

A Study on Deep Learning Techniques for Digital Image Processing

Nisha Shrivastava¹ and Dr.P. K. Rai²

Department of Computer Application, APS University, Rewa (MP)

Email - 1 er.shrivastavanisha@gmail.com and 2 pkrapso@gmail.com

Abstract:

There are a lot of images on the internet, and the number is always increasing day by day. The amount of data that needs to be stored, examined, and analysed has significantly increased because of AI producing more images than photographers have taken in the previous many decades. Because of their increased quantity and quality, processing these high-resolution photographs have become more time- and resource-intensive. In this paper we have explored deep learning techniques for digital image processing.

1. Introduction:

One of the main areas of artificial intelligence is deep learning. Studying adaptable computer algorithms with learning capabilities is the goal of deep learning [1]. The learning that takes place is always founded on a certain type of experience based on instructions or known input data.

Deep Learning can specifically increase the calibre and speed of research in the field of medical image analysis. The two primary goals of image processing are image analysis and image enhancement. Works demonstrating a variety of applications where the need for enhancement is evident are projected in the literature among image enhancing approaches.

To gain a deeper comprehension of the photographs, it is also beneficial to analyse them. In this process, methods such as clustering and classification are mostly taken into consideration. With the technology developed, a wide range of image processing applications emerged, from standard image processing to medical image analysis. Given that the number of medical cases is rising in proportion to the population, a thorough analysis of the photos is necessary to gain a better knowledge of the diseases. Many medical images segmentation techniques, including edge-based improvements, shape-based enhancements, region-based descriptors, texture-based techniques, and colour-based techniques, have been developed based on various methodologies to accomplish this goal [2-7].

The methods are regarded as degenerative models; in addition to these, other models highlighted in the literature are feature-based, depth-based, and model-based approaches. However, since the analysis in this process is predicated on parameter estimates and the entire analysis is predicated on these parameters, it is expected that model-based techniques or generative approaches are thought to be more effective.

In addition to these methods, deep learning techniques emerged with the most recent developments in machine learning. These methods believe to be more promising, especially when it comes to examining medical images. This paper presents the several image processing domains where deep learning methods are more appropriate through a thorough investigation [8-11]. An overview digital image processing is given in Section 2 of the paper. The description of Deep learning is given in Section 3. The applications of CNN in deep learning algorithms into content based digital image retrievals are presented in Section 4. The use of deep learning approaches for image processing has been explored in Section 5. The article is summarized in the Section 6.

2. Digital Image Processing:

Digital image processing is the collection of methods used to alter a digital image to enhance it, make it smaller, or extract information from it. Four categories of image processing can be: Image synthesis, image analysis, image coding with data compression, and restoration and enhancement. **Examples of images that are present in our environment includes:** Photographic photos of nature, Engineering and artistic drawings and Scientific photos (medical, satellite, etc.).

A digital image is made up of a limited number of components, each of which has a specific location and value. These components are referred to as pels, pixels, picture elements, or image elements. The most common term for the components of a digital image is "pixel." [12]

A two-dimensional function that depicts a measurement of a feature, like the colour or brightness of a picture being observed, is called an image. A 3D scene projected onto a 2D projection plane is called an image. The amplitude off at any pair of coordinates is known as the intensity of the image at that location. An image can be described as a two-dimensional function $f(x, y)$, where x and y are spatial (plane) coordinates.

The term grey level is used often to refer to the intensity of monochrome images. Colour images are formed by a combination of individual 2-D images. For example: In RGB colour system, a colour image consists of three (Red, Green and Blue) individual component images [12].

2.1 Digital image processing applications:

Applications for digital image processing are numerous and include:

- Remote sensing
- Business applications for image transmission and storage
- Processing medical data Radio Detection and Ranging (RADAR) and Sound Navigation and Ranging (SONAR)
- Acoustic image processing (underwater acoustics or hydro acoustics is the study of underwater sound.)
- Automated industrial part inspection and robotics. Earth resource tracking, geographic mapping, crop prediction, urban growth and weather monitoring, and flood and fire control are all aided by satellite imagery.

Applications for space images include:

- Identifying and analysing items in pictures taken by deep space probes.
- Applications for image transmission and storage take place in broadcast television.
- Conference calls Facsimile image transmission for office automation (printed papers and visuals). Interaction via computer networks,
- Security monitoring systems based on closed-circuit television and in communications within the military.

Applications in medicine:

- Chest X-ray processing.
- Cine angiograms.
- Trans axial tomography projection pictures.
- Nuclear Magnetic Resonance (NMR) medical imaging; and
- Ultrasonic scanning

3. Deep Learning:

Deep Learning (DL) techniques have been evolving in unison with this complex enhancement in the image collection. A fundamental concept in many fields, such as computer science, engineering, and medicine, is image processing. Algorithms used in deep learning techniques learn from experience, or what is referred to as input data. Neural networks are the focus of the machine learning subfield known as "deep learning." The purpose of this study is to investigate the latest advancements in Deep Learning techniques, including their applications in image processing [13]. It also offers details and perspectives on the most recent developments in the field as well as the challenges and limitations that come with using Deep Learning to process photos. The most recent advancements are also discussed in this paper [14].

3.1 Deep Learning architecture:

DL architecture can be divided into supervised DL and unsupervised DL:

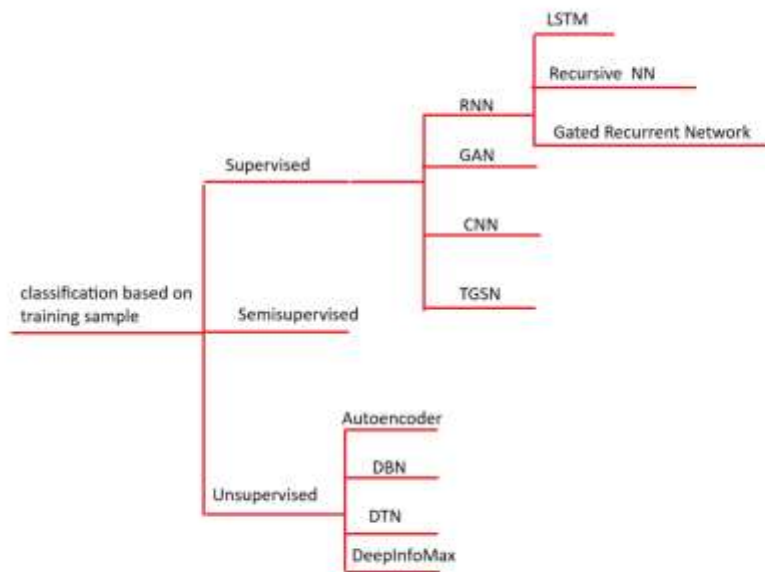


Fig – 1: Breakdown of Deep Learning architecture [16]

Supervised Deep learning models are:

- Recurrent Neural Networks (RNNs),
- Long Short-Term Memory (LSTM),
- Gated Recurrent Unit (GRU),
- Convolutional Neural Networks (CNNs),
- Generative Adversarial Network (GAN).

Unsupervised deep learning models are:

- Deep Belief Networks (DBN),
- Deep Transfer Network (DTN),
- Tensor Deep Stack Networks (TDSN),
- Autoencoders (AE).

4. Convolutional neural networks:

A paradigm for information processing, Artificial Neural Networks (ANNs) are primarily influenced by biological nervous systems. It is made up of numerous processing units known as neurons that cooperate to solve a particular assignment. Like a biological system, the learning process in an ANN entail modifying the connections between the processing units. A subfield of convolutional neural networks that is better suited for evaluating visual pictures is deep learning methodology. Convolutional Neural Networks are a component of machine learning techniques that deal with knowledge representation [15]. By splitting the photos into layers, these deep learning techniques aid in a better understanding of the parameters by allowing each layer to be examined and more precisely evaluated in contrast to the traditional analysis process. When the deep learning approach is applied, several layers are created, including input, output, vertical, and horizontal layers. The product of the horizontal vertical and the number of channels in use can be used to determine the number of hidden layers.

4.1 Convolution Layer:

It is possible to think of the image's passage through the Convolutional Layer as a feature extraction process. Before delving into the Convolutional Layer, let's examine how computer vision and human eyesight differ from one another [17]. For example, brightness, size, and contour help to visually a grayscale image of an 'a'. This 'a' image is a matrix of solely integer in computer vision, as seen in Fig - 2.

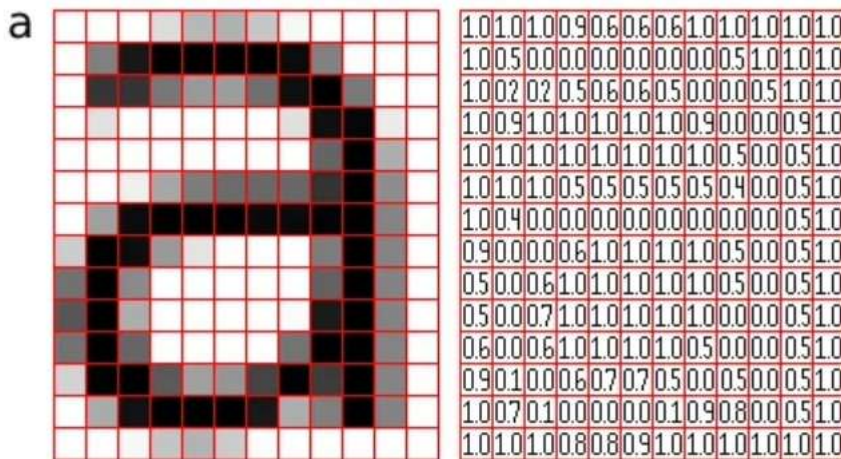


Fig - 2: Images in human vision (left) and computer vision (right) [20]

A computer must extract the image's features from this matrix to learn it. One such procedure is picture convolution. As an example, we select a 3x3 matrix known as a filter (or convolution kernel) that moves with a step size of 1 across a 5x5 image. Every time the filter slides, its values are multiplied by the image's value, and the total of these multiplications is calculated. An element of the feature matrix is the resultant value. The image's feature matrix can ultimately be obtained after traversing the complete image.

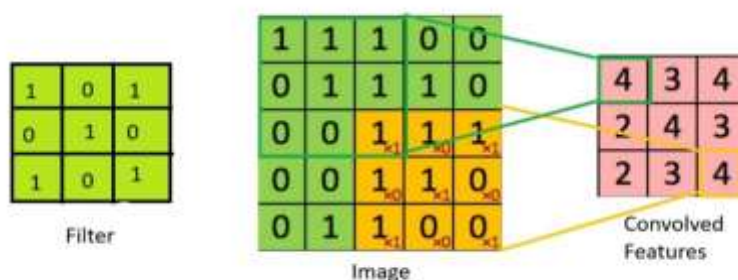


Fig - 3: Convolutional layer [23]

4.2 Pooling Layer:

A pooling layer is frequently inserted in between convolution layers in convolutional neural networks. Both the size of the parameter matrix and the number of parameters in the final fully linked layer can be efficiently decreased by the pooling layer. The computation can be sped up and over-fitting avoided by using the pooling layer. A pooling layer must be added between the convolution layers in image recognition when the trained image is too big. This lowers the number of training parameters. The image's depth doesn't change because pooling is done in all depth dimensions. Max pooling is the most widely used pooling form. Maximum pooling size 2x2 pixels make up a typical filter. Stride, measured in pixels, is the separation between each extracted tile [18]. It differs from convolutional ones in that the convolution filters slide pixel by pixel throughout the feature map, however in max pooling, step establishes the points

of origin for every tile. For a 2x2 filter size, a phase of 2 is specified, meaning that all non-overlapping 2x2 tiles are extracted from the function map by the max pooling operation.



Fig - 4: Pooling Layer[16]

4.3 Fully Connected Layer:

The fully connected layer, the last component of a convolutional neural network, is frequently employed in classification tasks. It transfers the outputs of formal layers into the classification task's goals. Assume that we have five outputs from the formal convolution and pooling layers, which we will map into three categories. The three categories are the targets for the classification task, and the outputs of the fully connected layer are the five outputs, which are known to be the key features that can help us determine which input image belongs to which category.

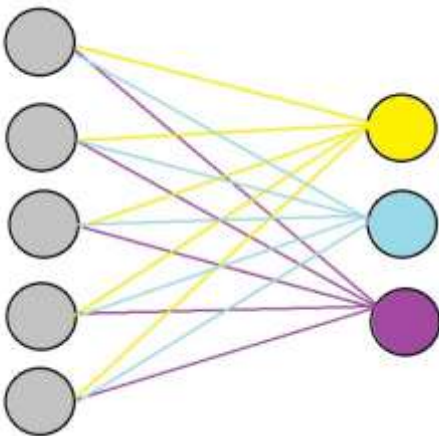


Fig - 5: Fully Connected Layer [22]

5. Digital Image Processing using Deep Learning techniques:

Image processing deals with a lot of complex concepts. There are two main application areas to think about: processing images datasets for transmission, storage, and analysis for machine interpretation, and improving and developing pictorial data for human comprehension and interpretation. Computational machines can examine additional electromagnetic spectrum ranges, such as radio waves, gamma waves, and x-rays, while humans can only observe one of the five senses, vision, which is limited to a specific range [19]. This holds greater significance and importance in the domains of astronomy and medicine. Consequently, the significance of image processing and related methods increases.

5.1 Data Acquisition and Preparation:

To apply deep learning techniques effectively, the first step is to collect an appropriate dataset, which typically begins with image collection. The size and quality of the dataset play a crucial role in determining the model's performance. Corrupt files, duplicate entries, and irrelevant data can negatively impact the training process [20]. Hence, the dataset must be cleaned before use. This process is referred to as data cleaning. For tasks such as object detection, image segmentation, and corruption analysis, data annotation is essential. To enable supervised learning, it is important to

label the objects accurately and consistently, ensuring that the model can learn meaningful patterns and generalize well to new data. After this, data pre-processing is done to give input to the neural network [21]. For this, techniques like Resize, Normalization, Data Augmentation are used. This technique helps in increasing the size of the data set and model generation from the artificial rope.

5.2 Choosing the Right Architecture:

The design of a neural network dictates the quality of fitting it may have on assimilating spatial, temporal, or generative characteristics of image data. Some of the deep learning architectures commonly used according to their applicability to various image processing tasks are below:

- **Convolutional Neural Networks (CNNs):** CNNs became the backbone of current image processing, and this kind of neural network is successful in image classification, object tracking and semantic segmentation tasks. They are made to extract spatial hierarchies of images with localized filters [22].
- **Recurrent Neural Networks (RNNs) and Transformers:** The tasks that need to work with time or sequence-based data are perfect to be performed with these architectures, including analysing video frames or captioning an image. They make context-sensitive use of image sequences possible when used along with CNNs.
- **Generative Adversarial Networks (GANs):** Generative Adversarial Networks (GANs) comprise the generator and discriminator which are trained adversarially. They excel especially when it involves creation or modifying pictures.
- **Autoencoders:** They are unsupervised models aimed at learning compact representations, thus referring to the input data. They can be applied in dealing with roles such as noise reduction, dimensionality reduction, and anomaly detection.

5.3 Customization and Layer Design:

Convolutional Neural Network (CNN) based models are usually built up of different types of layers such as Convolutional Layers, Pooling Layers, Fully Connected Layers, and Normalization and Regularization Layers. When developing a custom architecture design, it is very important to maintain a balance between the depth of the model, the total number of parameters, and its computational efficiency [20]. The purpose of this balance is to keep the performance of the model optimal, so that neither the problem of overfitting arises, nor the performance degrades due to underfitting. Appropriate layer combination and hyperparameter selection help in improving the generalization ability of the model [24].

5.4 Training the Model:

The process of training the Deep Learning model involves several important steps. The first step is to define the loss function, which quantifies the error between the prediction of the model and the ground truth label. The loss function controls the learning direction of the model, and its selection has a direct impact on the efficiency of training [25]. Then, the appropriate optimizer is selected, which updates the weights of the model sequentially such that the value of the loss function becomes minimum.

During the training process, input data is fed to the model in batches, and the model's weights are updated based on the gradients computed from the loss. Throughout this process, the model's performance is monitored on both the training and validation datasets to identify and address issues such as overfitting and underfitting. Additionally, the early stopping technique is employed, wherein the training is halted at the optimal point where the model's generalization ability is at its peak. This prevents the model from overtraining and helps maintain its performance on unseen data.

5.5 Model Evaluation:

After the training process, the performance of the model is evaluated on the Test Dataset, which contains data that the model would never have seen during training. Its objective is to impartially assess the generalization ability of the model. This evaluation is made through various evaluation metrics, which provide information about various aspects

of the model performance [26]. For example, specific metrics for tasks like Image Classification, Object Detection, and Image Segmentation are used such as Accuracy, Precision, Recall, F1-SCORE, Intersection Over Union (IOU) etc.

The results obtained after the evaluation are analysed, so that the strengths and weaknesses of the model can be understood. This analysis provides guidance towards the improvement and adaptation of models in future.

5.6 Deployment and Inference:

Deploying the model means making the trained Deep Learning model available for real-world applications. For this, mediums like Cloud Server, Edge Device, or Mobile Platform can be used. The selection of the appropriate platform depends on the requirements of the application and the available resources [27]. After the model is deployed, it processes new and previously unrecognized image inputs to make predictions. This stage is called the Inference stage. During inference the model performs the same function for which it was trained.

5.7 Monitoring and Maintenance:

Post-deployment monitoring and maintenance of the model are essential processes that ensure that the model continues to function effectively and reliably. The purpose of monitoring is to monitor the accuracy, reliability and response time of the model, especially when the model is processing a new image [28]. Additionally, the model may need to be retrained or fine-tuned from time to time, especially when the data distribution used in the real-world changes. This process helps in maintaining the generalizability of the model so that it can make accurate predictions even in unexpected or unseen scenarios [29].

6. Conclusion:

Deep Learning models are more appropriate for image processing jobs because they utilize neural networks, which are significantly more accurate and reliable. Furthermore, we observed that machine learning algorithms perform far better on data-driven sets than on big, complicated datasets. We also explored the several stages involved in deep learning techniques for image processing. By examining specific phases and layers, we discovered why DL is far more accurate in nature. We may discover more resources and opportunities to learn more about the topic as time goes on because the field is always evolving.

References:

1. Rao, Kschandra & Srinivas, Y & Krishna, M. (2020). Study on image processing using deep learning techniques. Materials Today: Proceedings. 10.1016/j.matpr.2020.09.536.
2. I. Goodfellow, Y. Bengio, A. Courville, Deep Learning, MIT Press, 2016.
3. A. Krizhevsky, I. Sutskever, G. Hinton, ImageNet classification with deep convolutional neural networks, NIPS (2012).
4. J. Deng, A. Berg, S. Satheesh, H. Su, A. Khosla, and L. FeiFei. ImageNet Large Scale Visual Recognition Competition 2012 (ILSVRC2012).
5. S. Ren, K. He, R. Girshick, J. Sun, Faster r-cnn: Towards real-time object detection with region proposal networks, NIPS (2015).
6. M. A. Abbas, "Improving deep learning performance using random forest HTM cortical learning algorithm," 2018 First International Workshop on Deep and Representation Learning (IWDRL), Cairo, 2018, pp. 13-18.
7. J. Wu, Yinan Yu, Chang Huang and Kai Yu, "Deep multiple instances learning for image classification and auto-annotation," 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, 2015, pp. 3460-3469.

8. L. Shao, D. Wu, X. Li, Learning Deep and Wide: A Spectral Method for Learning Deep Networks, IEEE Transactions on Neural Networks and Learning Systems 25 (12) (Dec. 2014) 2303–2308.
9. K. Nguyen, C. Fookes and S. Sridharan, “Improving deep convolutional neural networks with unsupervised feature learning,” 2015 IEEE International Conference on Image Processing (ICIP), Quebec City, QC, 2015, pp. 2270-2274.
10. Y. Wang, “Deep reasoning and thinking beyond deep learning by cognitive robots and brain-inspired systems,” 2016 IEEE 15th International Conference on Cognitive Informatics & Cognitive Computing (ICCI*CC), Palo Alto, CA, 2016, pp. 3-3.
11. B. Akay, O. Kaynar and F. Demirkoparan, “Deep learning based recommender systems,” 2017 International Conference on Computer Science and Engineering (UBMK), Antalya, 2017, pp. 645-648.
12. Hemeida, Ashraf & Hassan, Moatamad. (2022). “Image classification based deep learning: A Review.” Aswan University Journal of Sciences and Technology. 2. 10.21608/aujst.2022.259887.
13. J. Valente, J. António, C. Mora, S. Jardim, “Developments in Image Processing Using Deep Learning and Reinforcement Learning,”. Journal Imaging-2023.
14. RK. Santoshachandra, Y. Srinivas, M. Vamshi Krishna, “Study on image processing using deep learning techniques,” Materials Today: Proceedings-2021, Elsevier.
15. Sakshi Indolia; Anil Kumar Goswami; S. P. Mishra; Pooja Asopa. “Conceptual Understanding of Convolutional Neural Network- A Deep Learning Approach.” ScienceDirect-2018, Elsevier.
16. V. O. Owino, “Deep Learning - Convolutional Neural Networks for Image Classification,” no. June, pp. 1–22, 2020, [Online]. Available: <http://john-cd.com/blog/2017/03/08/Deep-Learning>.
17. D. Chitra, D. Kritika,” A Review on Deep Learning Applications,” IGI Global-2020. Page 21-31.
18. Z. Pala, V. Yamli and I. H. Ünlük, “Deep Learning researches in Turkey: An academic approach,” 2017 XIIIth International Conference on Perspective Technologies and Methods in MEMS Design (MEMSTECH), Lviv, 2017, pp. 120-123.
19. E. Çallı, E. Sogancioglu, B. van Ginneken, K. G. van Leeuwen, and K. Murphy, “Deep learning for chest X-ray analysis: A survey,” Med. Image Anal., vol. 72, p. 102125, 2021, doi: 10.1016/j.media.2021.102125
20. Hassan, Aya & Hemeida, Ashraf & Hassan, Moatamad. (2022). Image classification based deep learning: A Review. Aswan University Journal of Sciences and Technology. 2. 10.21608/aujst.2022.259887.
21. E. Alpaydin, Neural Networks and Deep Learning. 2021
22. F. Sultana, A. Sufian, and P. Dutta, “Advancements in image classification using convolutional neural network,” Proc. - 2018 4th IEEE Int. Conf. Res. Comput. Intell. Commun. Networks, ICRCICN 2018, pp. 122–129, 2018, doi: 10.1109/ICRCICN.2018.8718718.
23. Hassan, Aya & Hemeida, Ashraf & Hassan, Moatamad. (2022). Image classification based deep learning: A Review. Aswan University Journal of Sciences and Technology. 2. 10.21608/aujst.2022.259887.
24. ANDRIY BOROVKOV, “Image Classification with Deep Learning Thesis for degree in Master of Science in physics,” Univ. Hambg., p. 55, 2017.
25. Santosh, Kc & Das, Nibaran & Ghosh, Swarnendu. (2022). Deep learning models. 10.1016/B978-0-12-823504-1.00013-1.
26. Papas, Ioannis & Derakhshan, Behrouz & Mahdiraji, Alireza & Markl, Volker. (2021). Continuous Training and Deployment of Deep Learning Models. Datenbank-Spektrum. 21. 10.1007/s13222-021-00386-8.
27. Ketkar, Nikhil. (2017). Training Deep Learning Models. 10.1007/978-1-4842-2766-4_14.
28. Lv, Qing & Zhang, Suzhen & Wang, Yuechun. (2022). Deep Learning Model of Image Classification Using Machine Learning. Advances in Multimedia. 2022. 10.1155/2022/3351256.
29. K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 770-778, doi: 10.1109/CVPR.2016.90