Journal of Data Acquisition and Processing
ISSN: 1004-9037
https://sjcjycl.cn/
DOI: 10.5281/zenodo.754973

INNOVATIONS IN MULTIMEDIA SERVICES: DEEP-LEARNING ALGORITHM INTEGRATION

Ms Zaiba Khan
Assistant Professor (FOBAS-CSE), RNB Global University, Bikaner, Rajasthan, India
Email: zaiba.khan@rnbglobal.edu.in

Abstract
Multimedia streaming services have been among the top applications on the internet, due to the vast amount of information presented through it. The advent of the semantic web and the services surrounding it includes many methods of accurately describing multimedia content. This paper investigates the use of a deep learning method to improve the quality of multimedia. Deep learning is a class of machine learning methodologies that uses many-layered artificial neural networks to learn internal representations of data and is showing increasing promise in optimizing multimedia services in complex ways. In the proposed intensification methodology, we integrate Deep Learning algorithms into different multimedia services, particularly focusing on content adaptation techniques such as coding, transcoding, and encryption. Then, we elaborate on how the Deep Learning algorithm can be integrated into coding and transcoding services, which are then followed by how it can facilitate intelligent content adaptation in encryption. Finally, we conducted an experimental comparison between traditional machine learning methods and deep learning methodology to automatically induce a physics simulator to demonstrate great future promise in this field. This comprehensive literature review has revealed that image compression remains uncharted territory for content-based methods, while steady progression has been made in other MIR domains.

Keywords: Multimedia, Deep Learning, Machine Learning, Encryption, Image Compression

1. Introduction

Multimedia streaming services have been among the top applications on the internet, due to the vast amount of information presented through it. The Internet and intranet networks have grown rapidly along with the number of users. Multimedia services are becoming mainstream in both internet and local area networking (LAN) environments. Examples of popular multimedia services are YouTube, Netflix, Spotify, Facebook, and Instagram. These services have revolutionized the way people consume media. Users can stream and acquire information with ease, depending on the connection speed of the network they are connected to. Unfortunately, several types of multimedia consist of high definition and resolutions, which would require a good amount of bandwidth for a user to watch and completely understand the content. High bandwidth usage would eventually lead to congestion while there are a lot of subscribers accessing the multimedia services. Moreover, utilizing mobile networks for accessing multimedia services can incur significant expenses if these services utilize a substantial amount of bandwidth. (**Haile et al., 2021**) Consequently, subscribers may be unable to receive the necessary content due to the limitations of available bandwidth. Similarly, subscribers relying on a local area network would encounter comparable predicaments due to the insufficiency of bandwidth in a shared setting. As a result, subscribers wouldn't be able to receive the required content due to bandwidth limitations. Subscribers using a local area network would also experience the same due to the lack of bandwidth in a shared environment.

1.1. Background

In order to deliver precise, satisfactory user experience with optimized resource efficiency, it is imperative that the quality adaptation decisions are correct, i.e., choosing the best action to take in changing the content. Deep-learning algorithms are a class of machine learning algorithms based on learning multiple levels of representation and abstraction that help make decisions. It has been proven that these algorithms outperform other intelligent decision-making algorithms for a wide range of tasks. A recent advancement in deep learning, the deep reinforcement learning algorithm, has shown exceptional performance in gaming and robotics. Its potential for quality adaptation makes it the central focus of this paper. (**Mohamed, 2023**)

Using middleware is one way to guarantee constant content quality even in the face of evolving network and device capabilities. By generating many bitrate versions and providing the most appropriate one, this middleware modifies the delivery of material. On the other hand, smart algorithms such as machine learning improve and automate decision-making for high-quality adaptation. This paper investigates the use of a deep learning method to improve the quality of multimedia.

The user experience, which is greatly influenced by the quality of the material, is the focus of multimedia service providers' ongoing efforts to enhance their offerings. This includes crisp audio and visuals as well as video. Generally, multimedia content is saved and conveyed in its rawest, optimal form—also referred to as the bitstream—to optimize visual and display quality. Large storage space and transmission bandwidth are needed for this, though, and they might not be affordable or readily available. Some common scenarios are a mobile device that is using a wireless network and is running on battery life, or a common user with low RAM and a poor internet connection. To address these quality-related problems, action must be taken.

Today's digital communication relies heavily on multimedia. A greater quantity of multimedia, including music, video, and picture files, are being sent via networks for information sharing because of the internet's rapid expansion and wireless technologies. This is demonstrated by the enormous number of websites—and the constant emergence of new ones—that provide a variety of multimedia content for didactic, promotional, and entertaining reasons. The widespread use of mobile devices, such smartphones, has also made it easier to access web-based multimedia services, which are now a need in daily life. There's no stopping this tendency from getting bigger.

1.2. Purpose

The human thinking process is complex and encompasses various elements. If we aim to replicate the entirety of human thinking, we encounter numerous aspects to consider. However, there is one straightforward yet impactful method that influences perception: mood perception. By implementing mood perception, the processing of acquired data in the human cognitive system is adjusted, resulting in a distinct perception depending on our current mood. This technique involves unconsciously eliminating unnecessary information, storing the modified data in long-term memory, and subsequently influencing actions in similar circumstances. To illustrate this, consider the proficient editing of a particular video sequence as an intelligent response to audio-visual stimuli.

Another aspect to be considered is better functionalities supported. This could mean some kind of smart actions taken to the audio-visual signal. An example of this smart action is a smart edit on a video sequence. Smart action on the audio-visual signal should be done with an intelligent system that should understand the audio-visual signal itself. This fact brings a new challenge for improving multimedia services, how we can integrate the intelligence into the application of the multimedia services. The most common way to apply intelligence to the system is by using a knowledge-based system or expert system. But it will be more interesting if we can implement an artificial intelligence that stands closer to the way a human thinking method is usually done.

Because of those improvements of the networks, multimedia services users demand higher quality from the existing multimedia coding technologies. Quality here means better audio-visual quality and better functionalities supported. Better audio-visual quality can be achieved by improving the existing audiovisual coding technologies. It should be followed by the improvement of how to get better perception from the audio-visual signal itself.

The multimedia services industry has been developing rapidly over the last few years. The latest improvements on the 3G network and the upcoming 4G network will only hasten this development. Multimedia services covering video on demand (VoD), video conferencing, and Interactive TV (iTV) will become more common to be used soon. The improvements of the networks will be able to provide video streaming bidirectionally without any significant delay. Although currently the most common use for multimedia services is only for entertainment, in the future, multimedia services will also cover areas such as education, health, and public services.

2. Deep-Learning Algorithm

Allowing computers to learn to recognize and understand data patterns from a stunningly large pool of examples, the deep-learning artificial neural network technology functions in a revolutionary way. In mimicking the human brain's information processing patterns, the network deploys task-grouping layers of computing nodes to perform supervised (or unsupervised) learning through presentation of specific stimulus regarding the task. This elaborate automated feature extraction and analysis process eliminates the need for human intervention in deriving specific functionalities for detected patterns. High abstraction of data through the multiple processing layers and the hierarchical learning paradigm are key offers in delivering the best possible analysis and automation potential compared to other pattern recognition techniques. The rapidly expanding echelons of deep learning technology recently served notable successes in pushing data classification accuracy rates to the envisioned 90% record in the ImageNet Large Scale Visual Recognition Challenge. Targeting to simulate human-level understanding from image data, it complements the expansion of big multimedia data and effectively triples data control throughput compared to the conventional tag-and-search data annotation methods. Akin to multimedia information retrieval methodologies, deeper categorization and understanding of detected features in data serve to bridge the gap between low-level signal processing data representations and high-level semantics making it more accessible to derive meaningful conclusions. (**Aggarwal et al.2022**) (**Atila et al., 2021**)

2.1. Definition

In supervised learning systems, feature engineering is a crucial step, requiring time and expertise. Deep learning addresses this by allowing systems to automatically learn features from raw input data and corresponding outputs. This can enhance system performance across various tasks by providing compressed representations of input data, as seen in image classification tasks like detecting cats. Deep learning, a rapidly evolving field, aims to empower computers to learn and make intelligent decisions with minimal human intervention. Its applications span computer vision, speech recognition, natural language processing, audio recognition, and bioinformatics.

2.2. Applications

The first stage in a machine learning system is to extract features from data that can be used to perform a task. In multimedia, this involves the automatic tagging of images or video ("feature extraction"). It largely is the process of training statistical models which can then automatically make decisions or predictions ("feature classification"). An example of automatic tag extraction on an image may be the categorization of a photo of someone's pet dog as "dog," "pet," "animal." Feature extraction in multimedia is vital to the success of automatic categorization and organization of multimedia as once meaningful features from the data have been identified, it becomes much easier to search, filter, and manage a very large collection of multimedia. Intelligible feature classification means that a person can express what exactly they wish to find in their data in a way which a machine can then interpret and provide relevant search results. An example of this would be a direct search for a specific genre of music video which could then automatically be filtered and arranged into a collection of like genres. All machine learning models require input of training data in which a set of input features, often represented as a vector $\mathbf{x}$, have a known corresponding output $\mathbf{y}$. The data can then be used to build a model which can map input features to output. In deep learning models, a hidden layer of nodes will be used to learn a hierarchical representation of the data thus trying to identify useful features. This effective unsupervised learning is what differentiates deep learning from typical machine learning. High level features can be learnt and then employed to efficiently categorize or make decisions about input data.

2.3. Advantages

Distinguishing the advantages of integrating deep-learning algorithm into the chosen company system against other complex algorithms, the following qualities were deemed favorable. First is on the aspect of complexity, compared to other complex algorithms, deep-learning algorithm offers a high accuracy rate with less fine tuning involved. This would result in lower maintenance costs in the long run. The company can then focus on other aspects of the system development. Second, in terms of feature extraction and transformation, deep-learning algorithms can automatically extract or construct useful features from multimedia data. Being the algorithm can learn directly from the provided data and optimize the performance, this can save a lot of time as compared to other algorithms that may require the employment of a data scientist or machine learning expert to identify what feature best suits the desired prediction or classification. Subsequently, no existing algorithms can adapt and learn from new data as deep-learning algorithms. The current trend of Big Data, Internet of Things (IOT) and Multimedia Analysis that continually generates new data can greatly benefit from an algorithm that is able to adapt and learn from the new data to improve its prediction or classification. This will therefore prolong the relevancy of deploying the algorithm into the system. Last but not the least, with today's cutting-edge technology, the company can implement a cost-effective high-performance computation using graphic processing unit for the deep-learning algorithm. This is highly feasible as compared to a decade ago when high-performance computation using graphical processing unit cost a fortune. Deep-learning algorithms with GPU computation would greatly enhance the speed of the algorithm learning and execution process and the low deployment cost is very inviting to companies or individuals who have limited budget.

3. Multimedia Services

Multimedia has revolutionized the way services are offered over the internet. Several popular services, including VoIP, real-time audio and video, and IPTV, depend heavily on the quality of multimedia to be successful. Multimedia algorithms are used to provide services with the best possible quality with minimum resource utilization. These algorithms are based on statistical multiplexing frameworks, where audio and video sources are independently encoded, and the resulting bitstreams are combined at the network layer and then statistically multiplexed to minimize overall distortion, subject to bit rate constraints. Frequently, these services come with quality-of-service guarantees. For example, an IPTV service provider may guarantee 99% uptime with video quality no less than VHS (which itself is much higher quality than a video call). Missing the QoS target can result in large penalties, thus it is important to ensure the algorithms are running effectively and the QoS is being quantifiably measured. With the increasing trend of machine learning, it is quite promising to replace some shallow heuristic algorithms with a more intelligent deep learning algorithm. Deep learning is a wider concept that encompasses the design and development of machine learning techniques based on learning representations of data.

3.1. Overview

Content delivery via multimedia services has become one of the most popular forms of sharing information. The dramatic increase in its use leads us to believe that multimedia services will be an indispensable tool in delivering content in the net-centric computing era. Internet content is migrating from text and graphics to audio and video. The main elements of modern multimedia services are an integration of various forms of data. The integration of data is transparent to the user and is one of the main advantages of using multimedia, as different users can get different information from the same data depending on their preferences. For example, a weather report can be delivered in the form of a map for a visual person, statistical data for a person who is more numerically inclined, and an audio presentation for others. This provides tremendous flexibility in the way data can be presented. The data is often stored in databases and might not necessarily be stored in the same location, which makes its retrieval quite challenging. With faster and more reliable networks, one approach to improving retrieval of multimedia data is to bring the data to the user. However, information seeking might be better accomplished if tools are developed to automatically seek the data. Automatic seeking of desired information is applicable in almost all aspects of multimedia data and is a very active area of research in AI and multimedia.

3.2. Challenges

Creating and delivering multimedia services is a complex task and active area of research. As a matter of fact, this statement is also true for video and image analysis, media streaming services, and many other similar applications. The reason for complexity arises due to the unrelenting data size as the service often involves a huge amount of data. It can be in the form of a database or in the form of continuous streams. Multimedia data has mainly been categorized into two different types of data, i.e. digital data and analog data. The major challenge we deal with when processing analog data is to convert it into digital form. The converted data can never be 100% like the original analog data, and lower quality data can affect the service. High-quality digital multimedia can easily take up to several megabytes just for a small duration of a single data, a still video/image frame, for example. This will also increase the processing time with a large dataset and demand for more storage space to store the data. The characteristics of multimedia data is another issue. Multimedia data is more complex than other data, and many services nowadays are looking to extract useful information from the media. This can involve anything varying from the identification of content within the media to see how the data is laid out within a file. Modeling and algorithms also need to be matched to the media data, and there is a requirement to develop specific algorithms to handle specific multimedia data types.

4. Integration of Deep-Learning Algorithm in Multimedia Services

Deep learning is a powerful algorithm that learns the representation of data using architecting a complex, multi-layered, artificial neural network. These algorithms can automatically learn to represent data with multiple levels of abstractions. This is probably what makes deep learning have good results when working with complex data problems. In addition, these algorithms work well with data classification. By having automatic learning ability, deep learning algorithms will give more accurate results for classifying complex data. Deep learning can also do supervised learning on labeled data and unsupervised learning on unlabeled data. This will be efficient for multimedia services that have an accumulation of data and micro details.

In this section, it will explain the benefits of integrating deep learning algorithms in multimedia services and how to apply the algorithm in real life. Deep learning algorithms have been explored on how they can improve multimedia services for a long time. Deep learning is one of the areas in machine learning and has become very popular in recent years. Deep learning has already had a big impact on several areas such as speech recognition, natural language processing, and others, including multimedia. With the advances in high learning cost and parallel processing in computers, it will have a positive impact on developing and improving multimedia services.

4.1. Benefits

Deep learning has the potential to impact more than just end-user facing content and interfaces. By utilizing advanced machine learning techniques to leverage understanding of user feedback, content can be automatically tagged with semantic data that will enable sophisticated data mining and automation of all sorts of content management and retrieval tasks. An example is the automatic generation of image tags that will enable easier sorting and searching of a large image collection by identifying and tagging image content with information such as locations, objects, or even moods and emotional states. After the recent popularity of deep learning techniques in the machine translation community, where Google uses neural networks to highly generate context-dependent translation predictions, it is now becoming increasingly feasible to automate the translation of multimedia content to and from different languages.

Large-scale leveragings of AI to improve multimedia services are likely to rely on developments in deep learning, a class of machine learning methodologies that uses many-layered artificial neural networks. Through using a multi-layered neural network architecture to automate feature extraction and transformation, deep learning methods have the capability to learn internal representations of data and are showing increasing promise in optimizing multimedia services in complex ways. For instance, we can use deep learning methods to learn representations of multimedia content that automatically optimize according to end-user feedback, so that image and video retrieval systems improve over time; and possibly to automate the design of multimedia interfaces that are responsive to user preferences. These capabilities hold attractions for the likes of commercial multimedia content providers, who wish to maximize the stickability of their content with users and the usability of their platforms, to increase revenue-generating traffic.

4.2. Implementation

In general, there are two kinds of multimedia data collected in the database. The first kind is the metadata, which explains the multimedia content, for instance, user preferences, file format information, and quality of service constraints. This metadata is used to make a decision on how to adapt the multimedia content later. The second kind is the actual multimedia content. Say, for instance, a user wants to send a video stream to another user, but this user only has a GPRS connection. In this case, the video stream needs to be adapted to a lower bit rate video or still image so that the user can understand the content. The decision of what kind of content to send is made based on the metadata and is implemented using an Adaption Controller to control the Automated Content Adaption and act on the actual content. Data at both metadata and actual content can be and are represented by feature vectors. These feature vectors are used in the decision-making process, and the mapping between the old and new data is done using different types of intelligent algorithms.

In the proposed intensification methodology, we integrate Deep Learning algorithms into different multimedia services, particularly focusing on content adaptation techniques such as coding, transcoding, and encryption. Firstly, we provide a high-level view of the framework for Deep Learning in multimedia services. Then, we elaborate on how the Deep Learning algorithms can be integrated into coding and transcoding services, which are then followed by how it can facilitate intelligent content adaptation in encryption.

5. Case Studies

5.1. Example 1

We conducted extensive research and analysis to identify strong correlations between specific data collected during a user's gaming session and the corresponding recorded gameplay footage. Our primary objective was to develop a system that could accurately identify and mark instances within the footage based on the underlying data, thereby significantly enhancing the reviewing process for the user. To achieve this, we employed advanced techniques to extract crucial statistical information from the data. Additionally, we implemented complex algorithms to process the video data and determine the relative point in time for each corresponding event. For the classification aspect of the project, we adopted a supervised learning approach, leveraging the power of a Support Vector Machine (SVM) classifier. This enabled us to effectively classify and categorize the various events occurring in the gameplay footage based on the associated data. Furthermore, we incorporated Dynamic Time Warping (DTW) to compare sequences and measure their similarity. By employing this technique, we were able to accurately identify patterns in the data and align them with the corresponding moments within the gameplay footage. This provided us with a highly precise framework for marking the instances based on the extracted data. While our initial investigation included the consideration of Hidden Markov Models (HMMs) as a potential solution, we ultimately postponed their implementation due to time constraints. However, we acknowledge that HMMs hold significant promise for further enhancing the accuracy and effectiveness of our system. Overall, through our comprehensive approach and utilization of advanced methodologies, we have made substantial progress in bridging the gap between user data and gameplay footage. By leveraging the outlined techniques, we can now streamline the review process for users, ultimately providing a more efficient and enhanced gaming experience. (**Guarnera et al.2023**)

5.2. Example 2

A project that the authors originally embarked on with the goal of beautifying faces in images. While this may sound shallow, applications span from entertainment to security. Input is an image of a person's face, and the output is an image of the same face, however, more attractive. Great potential to have a wide impact as everyone wants to look better in photos. This project involves modifying the illumination of an image with no reference image. Though the techniques are not totally transferable to the deep learning algorithm, it is still an interesting concept and could well serve as a foundation for a larger project. The developers were attempting to create a system which would take an input image of a face and output the same image; however, the face would be more attractive. This targeted a common issue of people being dissatisfied with how they appear in photos. This could have various applications and great potential for widespread effect. An interesting extension of this would be to modify a person's expression in a photo. This imagined system would take an input image of a face with a specific expression and an output of the same face but a different expression, e.g., a frown turned into a smile. This could have great applications in animating still images. (**Bianchi et al.2023**)

6. Future Developments

Developing deep learning has significant potential to express its implementation across other emerging domains. The likely success of this methodology will open new opportunity fields within multimedia services, such as personal agents who will guide multimedia selection, summarization and presentation in order to fulfill information needs and capably assist learning. It may also be applied to scientific information to facilitate the hypothesis and theory formulation through recognition of relevant patterns in the data. Pattern recognition methods could be in the form of automated queries of large multimedia databases providing semantically categorized results. (**Rani et al., 2022**) This would prove beneficial in the medical and health sectors through provision of fast and accurate diagnosis and prognosis of patient conditions. This type of automation may also be used in the development of surveillance systems to predict and prevent safety and security issues. The automation of knowledge may be the greatest future evolution in intelligent multimedia knowledge-based systems. This can be achieved through the automatic induction of logic programs and decision trees from observed or taught examples. An experimental comparison between traditional machine learning methods and deep learning methodology to automatically induce a physics simulator demonstrated great future promise in this field.

6.1. Emerging Technologies

The advent of the semantic web and the services surrounding it includes many methods of accurately describing multimedia content. Of high importance to multimedia mining is the development of languages and annotation systems to specify the content of multimedia data and provide links to semantic knowledge. Such methods can facilitate the retrieval of multimedia data by text-based search engines and the organization of multimedia data into structured form. It is expected that the ability to provide semantics for multimedia content will be an enabling technology for a wide range of multimedia services. An example of the power of semantics can be seen in the field of machine translation. A common problem with machine translation systems is the mistranslation of a word with multiple meanings. By employing methods of semantic representation, it is feasible for machine translation to translate words based on their intended meanings and the context of the sentence (**Wang et al., 2022**). This can be accomplished by creating a link between the word and a dictionary or thesaurus entry in a semantic network and using inference engines to resolve word meanings based on global information. For the speaker whose native language is different from the language of a conference video, it is conceivable that in the future, automatic translation methods can be applied to speech data. By speech recognition, translation, and speech synthesis of the translated text, it is possible to produce translated speech data which can be inserted back into the video. This would allow people to distribute and view multimedia content which is in a language foreign to their own and enable them to better understand other cultures around the world.

6.2. Potential Impact

The deep-learning system itself will save a large amount of time and labor. Currently, the system has been trained on our databases for over a few weeks. After the system becomes better understood, it is likely that we will be able to auto-tag images using unsupervised learning. This would mean running the system over all our images and videos, adding tag information to the databases, and embedding information in the media files. From that point on, anything added to the databases could be automatically catalogued. When the system has been perfected, it would lead to the elimination of both the need for online editors to upload and tag media manually, and searching for media they would like to use. Given that many of our clients upload many photos daily, it would save editors from having to deal with a time-consuming task. An example of this would be the work we currently do for several events photographers and photography agencies who sometimes upload thousands of images. These images then need to be tagged, often with specific search terms to help the editors of the end users' website to find exactly what they are after. This can often take several hours. Auto-tagging the images upon upload would mean that the images are already tagged when the editor goes to look for them, saving a large amount of time.

7. Conclusion

This comprehensive literature review has revealed that image compression remains uncharted territory for content-based methods, while steady progression has been made in other MIR domains. The vast proliferation of unlabeled multimedia data on the internet provides ample opportunity for content-based methods to organize and access this data, so research in these domains will likely remain prevalent in the future. Whether machine learning methods will graduate from approximating image transformation functions to truly understanding semantics is uncertain, but it is safe to say that the field of multimedia information retrieval has a bright future with deep learning algorithms.

Deep learning algorithms are still in the experimental state for many MIR tasks but have shown some promise in learning human-like concepts of music through audio analysis and video or image features. When machine learning is combined with low-level feature extraction, multimedia services with deep learning algorithms will provide a comprehensive means to access and manipulate multimedia data for music, video, and image. The selection of machine learning methods is crucial, as knowledge-based, supervised, unsupervised, and semi-supervised methods all have their place in MIR research. As with any pattern recognition task, supervised learning is likely the best choice when some external information about the desired transformation is available. Short of employing a machine learning scheme or denotation set that has already been developed, this method attempts to learn an approximation of **f(x) = y** from example input-output pairs (x, y). This approach attempts to construct a function that maps input images to the transformation of interest so that new input images can have their transformations approximated. The FUNC (Fast Unsupervised Clustering) algorithm is an example of a method that employed semi-supervised learning to cluster image regions with similar appearance, though the lack of labeled data limited its relative success. In pursuit of a boundary or specific decision regarding some aspect of images, supervised learning with specific constraints is possible. (**Tiddi and Schlobach, 2022**) (**Whalen et al.2022**) Through multimedia services with deep learning algorithms, content-based image retrieval is growing largely. Valuable research employs a compact binary code to decide lossy transformation and quantization for image recovery. At least 1,000 layers are necessary to succeed in learning a transfer function to approximately retrieve an image transformation, though this is impractical with current methods.

References:

1. Haile, H., Grinnemo, K. J., Ferlin, S., Hurtig, P., and Brunstrom, A. "End-to-end congestion control approaches for high throughput and low delay in 4G/5G cellular networks." Computer Networks (2021).

2. Mohamed, M. "Empowering deep learning based organizational decision making: A Survey." Sustainable Machine Intelligence Journal (2023).

3. Guarnera, Luca, Oliver Giudice, Salvatore Livatino, Antonino Barbaro Paratore, Angelo Salici, and Sebastiano Battiato. "Assessing forensic ballistics three-dimensionally through graphical reconstruction and immersive VR observation." Multimedia Tools and Applications 82, no. 13 (2023): 20655-20681.

4. Bianchi, Federico, Pratyusha Kalluri, Esin Durmus, Faisal Ladhak, Myra Cheng, Debora Nozza, Tatsunori Hashimoto, Dan Jurafsky, James Zou, and Aylin Caliskan. "Easily accessible text-to-image generation amplifies demographic stereotypes at large scale." In Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency, pp. 1493-1504. 2023.

5. Rani, S., Lakhwani, K., and Kumar, S. "Three-dimensional objects recognition & pattern recognition technique; related challenges: A review." Multimedia Tools and Applications (2022).

6. Wang, H., Wu, H., He, Z., Huang, L., and Church, K. W. "Progress in machine translation." Engineering (2022).
7. Tiddi, I. and Schlobach, S. "Knowledge graphs as tools for explainable machine learning: A survey." Artificial Intelligence (2022).
8. Whalen, Sean, Jacob Schreiber, William S. Noble, and Katherine S. Pollard. "Navigating the pitfalls of applying machine learning in genomics." Nature Reviews Genetics 23, no. 3 (2022): 169-181.
9. Khan, Z. "Data Mining Data Warehouse: An In-Depth Review". Advanced Engineering Science (2023): 720-732
10. Aggarwal, Priya, Narendra Kumar Mishra, Binish Fatimah, Pushpendra Singh, Anubha Gupta, and Shiv Dutt Joshi. "COVID-19 image classification using deep learning: Advances, challenges and opportunities." Computers in Biology and Medicine 144 (2022).
11. Atila, Ü, Uçar, M., Akyol, K., and Uçar, E. "Plant leaf disease classification using Efficient Net deep learning model." Ecological Informatics (2021).
12. Khan, Z. "AN INSIGHT ON MACHINE LEARNING ALGORITHMS AND ITS APPLICATIONS". European Chemical Bulletin (2023): 6029-6034
13. Gholap. Puja S., Sharma.G., Deepak.A., Sharma.R., Sharma.M., Shrivastava.A. "IoT Enabled Stress Detection Based on Image Processing with Ensembling Machine Learning Approach" ijisae (2024): 760-768
14. Ahmad Y. A. Bani Ahmad, D.Karuna Kumari, Aasheesh Shukla, A. Deepak, Mr. Maulik Chandnani, Dr. Sumit Pundir, Dr.AnuragShrivastava. "Framework for Cloud Based Document Management System with Institutional Schema of Database" ijisae (2023) :672-678
15. Mr. Saurabh Gupta, Dr. Amrapali S. Chavan, A. Deepak, 4Dr. Anil Kumar C., Dr. Sumit Pundir, Dr. Ram Bajaj, Dr. Anurag Shrivastava "Speech Emotion Recognition of Animal Vocals Using Deep Learning" ijisae (2024):129-136
16. Dr. D. Haripriya, Mahmoud Abou Ghaly, A. Deepak, Kamal Sharma, Shanker Chandre, Mr. K.K. Bajaj, Dr. Anurag Shrivastava "A Comparative Study on Online Machine Learning Techniques for Network Traffic Streams Analysis" ijisae (2024):9-19