of object detection includes face detection and pedestrian detection.The progress of object detection has two historical periods.One is "traditional object detection period" which was before 2014 and the other one is "deep learning based detection period" which was after 2014 this is shown in the fig 2.

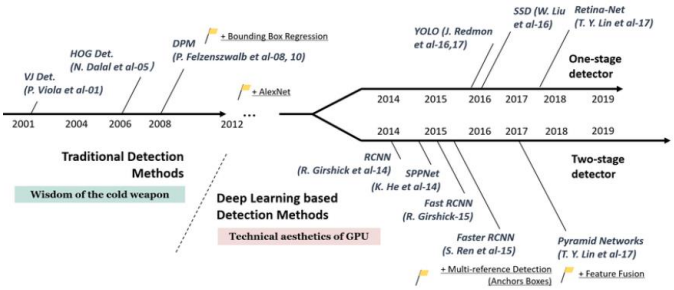


Fig 2: History of object detection

Viola Jones detectors, HOG detector and Deformable Part-based models are different traditional object detection .18 years ago P.Viola and M.Jones achieved real-time detection of human faces for the first time without any

constraints.It runs on a 700MHz Pentium III CPU the detector was tens times faster than other algorithms in its time.This algorithm was named VJ detector in honour of the condributor.The VJ detector follows a detection way called sliding windows.The VJ detector has improved its detection speed by using three techniques such as integral image,feature selection,detection cascades.Histogram of Oriented Gradients (HOG) is a feature descriptor which was proposed in 2005 by N.Dalal and B.Triggs. HOG is considered as an important improvement of the scale-invariant feature transform and shape context in those time. To balance the feature invariance and the nonlinearity HOG descriptor is designed. to be computed on. HOG can detect a variety of object classes, but it cant do pedestrian detection. For the detection of objects in different sizes the HOG detector is used.Here the detector will rescale the input images for multiple times without changing the size of detection window. The HOG detector was used for a long time been an important foundation of many object detectors and it became helpfull for a large variety of computer vision applications for years.In the year 2008 P. Felzenswalb proposed DPM is as an extension of HOG detector a variety of improvements was made by R. Girshick. The DPM follows a detection method of "divide and conquer" in which training is very simple.It does the learning process in a proper way of decomposing an object . The detecting of a car will be considered as the detection of its window, body and wheels. A DPM detector had a root-filter and a number of part-filters.

Deep learning object detection

Deep models are known as neural networks with deep structures.The history of neural networks started in the year 1940s and the original use was to simulate the human brain system to solve general problems which helps for learning.In the year 1980s and 1990s the back-propagation algorithm by Hinton was very popular.Due to the lack of large scale training data, less computation power and good performance the neural networks was widely used in early 2000s.Deep learning become popular from 2006 with a devolepment in speech recognition.CNN is the most popular model of deep learning. A CNN architecture is referred to as VGG16 .Different layers of CNN is known as a feature map The feature map of the input layer is a 3D matrix of pixel intensities for different color channels such as RGB.Different types of transformations can be done on feature map such as filtering and pooling.Convolution operation convolutes a filter matrix with the values of a field of neurons and takes a nonlinear function such as sigmoid , ReLU to obtain final responses.

R-CNN: The idea of RCNN is simple starting is with the extraction of a set of object proposal by selective search.The CNN model trained using ImageNet is fed with each proposal that was rescaled to a fixed size image for extracting features.In the final stage linear SVM classifiers are used for predicting the presence of an object within each region and to recognize object categories.RCNN gives a performance boost on VOC07, with a large improvement of mean Average Precision (mAP) from 33.7 percentage to 58.5 percentage.

SPPNet:In 2014 Spatial Pyramid Pooling Networks (SPPNet) was proposed. Old CNN models needs a fixed size input such as 224x224 image for AlexNet.The main contribution of SPPNet is the introduction of a Spatial Pyramid Pooling layer, which enables a CNN to generate a fixed-length representation regardless of the size of image of interest without rescaling it. When SPPNet is used for object detection, the feature maps can be computed from the entire image only once, and then fixedlength representations of arbitrary regions can be generated.This is done for to avoid computing the convolutional features. SPPNet is more than 20 times faster than R-CNN without sacrificing any detection accuracy.Although SPPNet has effectively improved the detection speed, there are some drawbacks such that the training is in multi-stage and SPPNet only fine-tunes its full layers and simply ignores all previous layers.

Fast RCNN: In 2015, R.Girshick proposed Fast RCNN detector, which is a further improvement of R-CNN and SPPNet. Fast RCNN is used to train a detector and a bounding box regressor. On VOC07 dataset, Fast RCNN increased the mAP from 58.5 percentage (RCNN) to 70.0 percentage while with a detection speed over 200 times faster than R-CNN. Fast-RCNN successfully combine the advantages of R-CNN and SPPNet but also its detection speed is still limited by the proposal detection.

Faster RCNN: In the year 2015 Faster RCNN detector was proposed shortly after the introduction of Fast RCNN. Faster RCNN is the first end-to-end framework and also the first realtime deep learning detector. A very importent contribution of Faster-RCNN is the Region Proposal Network (RPN) that enable cost-free region proposals. Detection, feature extraction, bounding box regression have been integrated into a single end-to-end learning framework.

Feature Pyramid Networks: Feature Pyramid Networks (FPN) was introduced in the year 2017 on basis of Faster RCNN. Detector run detection on top layer on the network before FPN.Features in deeper layers of a CNN are usefull for category recognition it is not conducive to localizing objects.Lateral connections are done in FPN in the end for making a topdown architecture.CNN naturally forms a feature pyramid by forward propagation the FPN shows great advances for detecting objects with a wide variety of scales. Faster R-CNN system uses FPN for achievements.FPN has now become a basic building block for many latest detectors.

III. Implementation

To better understand the visual content, we should not only know what is the object it is called classification task however additionally grasp wherever is that the terribly object it is called location task.The object detection task is to simultaneously provide these two information for a given image the fig 3 shows the block diagram of system.Depending on the pipeline most of themobject detection techniques could be divided into two categories is that one-stage method nd two-stage method. Generally one-stage methods focus on the speed performance while the dominant benefit of the two-stage strategies is that the precision performance. A two-stage methods is used for considering the outstanding precision performance. In the modern convolutional neural network (CNN) context, the regions with CNN features (RCNN) method should be the earliest two-stage detector. Just as its name implies the RCNN methods first output multiple region proposals using the selective search algorithms, then regress the bounding-box (bbox) coordinates and classify into a specified class based on the extracted CNN features of the proposed region with the matured support vector machine algorithm. To accelerate the pipeline the ResNet is proposed by claiming that the feature maps could be shared by different proposals and hence reducing the computation burden of the feature extraction process.

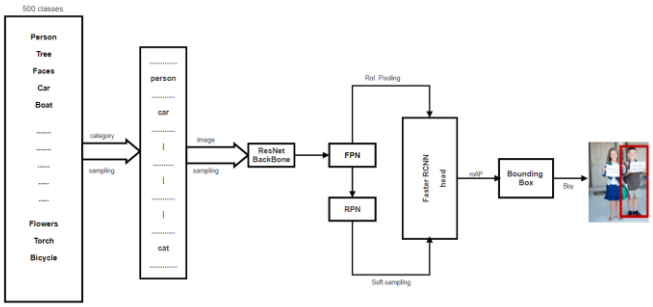


Fig 3: Block diagram

A. Open image data set

Open Images Dataset which contains images and ground-truth annotations for the three tasks. Open Images V4 has several attractive characteristics, compared to previously available datasets. The images were collected from Flickr without a predefined list of class names or tags, leading to natural class statistics and avoiding the initial design bias on what should be in the dataset. They were released by the authors under a Creative Commons Attribution (CC-BY) license that allows to share and adapt the material, even commercially; particularly so for models trained on these data, since it makes them more easily usable in any context.The Open image Dataset differs in 3 ways that from alternative datasets initial one is that each one image have inventive Commons Attribution (CC-BY) license and may so be a lot of simply used, with correct attribution . Second is that the images should be collected ranging

from Flickr so removing images that seem elsewhere on the net. This removes straightforward pictures that seem in search engines like Google Image Search, and thus the dataset contains a high proportion of attention-grabbing, advanced pictures with many objects. Third is that the pictures don't seem to be scraped supported a predefined list of sophistication names or tags, resulting in natural category statistics and avoiding the initial style bias on what ought to be within the dataset.

B. Bounding Boxes

Bounding boxes are annotated for the 600 boxable object categories. Google-internal annotators are used for drawing all boxes on the Open pictures information. it's crucial to coach annotators victimisation an automatic method within the spirit.Training consists of 2 elements one is supposed to show extreme clicking. Here annotators draw boxes on ten objects for every of the twenty PASCAL VOC categories once every category we have a tendency to mechanically give feedback on that boxes were properly or incorrectly drawn and why, by showing valid attainable positions of the intense points half two may be a qualification task during which the annotators follow each speed and accuracy. they're asked to draw 800 boxes and pass if their intersection-over-union (IoU) with the bottom truth is beyond 0:84 and drawing time per box is twenty seconds or less.

C. Evaluation Criteria

To validate object detection results two criterias are used overall precision and overall recall. Precision and recall is to validate object detection results. Overall precision is used in comparison with other state-of-art methods.

$$precision = \frac{tp}{tp + fp} \quad recall = \frac{tp}{tp + fn}$$

The computation of precision and recall is given in equation given below. The actual meaning of precision is how many predicting items are relevant elements, and Recall represents how many relevant elements are predicted.

IV. Result Analysis

The datasets of the Open Images challenge contains 1.7 million images ranging in 500 categories, where 100K images are the official suggested validation dataset.In our custom settings, to accelerate the evaluation process and also enlarge the train datasets, we only use 5000 images as the mini-validation dataset, and the other as the train dataset fig 4 and fig 5 shows the detection of objects in various pictures. The initial learning rate is 0.01 and reduced to 0.001 after 40K iterations. The training process will be terminated after 50K iterations. The batch size is 48. The default multiscale training and testing strategies in the Detectron framework. Six Tesla-V100 GPU are utilized for trainging.The data balance strategies could boost the performance heavily, 8.4 absolute points from the 46.5% baseline to the mAP of 54.9%. By further using the hierarchical strategy, the performance could be further improved by 2 absolute points, achieving the best single best model with a mAP of 56.9%. With a final ensemble strategy with 8 different models, we achieve the 62.2% mAP performance. A visual comparison among different single models such as baseline, baseline+data balance,baseline+data balance+HNMS.For the majority category such as the Person class all the model show good results.However for the minor category such as the Paddle and the Duck classes the baseline model produce inferior results with mis-located paddles and the missed ducks.The data balance strategy could greatly alleviate this problems as demonstrated in the middle of .With further hierarchical NMS, the label from the parent nodes could be correctly output as

shown in the middle region.In this Open Image challenge, we find that there exists server data imbalance and label missing problem. We use the re-sampling strategy to tackle the data imbalance problem, which could improve to 54.9% compared to the baseline model with a mAP of 46.5%. To alleviate the label missing problem, a hierarchical NMS strategy is proposed, increasing the performance of the single model to 56.9% from the 54.9%. Finally the

ensemble strategy is adopted,boosting the mAP to 62.2%. Besides, we find that for the label missing datasets, the popular OHEM strategy should be avoided.

When checking two modern object detection models with different capacities. The first model is Faster-RCNN with an Inception-ResNetV2 backbone, which performs feature extraction. The second model is SSD with MobileNetV2 feature extractor with depth multiplier 1:0 and input image size 300*300 pixels.The table shows report in the number of parameters and inference speed for each detection model.The annotation spectrum is data collection as proposed by the creators of Visual Genome (VG) and Visual Relationship Detection (VRD) datasets. The focus was on obtaining as much variety of relationships as possible by asking annotators. The annotations from several annotators were then merged and combined using various language and quality models.The difference in the two approaches naturally leads to difference in the properties of the two datasets while VG and VRD have a large number of relationship prepositions and object classes. The work shows that many of those are rather obvious the window on building. The top-10 most frequent relationship triplets in all three datasets in both VG and VRD the most frequent relationships can be predicted from object co-occurrence and spatial proximity while Open Images is more challenging in this respect. Second as follows from the free-form annotation process and lack of precise predefinitions, the annotations on VG and VRD contain multiple relationships with the same semantic meaning: for example, the difference between relationships ‘near’ and ‘next to’ is not clear. This leads to annotation noise as multiple instances of conceptually the same relationship have different labels.Since Open Images annotations were collected in a very controlled setting this kind of noise is much lower. Finally in VG and VRD annotations within an image are sometimes incomplete .If there are two chairs at a table in the same image, only one of them might be annotated).In Open Images for each image it is possible to know exactly if two objects are connected by a certain relationship or not.

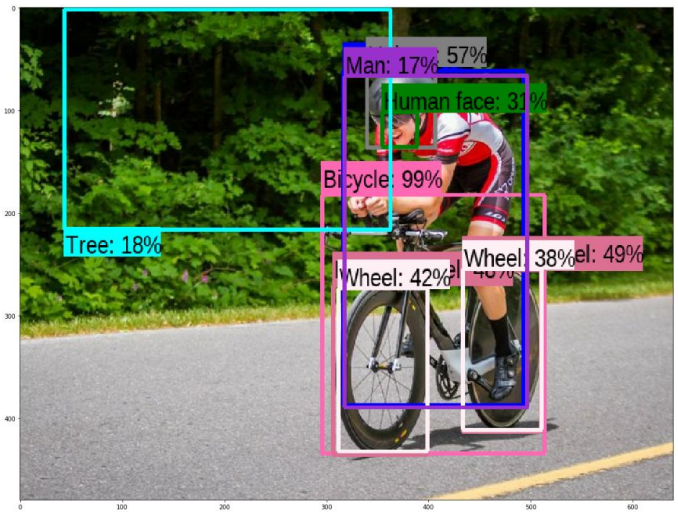


Fig 4: Object detection of man on a bycle

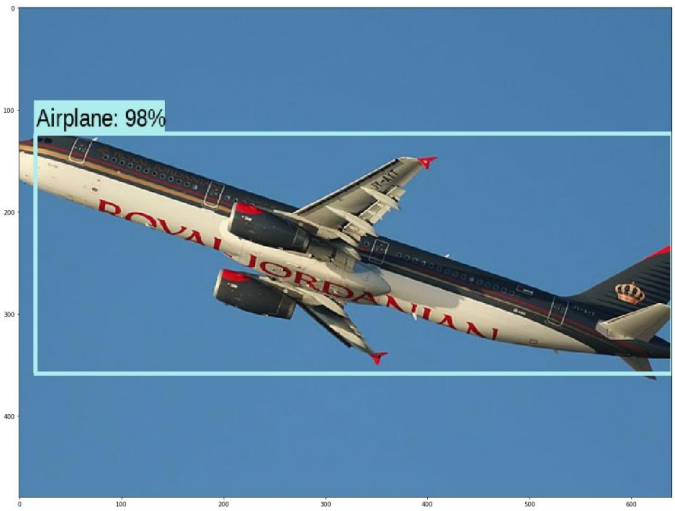


Fig 4:Object detection of aeroplane

V. Conclusion

The concept of a novel object detection system is for the purpose of providing the correct detection and localisation of different images in varied and complex image. The prototype of the object detection is done using a open image dataset which consist of a huge set of classes. The explanation of how the data was collected, annotated and presented comprehensive dataset statistics also evaluated its quality and reported the performance of several modern models for image classification and object detection. The scale, quality, and variety of Open Images V4 helps for research and innovation.

REFERENCES

[1] Zhong-Qiu Zhao, "Object Detection with Deep Learning: A Review", IEEE Transactions On Neural Networks And Learning Systems, 2162-237X c 2019 IEEE.

[2] Julian Muller "Disparity Sliding Window: Object Proposals From Disparity Images",arXiv:1805.06830v2 [cs.CV] 6 Aug 2018

[3] Stefan Stiene,Kai Lingemann, University of Osnabr¨uck Knowledge -Based Systems Research Group"Contour-based Object Detection in Range Images".|

[4] Reagan L. Galvez," Object Detection Using Convolutional Neural Networks",IEEE Region 10 Conference (Jeju, Korea, 28-31 October 2018), 2023-202X.

[5] Alina Kuznetsova Hassan Rom,"Unied image classication, object detection, and visual relationship detection",Google AI(2 Nov 2018).

[6] A. Bearman, O. Russakovsky, "What's the point: Semantic segmentation with point supervision",In ECCV, 2016.

[7] D. Ushizima, A. Carneiro, M. Souza, and F. Medeiros, "Investigating pill recognition methods for a new national library of medicine image dataset," in Proceedings of the International Symposium on Visual Computing (ISVC), Part II, LNCS 9475,2015.

[8] J.-S.Wang, A. M. Ambikapathi, Y. Han, S.-L. Chung, H.-W. Ting, and C.-F.Chen, "Highlighted deep learning based identification of pharmaceutical blister packages," in Proceedings of the 2018 IEEE 23rd International Conference on Emerging Technologies and Factory Automation (ETFA), 2018.

[9] Sandler M, Howard AG, Zhu M, Zhmoginov A, Chen L (2018) Mobilenetv2: Inverted residuals and linear bottleneck. In: CVPR.

[10] Sun C, Shrivastava A, Singh S, Gupta A (2017) Revisiting unreasonable effectiveness of data in deep learning era. In: ICCV.

[11] Deng, J., Berg, A., Satheesh, S., Su, H., Khosla, A., and Fei-Fei, L. Imagenet large scale visual recognition competition 2012 (ilsvrc2012), 2012.

[12] Girshick, R. Fast r-cnn. In Proceedings of the IEEE International Conference on Computer Vision (2015).

[13] Jain, V., and Learned-Miller, E. G. Fddb: A benchmark for face detec- tion in unconstrained settings. UMass Amherst Technical Report (2010).

[14] Szegedy, C., Reed, S., Erhan, D., and Anguelov, D. Scalable, high- quality object detection. arXiv preprint arXiv:1412.1441 (2014).

[15] Wang, X., Yang, M., Zhu, S., and Lin, Y. Regionlets for generic object detection. In Proceedings of the IEEE International Conference on Computer Vision (2013).

Naipunnya®
To reach the unreachable

COURSES OFFERED

- B.Com Finance (2 Batches)
- B.Com Computer Application
- B.Com Co-operation
- BBA
- M.Com
- B.Sc Computer Science
- BCA
- M.Sc. Computer Science
- BA English Language and Literature
- MA English
- B.Sc. Hotel Management and Catering Science (2 Batches)
- B.Sc. Hotel Management and Culinary Arts
- Bachelor of Hotel Administration
- Craftsmanship Course in Catering Management

NAIPUNNYA INSTITUTE OF MANAGEMENT AND INFORMATION TECHNOLOGY (NIMIT)

(Affiliated to the University of Calicut, Accredited by NAAC with B++, ISO 9001-2015 Certified)

Pongam, Koratty East,Thrissur District, Kerala State - 680 308.
Ph : 9605001987, 0480 2730340, 2730341. Website: www.naipunnya.ac.in
Email - mail@naipunnya.ac.in