Available online at www.ijpe-online.com

International Journal of Performability Engineering

vol. 19, no. 5, May 2023, pp. 342-349
DOI: 10.23940/ijpe.23.05.p6.342349

State of the Art Convolutional Neural Networks

Shreshtha Singh* and Arun Sharma

Indira Gandhi Delhi Technical University for Women, Delhi, India

Abstract

Convolutional Neural Networks (CNNs) have become a powerful tool for a wide range of computer vision tasks, such as image classification, object detection, and semantic segmentation. This paper provides an overview of the fundamental concepts and architectures of CNNs, highlighting recent advancements and applications. We discuss the key components of CNNs including convolutional layers, pooling layers, and activation functions. The journey of CNN from its genesis to its evolution to the one we are familiar with today is covered. Although CNN is highly capable on its own many researchers have benefitted by hybridizing CNN with quality models. Therefore, we explore various tailor-made hybrid applications of CNN that are designed to solve very specific problems. We also discuss various innovative fields of applications of CNN and discuss the wide range of fields wherein CNN is performing surprisingly well. The paper concludes with future challenges and endeavors with respect to CNN.

Keywords: convolution neural network (CNN); Deep learning applications; Deep neural network; GPU

1. Introduction

The significance of machine learning is conspicuous in today's world, from abetting real-life applications such as image recognition, speech recognition, traffic prediction, product recommendations, self-driving cars, email spam and malware filtering, to recreational areas such as gaming where the concept of procedural content generation comes into play. Either way, one cannot negate the importance and active involvement of machine learning in the current era. In deep learning, the training data and training period is increased for higher accuracy and better results. DL is a technique through which multilayer neural networks learn from a vast amount of data. Due to a generalized and scalable variance in DL, it is the most popular tool for multiple applications in the current era.

One of the most prominent algorithms of deep learning is known as Convolutional Neural Networks (CNN). Convolutional neural networks perform extravagantly on images but they are used in other areas such as speech recognition and language processing. But when it concerns Natural Language Processing (NLP), Speech Recognition and Image captioning RNNs are more dominant in these specific areas. The credibility for the multiplicity of usage of CNN lies in the fact that CNN is automatic, i.e., they learn with trivial human intervention. The main benefit of CNN over traditional applications is its weight sharing feature, which results in slighter overfitting problems and reliance on the extracted features.

This paper discusses the protuberant milestones in the evolution of CNN from its genesis, its architecture and applications to the revelation of various uncovered directions. The layout is as follows: The architecture of the CNN section deals with the building blocks of modules and layers of CNN, followed by applications of CNN wherein important and prominent applications of CNN are listed. Later, we compiled the variations within CNN that led to tailor-made solutions for complex problems, and the paper reaches its periphery with future challenges in the evolution of CNN.

2. CNN Architecture

The architecture of CNN comprises several significant layers or modules that serve at different levels and result in astonishing results. The layout of CNN Architecture is given in Figure 1.

* Corresponding author.
E-mail address: shreshtha011singh@gmail.com

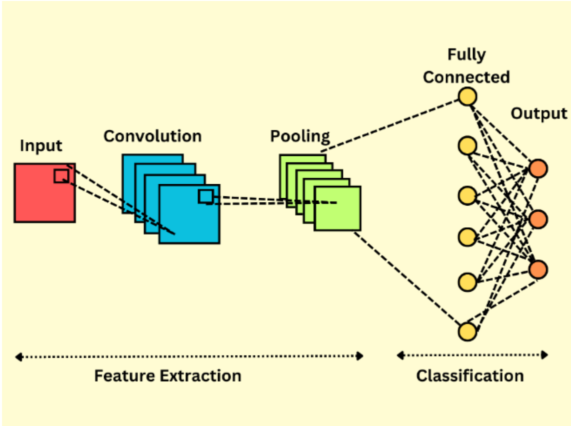

Figure 1. CNN Architecture

The crucial layer is called the convolutional layer, which comprises a set of gauzes also known as kernels. A kernel is a grid of kernel weights that are assigned entirely random numbers. Moreover, these weights are enhanced and modified at each training phase. The input of CNN is a multi-channelled image. In the convolution operation, the kernel scans the image vertically and horizontally, and later a scalar value is calculated. The role of stride becomes prominent when it comes to this operation because if the value of stride increases, it results in a feature map with lower dimensions.

The chief chore of the pooling layer is the down sampling of characteristic maps with the help of various methods like max pooling, average pooling, and global average pooling. All of these methods aim to subsample the feature map. The Activation Function takes charge of mapping the input to the output The activation functions like Tanh, Sigmoid, ReLU, Leaky ReLU, and Noisy ReLU are there to decide whether or not to fire a neuron.

This follows the approach that all neurons from current and previous layers are connected. The loss function pioneers the difference between the original and expected results. Popular loss functions are Cross-entropy or Softmax Loss function, Euclidean Loss Function, and Hinge Loss Function.
The concluding categorization is carried with the help of the output layer. Here, the loss functions help us determine the difference between forecasted and original results.

3. Evolution in CNN Architectures

The journey of CNN architectures started with LeNet [1] in 1989. Since then, various other architectures have evolved the working of CNN [2-7]. The AlexNet model [8] designed by Alex Krizhevsky overcame the major issue that older models dealt with and worked with less hardware quality. GPUs named NVIDIA (GTX-580) were used together for training to overcome lower hardware resolutions with an increase in feature extraction stages [9].

The Network in Network model used multiple layers of perception convolution that ran with the help of a 1 x 1 filter which ensured extra non-linearity. With ZefNet came the concept of getting above training models with Trial and Error and focusing on a multilayer deconvolutional neural network designed by Zeiler and Fergus in 2013 [10].

The embodiment was to visualize and quantify the interior of a network. This visualization helped in pinpointing the weaknesses and then employing appropriate changes to a model. Simonyan and Zisserman came up with the concept of Visual Geometry Group [11]. The abstract of the idea was to introduce the use of small-size filters, which led to a decrease in computational complexity, and the quality of the results did not decrease.

The GoogleNet architecture also called "Inception - V1" revolved around the concepts of consolidate, alter and cleave functions for feature extraction. The agenda of this architecture was to enhance the learning capacity while keeping the computational cost at its minimum [12]. The genesis of the Highway Network came into existence with a cross connectivity concept. The structure has two gating units inside a layer inspired by RNN based on LSTM. The highway network showcases an improved rate of convergence [13].

The ResNet or Residual Network was developed to create an architecture to reduce the vanishing gradient issue. Several ResNet architectures range from 34 layers to a massive 1202 layers [14]. The most customary of them all was ResNet50, which had an impressive number of 49 convolutional layers and an FC layer. With ResNet came the concept of introducing cross-layer connectivity by adding shortcut connections. Continuing on the trail of reducing vanishing gradient problems as acknowledged by Highway Network and ResNet, DenseNet came into existence [13,15,16]. The vital issue with ResNet was information conservation and a large number of weights. To solve this, DenseNet employed cross-layer connectivity.

ResNext, also called Aggregated Residual Transform Network, is an enhanced version of Inception Network [17]. The characteristic features of ResNext come from ResNet, VGG, and Inception. 'WideResNet' was developed by Zagoruyko and Komodakis to conquer the problem of feature reuse [18]. WideResNet gave substantial proof that widening a layer instead of deepening it can result in better performance. The number of parameters in WideResNet is double as compared to ResNet.

The Pyramidal Net was introduced to overcome the learning interference problem faced in ResNet [19]. This model slowly scales up the unconsumed unit to cup the most attainable areas, and hence the name pyramidal is agreed upon for the architecture due to gradual extension in the feature map. The Xception model reduced computational complexity by widening and exchanging a single dimension [20].

The High-Resolution Network has many implementations in the domain of object identification, semantic decomposition, and human posture detection [21,22]. The input images are made to go through a series of transformations from high-to-low resolutions using VGGNet and ResNet. Alternatively, a novel network is used to maintain the high-resolution representation. The HRNet is a boon for problems related to computer vision.

4. Applications of Convolutional Neural Networks

The applications of CNN vary among various fields from Image Processing, Object Classification, Image Segmentation, etc. In this section, the applications of CNN are discussed as well as various instances where CNN architectures were tailor made to serve a specific purpose.

In the domain of medicine, CNN is proving its usefulness. Researchers are applying CNN on various X-ray datasets [23] for classification. Also, CNN is used to detect 14 types of chest related diseases using CheXNet [16]. Since Manual detection of breast cancer is a tedious task, researchers have used CNN in order to confirm whether the tissue in question is benign or malignant. By training a CNN model, researchers were able to achieve a whopping accuracy of 99.86% [24]. Using high dimensional expression of genes, a novel CNN – NPR module was able to confirm the origin tissue of cancer whilst also deducing the type of cancer. [25]

CNN is also useful with its classification technique in the arena of agriculture by providing large scale snapshots of agriculture environments and prediction of the health and harvest of crops [26]. Researchers have also used CNN to classify Paddy crop [27]. CNNs are also used to predict shares and stock prices by applying LSTM with CNN [28]. Another application is predicting financial time series using CNN with chaos theory and Polynomial Regression (PR) [29]. The efficacy of the model is tested with the help of using different foreign rates of exchange, commodity prices and stock market indices.

On the legal front CNN has been used multiple times to classify, translate, predict and summarize information [30,31]. CNN is highly useful for people who have to pore over each legal document and comes in handy in legal cases where extensive and repetitive tasks are present in abundance. The legal judgements are also predicted with the help of hybridized models of CNN and BiGRU [32]. Results show evidently that using CNN and RNN alone does not output better results as compared to using the combination of CNN and BiGRU.

One of the major uses of CNN lies in Computer Vision [33], which epitomizes the concept of extracting information from a given visual data. The most common practical usage in this field comes into play with Face Detection. For instance, in 2016 Zhang implemented a modified "multitasking and cascaded" CNN to solve face discernment problems. CNN models are capable of producing 3D objects from 2D images, which is inspired from sculpting process [34]. A CNN model is also able to detect roofs from good quality satellite images [35].

In the field of Gaming CNN has proved its effectiveness multiple times from predicting outcomes in strategical games like StarTrek II [36] to object tracking in games [37]. Using GRAD-CAM the results obtained from the last layer of CNN is applied to obtain visual explanations with respect to StarTrek II. A generative model called Rodin is capable of creating 3D avatars using Diffusion [38]. With the help of a Deep ConvNet model for discriminative and generative voxel modeling the results are enhanced.

Wang proposed a three-dimensional CNN model to identify diverse activities from video frames. A CNN model is made to learn portrayals that center on the main object by changing the image with different versions of background, and the separation of the main object and the background is managed by Tobias [39]. In the domain of Speech Recognition, in 2012 Hamid presented a voice recognition system that was not speaker dependent [40]. CNN models are also implemented to classify videos on a huge scale [41].

5. Extensions to CNN

CNN has served well when serving alone in various applications [42]. CNN is great for the purposes of extraction of features but many researchers have observed better accuracy and better results when CNN is hybridized with other models. Many hybrid variations have been introduced in CNN models to serve some solutions to aim at extremely specific problems.

An example can be a Convolutional Network and conclusion deriving model for Visual fault Identification of High-Speed Train images [43], which was developed to cater to problems in identifying mechanical parts of trains that are in urgent need of repair due to exposure to environmental conditions, which leads to problems such as corrosion of mechanical parts. The model used is fundamentally similar to CNN Classification models but has been specifically constructed for this problem. To derive a system to retrieve images based on content, CNN is diffused with Computer Vision learning techniques for better efficacy and results [44].

A combination of CNN mixed with DNN and LSTM to detect Abnormal Flow in Software Defined Network based Smart Grid was constructed by Pengpeng Ding and others [45]. A combination of Deep CNN with a focal loss function is used for classification of skin lesion [46].

Another example of increasing efficiency of regular CNN models with the use of Hybrid CNN is seen in Hybrid CNN models for Sentiment Analysis [8]. CNN, LSTM and SVM are combined to improve a normal CNN model for sentiment Analysis. The famous combination of CNN, LSTM and SVM is also used in identifying patients suffering from novel coronavirus by studying X-radiational images [47,48]. Here, CNN is used for Deep Feature Extraction and LSTM is used for detection of the diseases using the deep feature. CNN and LSTM work in a combined fashion to give expected results with good efficacy.

Another example of bringing together two algorithms with striking differences is seen when a hybrid of circumstantial-dependent CNN with deep structure and pixel dependent multiple layer perceptions with shallow architecture are brought together [49]. MPL is again brought alongside CNN to differentiate between patients infected with Covid 19 and non-infected patients with the help of X-radiational images of patients, resulting in faster screening and early treatment [50]. CNN and MLP are also used in a model for analyzing network attacks [51] and claims better results than that observed from using separate CNN and MLP models.

In the Hybrid CNN-LSTM Model for small term House Charge Forecasting, LSTM layers are used for sequential learning with the layers of CNN for feature extraction [52]. In the case of forecasting PV Power prediction using CNN-LSTM Hybrid models [53], with the help of a five film CNN-LSTM hybrid model, photovoltaic power predictions are made. Again, for extracting the local features CNN acts as a filter, and for detecting temporal features the LSTM network comes into play [53]. A hybrid model of CNN and LSTM can also be used to forecast the amount of particulate matter in a city as shown in [54]. Here, CNN is used for figuring out the air quality and LSTM is used for showcasing the long term series process of input data [54].

CNN can also be assorted with twin support vector machines (TWSVM) thereby increasing the efficiency of a classic CNN model for traffic sign recognition tasks [55]. To predict human activity, CNN has been paired with TWSVM [56] where in features were extracted using CNN and Human Activity recognition with a weighted TWSVM. The weighted TWSVM ensures lower computational cost and gets the results more speedily.

For delivering a forecast of Remaining Useful Life for the betterment of maintenance strategies, a Hybrid model of CNN and BiLSTM can be used [57]. CNN provides spatial extraction of features whereas BiLSTM takes care of the time dimension of the data. This model is extremely useful for timely maintenance and development of parts and also in minimizing the industrial losses in quality and inspection check timelines. CNN is also used in Stock Price prediction along with BiLSTM [58]. Here, the quality for feature extraction of CNN comes in handy whereas the prediction is handled smoothly by BiLSTM.

In the field of autoencoders, a Deep AE can be used along with CNN to the diagnosis of fault in a data driven approach in induction motors [59]. A similar approach is followed for the detection of forgery in classified documents with the help of CAE and Logical Regression [60].

Based on the facts and information pertaining to a legal case, CNN is used with a BiGRU model to predict legal court judgements [32] that resulted in better accuracy than using CNN alone. Another example of a hybrid of CNN and BiGRU is seen in analyzing sentiments and intents in student texts [61].

In the field of Generative Adversarial Networks, a fully connected Convolutional GAN serves better picture quality than regular everyday convolution architecture [62]. The classification accuracy of CNN can also be improvised by using GAN to augment the existing dataset of Electroluminescence images [63].

Table 1. Hybrid CNN models for solving specific applications

Various Combinations of CNN	Area of Application
LSTM	• Short-Term Individual Household Load Forecasting [52] • Forecasting PV Power prediction [53] • Forecasting Particulate Matter [54]
LSTM + SVM	• Sentiment Analysis [47] • Discernment of novel coronavirus using X-radiation images [48]
TWSVM	• Traffic Sign Recognition [55] • Context-Aware Human Activity Recognition [56]
BiLSTM	• Prognostic Health Management (PHM) with (RUL)>Remaining Useful Life [57] • Stock Price Prediction [58]
Multilayer CNN	• Defect Identification in High -Train Images [43] • Improved CBIR System Using Multilayer CNN [44]
MLP	• Very Fine Resolution Remotely sensed image classification [49] • Differentiate between Covid positive and negative patients [50] • Novel Network Traffic Attack [51]
Deep Neural Network	• Uncommon course of flow in Software Described Network based Smart Grid [45] • Epidermis Lesion Classification of Dermoscopic photos [46]
Deep Autoencoder	• Fault Classification in Induction Motors [59] • Deep feature extraction for document forgery [60]
BiGRU Model	• Predict Legal judgements [32] • Sentiment Analysis [61]
GAN	• Fully Connected CGAN for higher picture Quality [62] • Improving classification efficiency of CNN [63]

It is quite evident that the Hybrid versions of CNN are quite prominent to better predict results and improve accuracy. The major contributions where CNN is hybridized for various applications is summarized in Table 1. Table 1 shows multiple instances where CNN has been hybridized and encouraging results are achieved.

CNN is working efficiently and achieving baffling results when intermixed with several other newfangled technologies and thereby broadening the horizon of the working of CNN in different fields. Here, the most evident use of CNN is its feature extraction, which when combined with other models produces desired results. The use of CNN is now visibly not limited to narrowed down fields but researchers are constantly reimagining the possibilities that come along with CNN.

6. Future Challenges with CNN

CNN and its applications are increasing and the areas of research are expanding day by day. Some of the major emerging research areas with respect to CNN have been discussed. There is no denial in the use of CNN in the picture segmentation tasks, but in this arena too CNN's generative learning ability is highly limited. There is a limited transition in the models being able to apply the concept of GAN [64] to speech and other forms of data.

One of the biggest challenges in the face of CNN is lack of sufficient computing resources and hardware deficiency. This issue is mainly visible in the implementation of Generative Adversarial Networks. The use of supercomputers and other modern GPU based equipment are gradually serving as solutions to this problem. The use of weighted TWSVM for lower computational costs is another helpful technique. Another major challenge with CNN is its inability to be used in real time, which can only be solved with using high computational machines and resources. Some common accelerators presented to overcome this problem are Eyeriss, FPGA, and some application specific Integrated Circuits [IC]. A distributed library of machine learning that is capable of executing model comparability was introduced by Google Group with GPipe. Another solution to the above-mentioned problems lies in Cloud-based platforms, which provide an exceptional environment to users at an extremely reasonable cost. Human activity recognition is another trending area of research in CNN and various researchers have contributed in human activity and pose recognition at an encouraging level.

The emergence and escalating use of CNN is extremely clear. CNN is widely popular because of its malleable characteristics and its ability to evolve for subsequent problems. As seen above, CNN and its applications are always in a constant rush to progress and advance, which results in better enactment of various applications.

7. Conclusion

The evolution of Convolutional Neural Networks has seen numerous ups and downs. It started from the genesis of LeNet in 1989 traversing through AlexNet, ZefNet, ResNet, ResNext, WideResNet, and ending with YOLOX. The extensions of CNN proposed by various eminent researchers for specific purposes have also seen gradual and steady changes. From solving the problem of Visible defect Identification for High-Speed Train Images, using Autoencoders to predict document forgery, using CNN to diagnose Covid 19 patients, to filtering out movie reviews according to the sentiments available by customer reviews, and estimating remaining life of important mechanical parts of devices, CNN has seen numerous extensions. Although CNN is capable of producing great results on datasets, the problems concerning limited amount of data still exists. Therefore, the area related with data augmentation and generating sample data should be researched more to alleviate this problem. The use of CNN not just alone but in a combination with LSTM, SVM, BiSLTM, MLP, Deep Neural Networks, BiGRU and Deep Autoencoders showed encouraging results. CNN is now not limited to its traditional capacity of extraction of features. The use of cloud-based architectures for the implementation of CNN based applications could prove to be immensely helpful with respect to the issues regarding enormous and gigantic level of data to be processed [1]. The problem of portability and computational costs is handled with this method.

CNN has been reimagined to serve specific purposes revolving around solving real world problems. From Medical, Agriculture, Architectural to Legal Security and Gaming and other recreations, CNN has proven its usefulness multiple times. The possibilities with CNN are endless and this is encouraging researchers to reconceive the solutions for numerous problems.

References

1. LeCun, Y., Jackel, L.D., Bottou, L., Cortes, C., Denker, J.S., Drucker, H., Guyon, I., Muller, U.A., Sackinger, E., Simard, P. and Vapnik, V. Learning algorithms for classification: A comparison on handwritten digit recognition. *Neural networks: the statistical mechanics perspective*, vol. 261, no. 276, pp. 2, 1995.
2. Khan, A., Sohail, A., Zahoora, U. and Qureshi, A.S. A survey of the recent architectures of deep convolutional neural networks. *Artificial intelligence review*, vol. 53, pp. 5455-5516, 2020
3. Hasan, R.I., Yusuf, S.M. and Alzubaidi, L. Review of the state of the art of deep learning for plant diseases: A broad analysis and discussion. *Plants*, vol. 9, no. 10, pp. 1302, 2020.
4. Xiao, Y., Tian, Z., Yu, J., Zhang, Y., Liu, S., Du, S. and Lan, X. A review of object detection based on deep learning. *Multimedia Tools and Applications*, vol. 79, pp. 23729-23791, 2020.
5. Ker, J., Wang, L., Rao, J. and Lim, T. Deep learning applications in medical image analysis. *IEEE Access*, vol. 6, pp. 9375-9389, 2017
6. Zhang, Z., Cui, P. and Zhu, W. Deep learning on graphs: A survey. *IEEE Transactions on Knowledge and Data Engineering*, vol. 34, no. 1, pp. 249-270, 2020
7. Shrestha, A. and Mahmood, A. Review of deep learning algorithms and architectures. *IEEE access*, vol. 7, pp. 53040-53065, 2019
8. Krizhevsky, A., Sutskever, I. and Hinton, G.E. Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, vol. 60, no. 6, pp.84-90, 2017
9. Alzubaidi, L., Zhang, J., Humaidi, A.J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M.A., Al-Amidie, M. and Farhan, L. Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. *Journal of big Data*, vol. 8, pp.1-74, 2021.
10. Zeiler, M.D. and Fergus, R. Visualizing and understanding convolutional networks. In *Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part I 13*, Springer International Publishing, pp. 818-833, 2014.
11. Simonyan, K. and Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.
12. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V. and Rabinovich, A. Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1-9, 2015
13. Srivastava, R.K., Greff, K. and Schmidhuber, J. Highway networks. *arXiv preprint arXiv:1505.00387*, 2015
14. He, K., Zhang, X., Ren, S. and Sun, J. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770-778, 2016
15. Wu, S., Zhong, S. and Liu, Y. Deep residual learning for image steganalysis. *Multimedia tools and applications*, vol. 77, pp.10437-10453, 2018.
16. Huang, G., Liu, Z., Van Der Maaten, L. and Weinberger, K.Q. Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4700-4708, 2017
17. Xie, S., Girshick, R., Dollár, P., Tu, Z. and He, K. Aggregated residual transformations for deep neural networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1492-1500, 2017

18. Zagoruyko, S. and Komodakis, N. Wide residual networks. *arXiv preprint arXiv:1605.07146*, 2016
19. Han, D., Kim, J. and Kim, J. Deep pyramidal residual networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 5927-5935, 2017
20. Chollet, F. Xception: Deep learning with depthwise separable convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1251-1258, 2017
21. Wang, J., Sun, K., Cheng, T., Jiang, B., Deng, C., Zhao, Y., Liu, D., Mu, Y., Tan, M., Wang, X. and Liu, W. Deep high-resolution representation learning for visual recognition. *IEEE transactions on pattern analysis and machine intelligence*, vol. 43, no. 10, pp.3349-3364, 2020
22. Cheng, B., Xiao, B., Wang, J., Shi, H., Huang, T.S. and Zhang, L. Higherhrnet: Scale-aware representation learning for bottom-up human pose estimation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 5386-5395, 2020.
23. Dong, Y., Pan, Y., Zhang, J. and Xu, W. Learning to read chest X-ray images from 16000+ examples using CNN. In *2017 IEEE/ACM International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE)*, IEEE, pp. 51-57, 2017, July.
24. Dabeer, S., Khan, M.M. and Islam, S. Cancer diagnosis in histopathological image: CNN based approach. *Informatics in Medicine Unlocked*, vol. 16, pp. 100231, 2019.
25. Gangurde, R., Jagota, V., Khan, M.S., Sakthi, V.S., Boppana, U.M., Osei, B. and Kishore, K.H. Developing an Efficient Cancer Detection and Prediction Tool Using Convolution Neural Network Integrated with Neural Pattern Recognition. *BioMed research international*, 2023.
26. Kamilaris, A. and Prenafeta-Boldú, F.X. A review of the use of convolutional neural networks in agriculture. *The Journal of Agricultural Science*, vol. 156, no. 3, pp. 312-322, 2018
27. Neforawati, I., Herman, N.S. and Mohd, O. Precision agriculture classification using convolutional neural networks for paddy growth level. In *Journal of Physics: Conference Series* (vol. 1193, no. 1, pp. 012026). IOP Publishing, 2019, April.
28. Mehtab, S. and Sen, J. Analysis and forecasting of financial time series using CNN and LSTM-based deep learning models. In *Advances in Distributed Computing and Machine Learning: Proceedings of ICADCML 2021* (pp. 405-423). Springer Singapore, 2022.
29. Durairaj, D.M. and Mohan, B.K. A convolutional neural network based approach to financial time series prediction. *Neural Computing and Applications*, vol. 34, no. 16, pp.13319-13337, 2022
30. Bansal, N., Sharma, A. and Singh, R.K. A review on the application of deep learning in legal domain. In *Artificial Intelligence Applications and Innovations: 15th IFIP WG 12.5 International Conference, AIAI 2019, Hersonissos, Crete, Greece, May 24–26, 2019, Proceedings 15* (pp. 374-381). Springer International Publishing, 2019.
31. Undavia, S., Meyers, A. and Ortega, J.E. A comparative study of classifying legal documents with neural networks. In *2018 Federated Conference on Computer Science and Information Systems (FedCSIS)*, IEEE, pp. 515-522, 2018, September.
32. Wang, C. and Jin, X. Study on Prediction of Legal Judgments Based on the CNN-BiGRU Model. In *Proceedings of the 2020 6th International Conference on Computing and Artificial Intelligence*, pp. 63-68, 2020, April.
33. Bhatt, D., Patel, C., Talsania, H., Patel, J., Vaghela, R., Pandya, S., Modi, K. and Ghayvat, H. CNN variants for computer vision: history, architecture, application, challenges and future scope. *Electronics*, vol. 10, no. 20, pp. 2470, 2021.
34. Stigeborn, P. Generating 3D-objects using neural networks, 2018.
35. Hang, L. and Cai, G.Y. CNN based detection of building roofs from high resolution satellite images. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 42, pp.187-192, 2020.
36. Baek, I. and Kim, S.B. 3-Dimensional convolutional neural networks for predicting StarCraft II results and extracting key game situations. *Plos one*, vol. 17, no. 3, pp. e0264550, 2022.
37. Venkatesh, A. Object tracking in games using convolutional neural networks, 2018.
38. Wang, T., Zhang, B., Zhang, T., Gu, S., Bao, J., Baltrusaitis, T., Shen, J., Chen, D., Wen, F., Chen, Q. and Guo, B. Rodin: A generative model for sculpting 3d digital avatars using diffusion. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 4563-4573, 2023.
39. Cao, Y.H. and Wu, J. A Random CNN Sees Objects: One Inductive Bias of CNN and Its Applications. In *Proceedings Of The AAAI Conference On Artificial Intelligence*, vol. 36, no. 1, pp. 194-202, 2022, June.
40. Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., Liu, T., Wang, X., Wang, G., Cai, J. and Chen, T. Recent advances in convolutional neural networks. *Pattern recognition*, vol. 77, pp.354-377, 2018.
41. Karpathy, A., Toderici, G., Shetty, S., Leung, T., Sukthankar, R. and Fei-Fei, L. Large-scale video classification with convolutional neural networks. In *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*, pp. 1725-1732, 2014.
42. Alawad, W., Alburaidi, B., Alzahrani, A. and Alflaj, F. A comparative study of stand-alone and hybrid CNN models for COVID-19 detection. *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 6, 2021.
43. Wang, Z., Peng, J., Song, W., Gao, X., Zhang, Y., Zhang, X., Xiao, L. and Ma, L. A convolutional neural network-based classification and decision-making model for visible defect identification of high-speed train images. *Journal of Sensors*, pp.1-17, 2021.
44. Kapadia, M.R. and Paunwala, C.N. Improved CBIR system using Multilayer CNN. In *2018 International Conference on Inventive Research in Computing Applications (ICIRCA)* (pp. 840-845). IEEE, 2018, July.
45. Ding, P., Li, J., Wang, L., Wen, M. and Guan, Y. HYBRID-CNN: An efficient scheme for abnormal flow detection in the SDN-Based Smart Grid. *Security and communication networks*,, pp.1-20, 2020.
46. Nie, Y., Sommella, P., Carratù, M., O'Nils, M. and Lundgren, J. A Deep CNN Transformer Hybrid Model for Skin Lesion Classification of Dermoscopic Images Using Focal Loss. *Diagnostics*, vol. 13, no. 1, pp. 72, 2022.

47. Dang, C.N., Moreno-García, M.N. and De la Prieta, F. Hybrid deep learning models for sentiment analysis. *Complexity*, pp.1-16, 2021.
48. Islam, M.Z., Islam, M.M. and Asraf, A. A combined deep CNN-LSTM network for the detection of novel coronavirus (COVID-19) using X-ray images. *Informatics in medicine unlocked*, vol. 20, p.100412, 2020.
49. Zhang, C., Pan, X., Li, H., Gardiner, A., Sargent, I., Hare, J. and Atkinson, P.M. A hybrid MLP-CNN classifier for very fine resolution remotely sensed image classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 140, pp.133-144, 2018.
50. Ahsan, M.M., E. Alam, T., Trafalis, T. and Huebner, P. Deep MLP-CNN model using mixed-data to distinguish between COVID-19 and Non-COVID-19 patients. *Symmetry*, vol. 12, no. 9, pp. 1526, 2020.
51. Kumar, V., Rana, K., Malik, J. and Tomar, A. Evaluating hybrid cnn-mlp architecture for analyzing novel network traffic attacks. *International Journal of Scientific & Technology Research*, vol. 9, no. 3, pp. 2277-8616, 2020.
52. Alhussein, M., Aurangzeb, K. and Haider, S.I. Hybrid CNN-LSTM model for short-term individual household load forecasting. *IEEE Access*, vol. 8, pp.180544-180557, 2020.
53. Tovar, M., Robles, M. and Rashid, F. PV power prediction, using CNN-LSTM hybrid neural network model. Case of study: Temixco-Morelos, México. *Energies*, vol. 13, no. 24, pp. 6512, 2020.
54. Li, T., Hua, M. and Wu, X.U. A hybrid CNN-LSTM model for forecasting particulate matter (PM2. 5). *IEEE Access*, vol. 8, pp.26933-26940, 2020.
55. Sun, Y. and Chen, L. Traffic sign recognition based on CNN and twin support vector machine hybrid model. *Journal of Applied Mathematics and Physics*, vol. 9, no. 12, pp.3122-3142, 2021.
56. Chui, K.T., Gupta, B.B., Torres-Ruiz, M., Arya, V., Alhalabi, W. and Zamzami, I.F. A Convolutional Neural Network-Based Feature Extraction and Weighted Twin Support Vector Machine Algorithm for Context-Aware Human Activity Recognition. *Electronics*, vol. 12, no. 8, pp. 1915, 2023
57. Zhao, C., Huang, X., Li, Y. and Yousaf Iqbal, M. A double-channel hybrid deep neural network based on CNN and BiLSTM for remaining useful life prediction. *Sensors*, vol. 20, no. 24, pp. 7109, 2020
58. Lu, W., Li, J., Wang, J. and Qin, L. A CNN-BiLSTM-AM method for stock price prediction. *Neural Computing and Applications*, vol. 33, pp.4741-4753, 2021.
59. Toma, R.N., Piltan, F. and Kim, J.M. A deep autoencoder-based convolution neural network framework for bearing fault classification in induction motors. *Sensors*, vol. 21, no. 24, pp. 8453, 2021
60. Jaiswal, G., Sharma, A. and Yadav, S.K. Deep feature extraction for document forgery detection with convolutional autoencoders. *Computers and Electrical Engineering*, vol. 99, pp. 107770, 2022
61. Yan, W., Zhou, L., Qian, Z., Xiao, L. and Zhu, H. Sentiment analysis of student texts using the CNN-BiGRU-AT model. *Scientific Programming*, pp.1-9, 2021
62. Barua, S., Erfani, S.M. and Bailey, J. FCC-GAN: A fully connected and convolutional net architecture for GANs. *arXiv preprint arXiv:1905.02417*, 2019.
63. Luo, Z., Cheng, S.Y. and Zheng, Q.Y. GAN-based augmentation for improving CNN performance of classification of defective photovoltaic module cells in electroluminescence images. In *IOP Conference Series: Earth and Environmental Science* (vol. 354, no. 1, pp. 012106). IOP Publishing, 2019, October.
64. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł. and Polosukhin, I. Attention is all you need. *Advances in neural information processing systems*, 2017.