Neural Networks for Image Recognition

Sonal Borase, PhD Scholar, JJT University, Jhunjhunu, India, sonalborase@gmail.com

Javed Shaikh, Associate Professor, SKN Singhgad Institute of Technology & Science, Lonavala, India, javedsheikh1987@gmail.com

Anupama Deshpande, Professor, JJT University, Jhunjhunu, Rajasthan, India

Mangala.d.2000@gmail.com

Abstract:

Pattern recognition involves study from image processing fields. These fields include machine learning i.e. a branch of Artificial Intelligence. Pattern recognition is used in computer-aided diagnosis, recognition of handwritten words, recognition of images, etc. Image recognition is a subset of pattern recognition. It has become prominent in many machine vision applications. This is designed in such a way that it should learn automatically. Spatial hierarchies of features are used in algorithms through backpropagation by using multiple layers. This article offers a perspective on the basic concepts of Image recognition using neural networks and its applications in the image processing domain. Neural networks have been successful in identifying faces, objects in the image. The neural network provides a very exciting alternative and other application that can play an important role in today's computer science field. There are some limitations also which are mentioned.

Keywords *— Artificial Neural Network, ANN, Convolutional Neural Network, CNN, LeNet-5 Structure, AlexNet Structure, VGG-16 Structure, Inception-v*

Introduction:

In terms of machine vision, Image recognition is nothing but an identifying object in the image. Here objects in the image can be a specific action, place, or person. This recognition can be done by software or technology. The technology can be a combination of Artificial intelligence software along with a camera to achieve image recognition. Under the hood, image recognition is powered by deep learning, specifically Convolution Neural Networks, a neural network architecture that emulates how the visual cortex breaks down and analyzes image data. Convolution Neural Network and neural network image recognition is a core component of deep learning for computer vision. It has many applications including e-commerce, gaming, automotive, manufacturing, and education.

I. WHAT IS IMAGE RECOGNITION

Image recognition uses artificial intelligence technology. It identifies objects, people, places, and actions in images automatically. Image recognition is used to perform tasks like searching for content in images, labeling images with descriptive tags. It also performs guiding robots, autonomous vehicles, and driver assistance systems. Image recognition is natural for humans and animals. It is an extremely difficult task for computers to perform. Over the past two decades, tools and technologies have been developed in the field of Computer Vision has emerged, which can rise to the challenge. A deep neural network is the most effective tool found for the task for image recognition. Specifically a Convolution Neural Network. Convolution Neural Network is an architecture designed to understand high-resolution images efficiently. It can correlate and understand a large amount of data in high-resolution images.

II. HOW DOES IMAGE RECOGNITION WORK?

When the human eye sees an image it is interpreted by the brain as a set of signals. That is retained in memory linked to objects and concepts as the outcome is an experience of a scene. Image recognition imitates this process. In the process of image recognition using a neural network, it depicts physical objects and features in the vector encoding of the image. It is turned into constructs that can logically analyze by Computer vision systems. First it simplifies images and extracts the most important information. Then organize the data through feature extraction and classification. Finally, the use of classification algorithms to decide on the image in computer vision systems.

III. IMAGE RECOGNITION ALGORITHMS

Image classifier is one type of image recognition algorithm. It selects an image as an input and predicts the content of the image. The class label a type of class output like a classification of an object in the image whether it's living or non-living thing. For this the algorithm needs to be trained. Training is a learning process. It can distinguish between classes. E.g., for a classification algorithm we have to train a neural network with thousands of images of living and non-living things. The algorithm will learn to extract the features that identify living things object and classify images that contain non-living things too. Here other algorithms can be used to perform more complex activities

IV. IMAGE DATA PRE-PROCESSING STEPS FOR NEURAL NETWORKS

Image recognition using neural network algorithms depends on the quality of the dataset. The images used for training and testing purpose of the model. Here for the image data preparation, important parameters are considered.

- Image size: The main purpose of higher quality image is to give more information to the model. But at the same time it requires more neural network nodes and more computing power to process as well.
- Number of images: For the more accurate result we have to feed more data to a model. It ensures the training set represents the real population.
- Number of channels: Color images typically have 3 color channels i.e. Red, Green, Blue with colors represented in the range [0,255]. Similarly grayscale image have 2 channels i.e. black and white
- Aspect ratio: As the neural network assume a square shape input image. So the size and aspect ratio of the image is supposed to be same.
- Image scaling: Once all images are squared you can scale each image. There are different deep learning library functions such as up-scaling and down-scaling techniques.
- Mean, standard deviation of input data: you can look at the ‘mean image’ by calculating the mean values for each pixel, in all training examples, to obtain information on the underlying structure in the images.
- Normalizing image inputs: ensures that all input parameters (pixels in this case) have a uniform data distribution.
- Dimensionality reduction: you can decide to collapse the RGB channels into a gray-scale channel.
- Data augmentation: involves augmenting the existing data-set, with perturbed types of current images, including scaling and rotating.

V. BUILDING THE PREDICTIVE MODEL FOR IMAGE WITH NEURAL NETWORKS

After the training of images, we need a system that can process them and use them to make a prediction on new, unknown images. That system is an artificial neural network. Image recognition using neural network algorithms can classify anything. It can classify text, images, audio files, and videos as well. Perceptron's are nothing but interconnected collection of nodes of neural network. Every neuron takes one minute sample of the input data i.e. one pixel of the image. Then it applies to an activation function to generate a result. An activation function is a simple computation. Each neuron carries a numerical weight which affects its result. That result is then fed to additional neural layers, unless and until at the end of the process the neural network generates a prediction for each input or pixel.

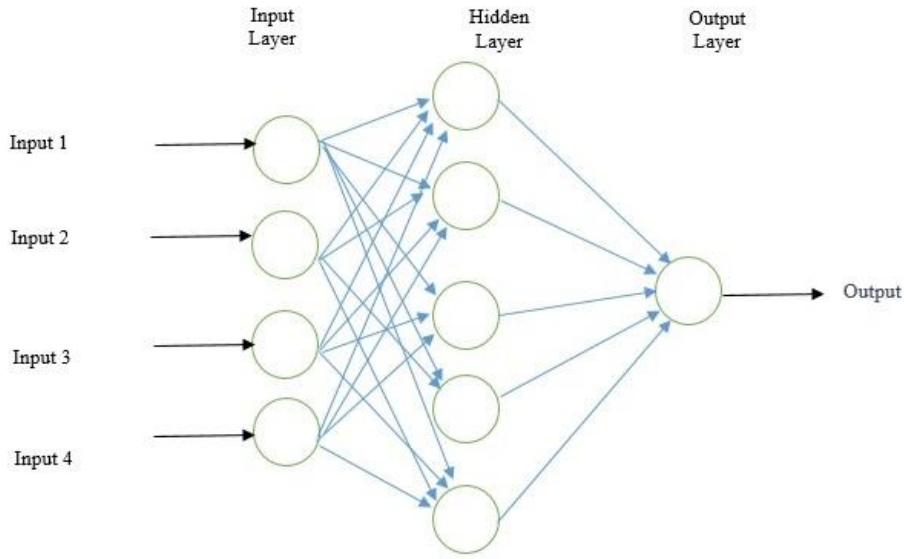

Fig: Perceptron connection

In this process of repetitions, the network learns the most appropriate weights for each neuron for a large number of images, which led to the accurate predictions, in a process called backpropagation. Once a model is trained, it is applied to a new set of images which did not participate in training, to test its accuracy. The model can be used to classify real-world images after some tuning.

VI. LIMITATIONS OF REGULAR NEURAL NETWORKS FOR IMAGE RECOGNITION

Traditional neural networks use a fully-connected architecture. Every neuron in one layer connected to all the neurons in the next layer. For processing the image data, the architecture of the neural network is inefficient.

- Overfitting is another limitation of the traditional neural network, as for an average image with hundreds of pixels and three channels, it will generate millions of parameters.

- The model of the neural network would be very computationally intensive.
- Improve its performance based on the interpretation of results may be difficult.

VII. IMAGE RECOGNITION APPLICATIONS

Implementations of image recognition using neural networks requires security and surveillance. It includes large databases for face recognition, object recognition, medical image analysis, visual geolocation, driver assistance, gesture recognition and image tagging. It also helpful for organization websites. Image recognition has entered the mainstream. In production by Facebook, Google and many other high profile consumer applications are used in for the recognition of face, photo, and video frame. Toolkits and cloud services have emerged. It can help smaller players integrate image recognition into their websites. There are different application areas of neural network image recognition such as Manufacturing, E-commerce industry, Education, Gaming Industry, Automotive Industry.

Image recognition is used to automatically process, categorize and tag product images, and enable powerful image search. For example, consumers can search for a chair with a particular armrest and receive relevant results. Image recognition can be used to transpose a digital layer on top of images from the real world. Augmented reality adds details to the existing environment. Pokémon Go is a popular game that relies on image recognition technology. Autonomous vehicles are in testing phases in the United States and are used for public transport in many European cities. To facilitate autonomous driving, image recognition is taught to identify objects on the road, including moving objects, vehicles, people and pathways, as well as recognize traffic lights and road signs. Image recognition is employed in different stages of the manufacturing cycle. It is used to reduce defects within the manufacturing process, for example, by storing images of components with related metadata and automatically identifying defects. Image recognition can help students with learning difficulties and disabilities. For example, applications powered by computer vision provide image-to-speech and text-to-speech functions, which can read out materials to students with dyslexia or impaired vision.

VIII. CONCLUSION

By studying Image recognition using Neural Network we had concluded that as per as technology is developing day by day the need of Artificial Intelligence is increasing because of only parallel processing. In present time parallel processing is needed because with the help of this only we can save more time and money in any work related to computers and robots. In future, we have to develop more algorithms and problem solving techniques so that we can remove the limitations of the Artificial Neural Network. And if the Image processing concepts combined with the Machine learning we will definitely solve some limitations of this excellent technology.

REFERENCES

[1] Loan N. N. Do Neda Taherifar Hai L. Vu, “Survey of neural network-based models for short-term traffic state prediction”, WILEY, 2018

[2] Gözde ÖZSERT YİĞİT , Buse Melis ÖZYILDIRIM “Comparison of Convolutional Neural Network Models for Food Image Classification”, IEEE 2017

[3] Kasthurirangan Gopalakrishnan, Siddhartha K., Khaitan b, Alok Choudhary a, Ankit Agrawal, “Deep Convolutional Neural Networks with transfer learning for computer vision-based data-driven pavement distress detection”, ELSEVIER 2017

[4] J. F. Velez, A. Sanchez, A. B. Moreno, and L. Morillo-Velarde, “Comparing elastic alignment algorithms for the off-line signature verification problem.” in IWINAC (2), ser. Lecture Notes in Computer Science, J. M. Ferrndez, J. R. l. Snchez, F. de la Paz, and F. J. Toledo, Eds., vol. 6687. Springer, 2011, pp. 233–242.

[5] J. F. V. Bonilla, M. A. F. Ballester, C. M. T. Gonzalez, and J. B. A. Hernandez, “Offline signature verification based on pseudo-cepstral coefficients,” in Proceedings of the 2009 10th International Conference on Document Analysis and Recognition, ser. ICDAR ’09. Washington, DC, USA: IEEE Computer Society, 2009, pp. 126–130.

[6] J.F. Vargas, C.M.Travieso, J.B.Alonso, M.A. Ferrer, "Offline Signature Verification Based on Gray level Information using Wavelet transform and Texture Features", proceedings of 12th International Conference on Frontiers in Handwriting Recognition, IEEE 2010, pp 587-592

[7] J. Hu and Y. Chen, “Offline signature verification using real adaboost clas- sifier combination of pseudo-dynamic features,” in Document Analysis and Recognition (ICDAR), 2013 12th International Conference on, Aug 2013, pp. 1345–1349.

[8] Karen Simonyan, Andrew Zisserman,“Very Deep Convolutional Networks for Large-Scale Image Recognition”, Cornell University April 2015

[9] Wang et al,“Development of convolutional neural network and its application in image classification: a survey”, SPIE 2019

[10] J. U. Duncombe, “Infrared navigation—Part I: An assessment of feasibility (Periodical style),” *IJREAM* vol. ED-11, pp. 34–39, Jan. 1959.

[11] K. Teilo, “ An Introduction to Convolutional Neural Networks,” no. NOVEMBER 2015, pp. 0–11, 2016.

[12] I. Kokkinos, E. C. Paris, and G. Group, “Introduction to Deep Learning Convolutional Networks, Dropout, Maxout 1,” pp. 1–70.