Image Processing Using Deep Learning Techniques

Nisha Shrivastava[1], Dr.P. K. Rai[2]
Department of Computer Application, APS University, Rewa (MP)

Abstract: The original purpose of digital image processing was to make it easier for a human observer to catch specific events happening through such images, yet we tend to conceal the observer. In the same way as a human observer could, we wish to ask the computer to automatically analyze photographs. In image processing, methods based on mathematics are used to process images. In image processing, many approaches are used to the image to produce a better image. The fundamental objective of image enhancement is to process a certain image so that the finished product is better suitable for a particular application than the original image. The first section of the paper is an overview of digital image processing, and the second introduces the idea of Deep Learning (DL) methods and contrasts them.

Keywords: Digital image processing, Deep Learning (DL)

1. INTRODUCTION

Image processing using Deep Learning techniques has revolutionized the field of computer vision by enabling computers to understand and interpret visual information with unprecedented accuracy and efficiency[1]. Deep Learning, a subset of machine learning, utilizes artificial neural networks with multiple layers to learn hierarchical representations of data[2]. When applied to image processing, deep learning algorithms can automatically extract meaningful features and patterns from images, making it possible to perform a wide range of tasks, including image classification, object detection, segmentation, and image synthesis. Figure 1 shows the steps of image processing.

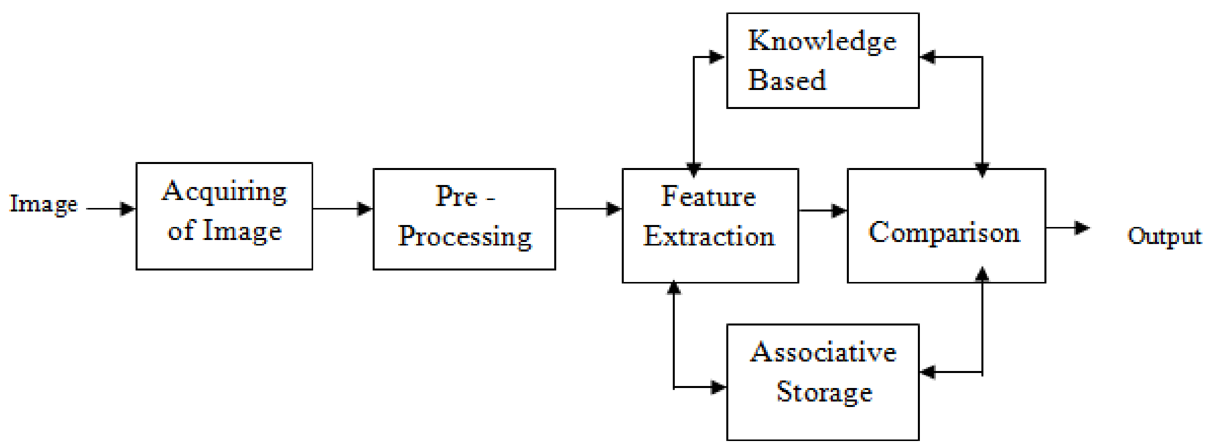

Figure1: Steps of image processing

One of the key advantages of Deep Learning in image processing is its ability to learn directly from raw pixel data, eliminating the need for manual feature engineering. This allows the models to discover complex relationships and capture intricate details that were previously challenging to extract using traditional image processing techniques. Convolutional Neural Networks (CNNs) are particularly well-suited for image processing tasks, as they exploit spatial hierarchies and local receptive fields to effectively capture local patterns and global structures[3]. To train a deep learning model for image processing, a large labeled dataset is typically required. This dataset is used to optimize the model's parameters through a process called backpropagation, where the errors made by the model are propagated backward through the network to adjust the weights and biases[4]. Deep learning models can be trained on powerful hardware such as Graphics Processing Units (GPUs) or specialized accelerators to handle the computational demands of processing large amounts of image data.

Once trained, Deep Learning models can be applied to a variety of image processing tasks. For instance, in image classification, a model can accurately categorize images into predefined classes or categories. Object detection algorithms can identify and localize multiple objects within an image, while segmentation techniques can separate an image into distinct regions or segments based on semantic information. Deep Learning has also been used to generate realistic.

2. HISTORICAL EVOLUTION OF DIGITAL IMAGE PROCESSING

Earlier, the newspaper industry mostly used digital image processing to improve the appearance of photos or to simply turn black and white images into color images. Digital images were electronically sent between London and New York in the 1920s. The first Bartlane cable picture systems could code an image using five grey levels; in 1929, this capability was increased to 15 grey levels. After the development of digital computers and related technologies, such as picture storage, display, and transmission, actual digital image processing began. Powerful computers created useful digital image processing in the 1960s. The Ranger 7 American spacecraft captured satellite images of the moon, which were processed at the Californian Jet Propulsion Laboratory.

At the same time, numerous tasks related to astronomy, medical image processing, remote sensing, etc. started using digital image processing. The use of digital image processing techniques has increased dramatically since 1960. Nowadays, these methods are applied in practically all facets of our lives. They have uses in the domains of law, astronomy, medicine, and defense, among others[5].

Digital image processing applications include recognizing fingerprints, processing satellite images, predicting the weather, character and face identification, product inspection, and assembly.

3. IMAGE PROCESSING

Image processing refers to the manipulation, analysis, and enhancement of digital images using various algorithms and techniques. It involves applying mathematical operations and transformations to images to extract useful information, improve visual quality, or facilitate subsequent analysis. Digital images are composed of pixels, which are individual picture elements that contain information about color and intensity. Image processing techniques operate on these pixels or groups of pixels to achieve desired effects. Some common tasks in image processing include:

i) Image Filtering: Filtering operations involve applying mathematical operations to pixels or neighborhoods of pixels to modify image properties. Common filters include blurring (to reduce noise or smooth images), sharpening (to enhance edges), and edge detection (to identify boundaries between objects) [6].
ii) Image Enhancement: Enhancement techniques aim to improve the visual quality or highlight specific features of an image. These methods may involve adjusting brightness, contrast, or color balance to make the image more visually appealing or improve its interpretability.
iii) Image Restoration: Restoration techniques are used to recover degraded or corrupted images. These methods can remove noise, artifacts, or other distortions caused by factors such as sensor limitations, compression, or transmission errors.
iv) Image Compression: Compression algorithms reduce the storage space required for images by removing redundant or irrelevant information. Lossless compression methods preserve all image details, while lossy compression methods discard some information to achieve higher compression ratios [7].
v) Image Segmentation: Segmentation techniques divide an image into meaningful regions or objects based on properties such as color, texture, or intensity. It is useful for object recognition, image understanding, and computer vision applications.
vi) Object Detection and Recognition: These techniques involve identifying and classifying specific objects or patterns within an image.

Object detection algorithms localize and identify instances of objects, while recognition algorithms associate objects with specific classes or categories [8].

vii) Image Registration: Registration aligns multiple images or different modalities of the same scene to facilitate comparison or fusion of information. It is useful in medical imaging, remote sensing, and panorama stitching.

Image processing plays a crucial role in numerous fields, including medical imaging, surveillance, robotics, remote sensing, digital photography, and computer vision [9]. It enables the extraction of meaningful information from images, aids in decision-making, and helps automate tasks that would otherwise require manual intervention [10].

4. DEEP LEARNING TECHNIQUES

Deep learning techniques are a subset of machine learning methods that use artificial neural networks with multiple layers to learn and extract complex representations from data. These techniques have gained significant attention and achieved remarkable success in various domains, including computer vision, natural language processing, speech recognition, and more [11,12]. Here are some key deep learning techniques:

a) Artificial Neural Networks (ANNs): Artificial Neural Networks form the foundation of deep learning. ANNs are composed of interconnected nodes, or "neurons," organized into layers. Each neuron applies a weighted sum of inputs, followed by a non-linear activation function, to produce an output. ANNs can have multiple layers, including input, hidden, and output layers[13].
b) Convolutional Neural Networks (CNNs): CNNs are specialized neural networks designed for processing grid-like data, such as images [14]. They consist of convolutional layers that apply filters to extract spatial hierarchies of features, pooling layers that downs ample the spatial dimensions, and fully connected layers that perform classification or regression tasks. CNNs excel at tasks like image classification, object detection, and image segmentation [3].
c) Recurrent Neural Networks (RNNs): RNNs are neural networks specifically designed for sequence data processing[15]. They possess recurrent connections that allow them to store and utilize information from previous time steps. RNNs are particularly useful for tasks like speech recognition, natural language processing, and time series analysis[16, 17.]
d) Generative Adversarial Networks (GANs): A generator network and a discriminator network are the two neural networks that make up a GAN. The discriminator network looks for differences between actual and fake data, while the generator network creates synthetic data samples.
e) Transfer Learning: Transfer learning involves using a pre-trained deep learning model that has been trained on a large dataset and then applying it to a different but related task or domain. By leveraging the learned representations from the pre-trained model, transfer learning can significantly reduce the amount of data and training time required for a new task.

In many circumstances, machine intelligence can be just as beneficial as human specialists, if not more so [5]. For this reason, deep learning may be able to address the following issues:

i) Cases where human experts are not available.
ii) Cases where humans are unable to explain decisions made using their expertise (language understanding, medical decisions, and speech recognition) [18].
iii) Cases where the problem solution updates over time (price prediction, stock preference, weather prediction, and tracking).
iv) Cases where solutions require adaptation based on specific cases (personalization, biometrics).
v) Cases where size of the problem is extremely large and exceeds our inadequate reasoning abilities (sentiment analysis, matching ads to Face book, calculation webpage ranks).

Numerous performance characteristics could provide the requirement of deep learning, for example:

i) The Universal Learning Approach DL is frequently called universal learning because of its capacity to function in nearly all application domains.

ii) DL approaches do not necessitate precisely constructed features. Rather, an automated process is used to learn the optimized properties. Deep learning's performance on the task at hand in comparison to human performance. As a result, the input data becomes robust against typical changes [19].

iii) The same deep learning (DL) technology can be used to new data types or applications. This technique is sometimes called transfer learning (TL). It is helpful strategy in situations where there is not enough data.

5. CONCLUSION

The application of deep learning techniques in image processing has ushered in a transformative era, revolutionizing the way we analyze and interpret visual data. The remarkable capabilities of deep neural networks, particularly Convolutional Neural Networks (CNNs), have significantly enhanced the accuracy and efficiency of tasks such as image classification, object detection, segmentation, and image generation[21].One of the key strengths of deep learning in image processing lies in its ability to automatically learn hierarchical features from raw data, eliminating the need for manual feature engineering.

In the future, as technology continues to advance, we can anticipate even more sophisticated deep learning models that push the boundaries of image processing capabilities. The integration of multi-modal data, the development of explainable AI, and the exploration of novel architectures are likely to shape the next phase of advancements in this dynamic and rapidly evolving field [22].

REFERENCE

[1] Al-Azzawi A, Ouadou A, Max H, Duan Y, Tanner JJ, Cheng J.. BMC Bioinform. 2020;21(1):1–38.

[2] Glorot X, Bengio Y. In: Proceedings of the thirteenth international conference on artificial intelligence and statistics; 2010. p. 249–56.

[3] Jhong SY, Tseng PY, Siriphockpirom N, Hsia CH, Huang MS, Hua KL, Chen YY. In: 2020 international conference on advanced robotics and intelligent systems (ARIS). IEEE; 2020. p. 1–6.

[4] Goller C, Kuchler A.In: Proceedings of international conference on neural networks (ICNN'96), vol 1. IEEE; 1996. p. 347–52.

[5] De Fauw J, Ledsam JR, Romera-Paredes B, Nikolov S, Tomasev N, Blackwell S, Askham H, Glorot X, O'Donoghue B, Visentin D,. Nat Med. 2018;24(9):1342–50. 51.

[6] Li HC, Deng ZY, Chiang HH. Sensors. 2020;20(21):6114.

[7] Zhou DX. Theory of deep convolutional neural networks: down sampling. Neural Netw. 2020;124:319–27.

[8] Gu J, Wang Z, Kuen J, Ma L, Shahroudy A, Shuai B, Liu T, Wang X, Wang G, Cai J,. 2018;77:354–77.

[9] Kermany DS, Goldbaum M, Cai W, Valentim CC, Liang H, Baxter SL, McKeown A, Yang G, Wu X, Yan F, 2018;172(5):1122–31.

[10] Mnih V, Kavukcuoglu K, Silver D, Rusu AA, Veness J, Bellemare MG, Graves A, Riedmiller M, Fidjeland AK, Ostrovski G,. 2015;518(7540):529–33.

[11] BaturDinler Ö, Aydin N. Appl Sci. 2020;10(4):1273.

[12] Jagannatha AN, Yu H. In: Proceed-ings of the conference on empirical methods in natural language processing. Conference on empirical methods in natural language processing, vol. 2016, NIH Public Access; 2016. p. 856.

[13] Van Essen B, Kim H, Pearce R, Boakye K, Chen B. Lbann:. In: Proceedings of the workshop on machine learning in high-performance computing environments; 2015. p. 1–6.

[14] Li G, Zhang M, Li J, Lv F, Tong G. Pattern Recogn. 2021;109:107610.

[15] Urban G, Subrahmanya N, Baldi P. J ChemInf Model. 2018;58(2):207–11.

[16] Jiang Y, Kim H, Asnani H, Kannan S, Oh S, Viswanath P. IEEE J Sel Areas Inf Theory. 2020;1(1):207–16.

[17] Sadr H, Pedram MM, Teshnehlab M. Neural Process Lett. 2019;50(3):2745–61.

[18] Pascanu R, Gulcehre C, Cho K, Bengio Y. In: Proceedings of the second international conference on learning representations (ICLR 2014); 2014.

[19] Palaz D, Magimai-Doss M, Collobert R. Speech Commun. 2019;108:15–32.
[20] Saeed MM, Al Aghbari Z, Alsharidah M.. PeerJComput Sci. 2020;6:321.
[21] Socher R, Perelygin A, Wu J, Chuang J, Manning CD, Ng AY, Potts C. In: Proceedings of the 2013 conference on empirical methods in natural language processing; 2013. p. 1631–42.
[22] Socher R, Lin CCY, Ng AY, Manning CD. In: ICML; 2011.