Neural Networks in Image Recognition

Oluchi Emmanuella Obinna

ITAI-1370: Artificial Intelligence History

Prof. Machuria Johnson

September 23, 2023

Abstract

Neural networks are computational models used in image identification that draw inspiration from the visual processing skills of the human brain. By successfully learning and extracting complex features from images, these networks—in particular convolutional neural networks [CNNs]—have transformed the industry. CNNs are capable of capturing both low-level details like edges and textures and high-level, complicated features like object forms and patterns through a sequence of layers that conduct operations like convolution and pooling. These networks can accurately categorize and recognize objects within images after being trained on large datasets, enabling applications like facial recognition, object identification, medical image analysis, and many more. In reality, neural networks, a method of artificial intelligence that has been in and out of favor for more than 70 years, are what deep learning is now known as. Warren McCullough and Walter Pitts, two University of Chicago scholars who transferred to MIT in 1952 as founding members of what is commonly referred to as the first cognitive science department, made the first theoretical proposal for neural networks in 1944 (Hardesty, 2017). Neural networks are an essential tool for automating and enhancing image-related operations in a variety of sectors due to their adaptability and scalability. This essay will shed light on CNNs, their uses, advantages, and disadvantages.

Neural Networks in Image Recognition

Over the past ten years, neural networks have transformed image recognition and computer vision, enabling machines to recognize and interpret visual data with astounding accuracy. A network of artificial neurons organized into layers using software is referred to as a deep neural network. Object identification, facial recognition, medical image analysis, and even self-driving cars have all benefited greatly from the use of neural networks for image recognition. Deeper architectures, regularization methods, and advancements in data augmentation and transfer learning all contribute to their continued progress. As a result,

neural networks have revolutionized the way we process and interact with visual information, providing previously unheard-of automation and accuracy in image identification tasks. The artificial neuron's firing rate is influenced by the weight of each connection between layers, each of which is connected to the one before it. In general, neurons that fire together have stronger connections, and this firing helps define how strong the connections are between layers. similar to how actual neurons work. Modern image recognition systems are built around these artificial neural networks, which were inspired by the structure and operation of the human brain (Lee, 2020).

What is the CNN

Convolutional neural networks [CNNs] are at the core of neural networks for image recognition. CNNs are made to handle and extract features from images in a structured way. They scan and learn local patterns and features, such as edges, textures, and forms, from various areas of an image using convolutional layers. This is a fundamental ConvNet building block that handles the majority of the computing work. To maintain the topology of the image, weight sharing is applied in this layer (Huynh, 2022). In order to minimize the spatial dimensions of the feature maps and increase the network's computational efficiency, pooling layers are frequently used. This is a layer that is frequently placed between two Conv layers. To lower resolution, this layer calculates a mean or maximum over narrow windows. As a result, it lessens the amount of processing and parameters in the network (Huynh, 2022)The network gradually learns to recognize increasingly complicated and abstract properties as the data moves through successive levels. Fully connected layers often make up the last layers of a CNN, which use the learnt characteristics to produce predictions or categorize data. Giving the input image a label or category is a common step in image recognition. In order to reduce prediction errors, the network is trained on a sizable dataset of labeled images and its internal parameters are adjusted using a technique known as backpropagation and gradient descent.

Gradient Descent

Gradient descent is used in backpropagation to ensure that the error rate for each connection is lowered as near to zero and to assess the rate-of-change of the loss function with regard to the weighting of each connection. The network is ought to eventually settle on a solution with the least amount of overall error (Lee, 2020). A key optimization method for convolutional neural networks [CNNs] training is gradient descent. CNNs modify their internal parameters, or weights, during the training process to reduce the difference between the expected and actual values [such as when classifying images]. Gradient descent determines the direction and size of modifications required for the loss function's gradient to be as small as possible with regard to these parameters. CNNs eventually converge towards an ideal set of parameters that yield precise predictions by iteratively updating the weights in the direction opposite to the gradient. This method is refined by variations like stochastic gradient descent [SGD] and adaptive optimizers, which enable CNNs to effectively understand the features and patterns present in picture data during training.

Applications and Limitations

In many different real-world sectors, neural networks in image identification have found useful and diverse uses. Due to their capacity to automatically recognize and extract characteristics from images, CNNs have a remarkable range of applications. They are used in a wide range of industries, including security for facial recognition, e-commerce for visual search and recommendation, healthcare for diagnosing diseases from medical images, autonomous vehicles for safe navigation, entertainment for content recommendation and augmented reality, and even in environmental monitoring for analyzing satellite imagery. Because of its adaptability and ability to absorb and understand visual data, CNNs are a foundational technology with the potential to revolutionize a wide range of industries.

Applications: Healthcare

They are used in the healthcare industry to help radiologists identify and diagnose disorders from medical images. To find anomalies like cancers, fractures, or irregularities in a patient's anatomy, CNNs can, for example, examine X-rays and MRI scans. The trained CNN can interpret new photographs by spotting patterns in the individual images that point to specific diseases. It mimics medical training in this sense, but the hypothesis holds that the CNN approach yields more accurate results since it can learn from a much larger number of images than any human. (Hou et al., Gao). By enabling early detection, this technology expedites diagnosis, improves accuracy, and may even save lives.

Applications: Autonomous Decision Making

Neural networks are essential to the development of self-driving vehicles in the automotive sector. Autonomous vehicles' cameras and sensors record live photos of the road and surroundings. The autonomous decision-making mechanisms in self-driving cars. They are able to handle data streams from a variety of sensors, including cameras, LiDAR, RADAR, GPS, and inertia sensors. Deep learning algorithms are then used to model this data, and as a result, decisions are made that are appropriate for the driving situation (Barla, 2023). Through the processing of these images, CNNs can identify lane lines, pedestrians, other vehicles, and traffic signs, enabling the vehicle to safely travel. Through the reduction of accidents and the provision of more effective and practical commuting, this application has the potential to revolutionize transportation.

Applications: Facial Recognition

Another well-known instance is facial recognition technology used in security and surveillance systems, which is powered by neural networks. In order to build an individual record of a person that can be saved in a database, facial recognition gathers many photos via camera or video, parses the images, and attempts to record the data of specific places on a

human face. In particular, facial recognition makes use of convolutional neural networks to identify faces by correctly classifying the coefficients obtained by the eigenface technique. Before being used to match the face pictures that are submitted to it, the network must first be trained on the images from the face database. Up to 80 elements on the face are examined by facial recognition systems to assist in detecting distinctive traits and ultimately match these aspects to identify a person (Adinma Chidumije, Fatima Gowher, Ehsan Kamalinejad, Justin Mercado, Jiwanjot Soni, & Jiaofei Zhong, 2021). This technology is used at airports, public areas, and private businesses to identify people, improve security, and simplify access control.

Limitations

There are various drawbacks to using neural networks for image identification. They are frequently data-hungry and unusable for domains with small datasets since they frequently need huge volumes of labeled data for training. Their deployment on devices with limited resources is limited by their complicated structures, which demand a lot of computational power. Interpretability is still a problem because neural networks are frequently thought of as "black boxes," making it hard to comprehend how they make decisions (Urwin, 2023)They are also vulnerable to adversarial attacks, in which minute changes to an image can produce inaccurate predictions and cause security issues.

Image Recognition: Can we trust it?

The potential of neural network image identification technology is enormous. We may anticipate the emergence of more complex systems, including capsule networks, which will enhance our capacity to comprehend spatial relationships in images. There will be less need for large labeled datasets as transfer learning and

domain adaptation techniques improve and enable models trained on one type of picture data to be used in other domains. Applications like augmented reality will become more widely used as real-time, on-device picture recognition becomes more common. In addition, as research into explainable AI advances, neural networks may become increasingly important in addressing ethical problems by providing more transparent and understandable models. However, moral issues continue to be crucial. Image recognition systems' privacy, surveillance, and data security concerns will continue to be hotly contested topics. To ensure that these mechanisms do not sustain discrimination, especially in touchy areas like criminal justice or hiring, concerns about bias and fairness must be addressed. Finding a balance between technical innovation and defending individual rights and societal values will be a big problem as image recognition becomes more and more ingrained in our daily lives. The use of neural networks for image identification will continue to have a significant social influence. Industries including healthcare, driverless vehicles, entertainment, and agriculture will all be transformed by them. The broad use of facial recognition will prompt concerns about surveillance and personal privacy. Furthermore, to protect the integrity of the visual content, strong countermeasures will be required due to their significance in the development of deep fakes and misinformation. In order to ensure that the advantages are shared fairly and any possible risks are reduced, society must consider the implications and obligations that come with the use of new technologies as they grow more pervasive in our daily lives. The future of neural networks in image recognition will be significantly shaped by public policy, regulation, and ethical standards.

Conclusion

A new era of computer vision has begun thanks to neural networks in image recognition, which allows machines to comprehend and interpret visual input with astounding accuracy. They are excellent at discovering and automatically extracting pertinent features from images, doing away with the need for laborious feature engineering. Due to their hierarchical nature, they are able to effectively recognize complicated patterns and objects by being able to capture both low-level and high-level visual information. In addition to being highly scalable, CNNs can be used for a variety of image-related tasks, including segmentation and object detection. They are also flexible to a variety of domains due to their capacity to generalize from big datasets, which contributes to their widespread use in computer vision applications. That being said, we must be reminded of their limits, Such as their need for sizable annotated datasets for efficient training, which can be costly and unworkable for specialized topics. They also require strong hardware that may not be readily available to all developers because they are computationally taxing. Furthermore, these networks may be hard to read, making it difficult to comprehend why they generate particular predictions, particularly in high-stakes applications like autonomous vehicles or healthcare. Unrepresentative training data can also cause bias and fairness problems, which can produce unfair results and raise moral dilemmas. Ethical alternatives are being actively investigated to address these limits. Increasing the diversity and objectivity of training datasets is one method for ensuring that neural networks learn from a larger variety of perspectives and experiences. The goal of explainable AI is to increase the transparency of neural networks by revealing how they make decisions. In addition, rules and guidelines for the responsible application of image recognition technology are developing, particularly in delicate contexts like

surveillance or law enforcement. In order to reduce bias and prejudice in image recognition systems, it is essential to promote research into robust evaluation techniques and fairness-aware algorithms. In conclusion, neural networks have transformed computer vision through picture recognition, but they have both practical and ethical drawbacks. Continuous efforts in data collecting, transparency, and ethical usage will be crucial in determining the future of this breakthrough technology in order to maximize their potential while protecting against any harm.

References

Larry Hardesty | MIT News Office. (n.d.). *Explained: Neural networks*. MIT News | Massachusetts Institute of Technology.

https://news.mit.edu/2017/explained-neural-networks-deep-learning-0414

Larry Hardesty | MIT News Office. (n.d.). *Explained: Neural networks*. MIT News | Massachusetts Institute of Technology.

https://sandy-lee.medium.com/how-to-train-neural-networks-for-image-classification-part-1-21327fe1cc1

Huynh, N. (2022, April 9). *Convolutional Neural Networks for Image recognition*. Medium.

https://medium.com/mlearning-ai/convolutional-neural-networks-for-image-recognition-7148a19f981f

Hou, J., & Gao, T. (2021, August 9). *Explainable DCNN based chest X-ray image analysis and classification for COVID-19 pneumonia detection*. Nature News.

https://www.nature.com/articles/s41598-021-95680-6

Barla, N. (2023, August 10). *Self-driving cars with Convolutional Neural Networks (CNN)*. neptune.ai.

https://neptune.ai/blog/self-driving-cars-with-convolutional-neural-networks-cnn

A survey of CNN and facial recognition methods in the age of Covid-19*. (n.d.).

https://dl.acm.org/doi/fullHtml/10.1145/3471287.3471292

4 disadvantages of Neural Networks. Built In. (n.d.).

https://builtin.com/data-science/disadvantages-neural-networks