Deep Learning and Computer Vision in Image Processing.

*Akshat Kakkad(38)*  
*Information and communication Technology*  
*Marwadi University*Rajkot, India  
akshat.kakkad107917@marwadiuniversity.ac.in

*Abstract*—Over the past few years Deep Learning and Computer Vision gained a good rise in the field of Image Processing. DL has been applied widely for image analysis task which includes Image Fusion, Image Registration, Scene Classification, Object Detection, Segmentation, and at last Object Based Image Analysis (OBIA). Computer Vision is further studied in many different perspectives like Pattern Recognition, Machine Learning, Computer Graphics

Keywords—Deep Learning, Image Processing, Computer Vision, Image Fusion, Image Registration, Scene Classification, Object Detection, Segmentation, and Machine Learning.

# Introduction

Computer Vision is a very vast field of study which includes different concepts, techniques, and ideas for recording raw data and then performing extraction of pattern and image interpretation.[1] Computer Vision is used to perform tasks that help us to obtain information about some event from digital images. The methods used to analyze any event depends on the nature of data in the event.[2][3] Computer Vision is also gaining importance to perform tasks like creating models and data extracts. Development in this field is done by deep learning of characteristics of Human Vision.[4][5]

Image Processing is to implement computational Transformations for images such as contrast, sharpening etc. This is done under HCI ‘Human Computer Interaction [6].HCI focuses on all aspects of technologies that are related to interaction of Human and Computer. However, Computer Vision is not capable to analyze any image as good as Human Vision (Human eye).[7] One of the main challenge faced in the development is accuracy in results, less sensitivity of algorithms as well as less accuracy in the results obtained.

Deep learning (DL) mainly focuses on algorithmic advancements. DL itself is a learning algorithm based on Neural Networks. These networks are composed of different layers that transform input data(Image) to output (Based on different categories) while learning progressively features of input data(image).[8]

DL can be applicable to several other applications; therefore, we can say that it could be applied to almost every step of remote sensing image processing. Below figure is Taxonomy related to DL in remote sensing image processing.[9]

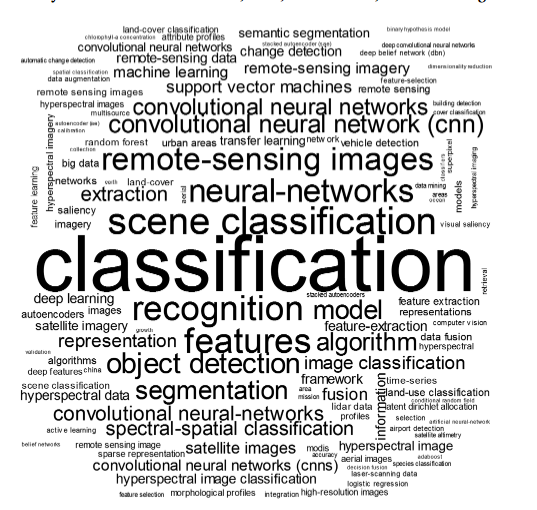


Fig 1. Taxonomy of DL in image processing

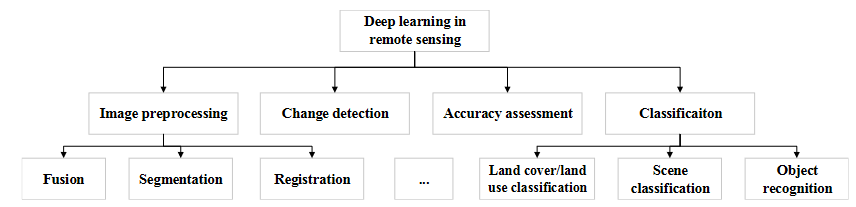


Fig 2. Four Main Tasks of Deep Learning (DL)

# Deep Learning in Image Processing

## Image Fusion

Image Fusion helps us to obtain an image that has both high spectral and spatial resolutions simultaneously. For example, pan-sharpening. It can be regarded as super resolution problem for a low-resolution source image, aided by high resolution source image[10]. There are many methods for DL based image fusion, some most popular are: -

* MS and Pan Image Fusion.
* HS and MS image Fusion.

## Image Registration

In Image Registration we try to align two or more images captured by a sensor at different time or captured by didderent sensors from different viewpoints. It generally includes four steps: -

* Feature Extraction: - Decides what feature is to be matched for analysis.
* Feature Matching: - Matches Features from different images according to given parameters.
* Transformation Model Estimation: - Based on Harmony Search Algorithm[11].
* Image Resampling: - Used to determine similarity between images and resample them.

## Scene Classification and Object Detection.

Scene classification basically helps us in determination of categories from numerous pictures like: -

* Agriculture Scenes
* Forest Scenes
* Beach Scenes etc.

Object Detection mainly focuses on detecting different objects in a particular given scene such as: -

* Villages
* Urban areas
* Aero planes
* Other Automobiles etc.

## Sematic Segmentation

A picture or an image to be analyzed is divided in parts known as pixels. This Sematic Segmentation helps us in assigning land cover label to each pixel of the image. It refers to analyzing dense prediction of pixel labels rather then pixel by pixel analysis. (Pixel by pixel is one of the traditional methods for classification of images).

## Object based Image Analysis.

First it generates patches and then represents the type of object through pixel classification. This all is done with the help of CNNs (Convolutional Neural networks).[12]

# Computer Vision In Image Processing

Computer Vision helps us to extract valuable information from any given image or Object. It does this by using optical sensors and algorithms to stimulate human visualization. Computer vision has been expanded to a branch of Artificial Intelligence (AI) which makes it more stronger as compared to other conventional methods that take huge time to find results and to analyze image.[13]

There are nearly five stages for analysis of image using Computer Vision they are: -

* Image Formation – Capturing of image of our object and storing it in our system (database).
* Image Preprocessing – Quality of the captured image is upgraded or enhanced to clearly view the details in image.
* Image Segmentation – Object is classified on certain given parameters and the separated from its background in the image.
* Image Measurement – Quantization of image takes place based on several characteristics and significant features.
* Image Interpretation – At last deep and detailed interpretation of the filtered image takes place and then we get our result / output.

# Uses of Image Processing using Computer Vision.

Here are some of the uses of image processing in which study is going on: -

* Image Processing is applied in Digital Sound System. [14]
* Image Processing is used in food analysis. (Has a very good impact on food and beverages industry).[15][16]
* Convolutional Neural Networks (CNN) is used for object detection. (To solve edge detection problems) [17]

# Pattern Recognition

It is a process which extracts information form the images obtained from sensors and based on it performs analysis and take necessary steps / decisions. It deals with object identification through image transformation to get better quality of interpretation of image[18]. Computer Vision framework consist of two groups: -

* 3D Morphological Analysis – standard theory for computer vision image processing and pattern recognition
* Pixel Optimization – related to pixel morphology for better structural analysis and characterization as well as better understanding of vector functions of image.

# Segmentation and Retrival of image data

Segmentation means dividing image into small areas and such that they do not overlap with the help of certain algorithms. Then grouping is performed on them. (A city is a collection of these pixels that have nearly same characteristics as gray level, texture, and color).

Fundamentally Segmentation consists of Five main stages: -

* Input image
* Segmented map before integration
* Edge map before integration
* Segmented map and edge map after combination
* Pixel clustering

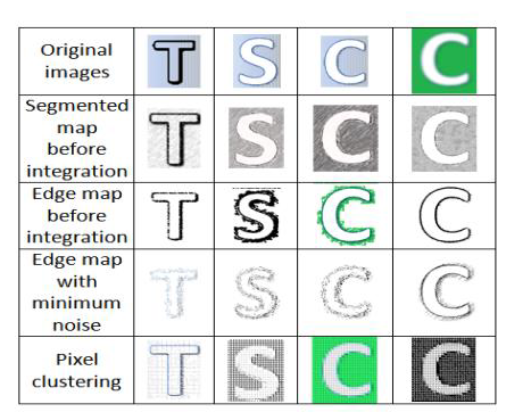


Fig 3. Segmentation stages



Fig 4. Example of image retrievals using query image from big datasets of PASCAL MTH, MSD, SLAR, CDH, and RADAR with various RGB pixel values.

Object Detection often has some errors as there is no perfect algorithm which specifies a particular pattern of any given object. Example – Here in the above figure of face recognition algorithm is applied to find supplements like head, nose, eyes, ears etc. But as we are aware no face is similar so sometimes these algorithms fail in performing their task properly.

Image segmentation approach that is often used is Intensity, color approach and shape approach. In computer vision and image processing segmentation and edge or border detection is very useful for proper interpretation of the image. Segmentation of any given image is done with the help of feature space and gradient texture.

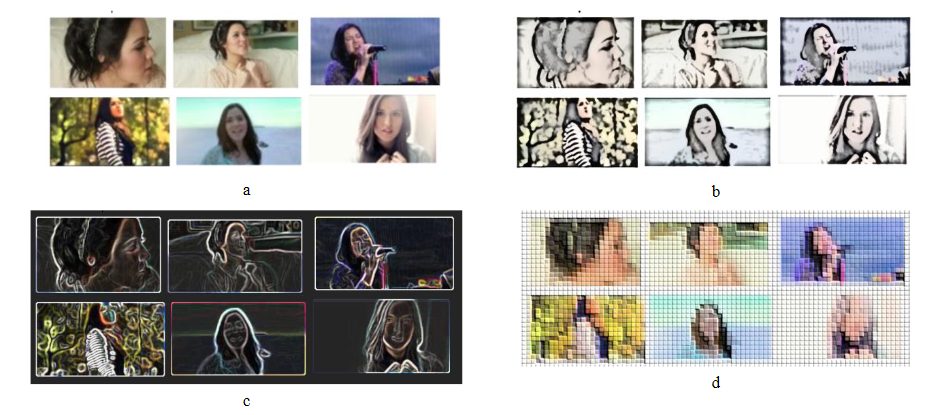


Fig 5. Example of Image processing (sample extraction using segmentation)

1. After scaling objects have different size as compared to their original size.
2. Scaling and arrangement of objects after BW color.
3. Object is changed to black and white.
4. Analysis and matching of colors. (Bitmap)

This method is based on the theory that intensity of the pixels of image will change rapidly.

# Bitmap Image

Bitmap images are referred as raster images. They are stored in our system as a set of pixels. They are very small in size and get blurred on enlargement. These images are used to determine color depth and resolution of a particular image. (As each pixel in a Bitmap image consist of only one color). Sometimes due to unstructured pixel correct intensity of a real seen is not obtained in a bitmap image. But bitmap images are read by computers, so they are important for us to perform analysis. During transmission and reception of these data and images, noise enters in our image. Or when an image is scanned using scanner it may generate some noise (Distortion). Several algorithms are used to reduce this type of noises like gaussian based algorithm, linear filtering, spam filtering, optical character recognition etc.

##### Conclusion

Several main subfields of deep learning are summarized, including Image Fusion, Image Registration, scene classification, object detection segmentation, OBIA. DL models which are supervised need a large quantity of training samples (E.g. - For CNN), For image processing remote sensing, DL models including SAEs and CNNs have been considerably successful recently in remote sensing image fusion, and therefore, more DL models such as RNNs and GANs are expected to be introduced into this field for further developments.

Machine Learning and Image Processing are related to computer vision. This has brought efficiency and enhancement in many of the areas like analysis of images to obtain information and then taking necessary steps or decisions based on it. This is used more extensively in areas of studies and detect object behavior and characteristics including natural events as well as human activities. It can further be applied to many fields like satellite communication, health care, Robotics, geographical remote sensing etc.

##### Acknowledgment

This paper is written under the guidance of prof. Vishal Sorathiya under the subject of technical writing work was supported by Information and communication technology department of Marwadi University Rajkot.

##### References

[1] G. Scarpa, S. Vitale, and D. Cozzolino, “Target-Adaptive CNN-based pansharpening,” *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 9, pp. 5443–5457, 2018, doi: 10.1109/TGRS.2018.2817393.

[2] V. Wiley and T. Lucas, “Computer Vision and Image Processing: A Paper Review,” *Int. J. Artif. Intell. Res.*, vol. 2, no. 1, p. 22, 2018, doi: 10.29099/ijair.v2i1.42.

[3] O. Cosido *et al.*, “Hybridization of convergent photogrammetry, computer vision, and artificial intelligence for digital documentation of cultural heritage-A case study: The magdalena palace,” *Proc. - 2014 Int. Conf. Cyberworlds, CW 2014*, no. October 2014, pp. 369–376, 2014, doi: 10.1109/CW.2014.58.

[4] B. Oluleye, A. Leisa, D. Dean, and L. Jinsong, “A Neuronal Classification System for Plant Leaves Using Genetic Image Segmentation,” *Br. J. Math. Comput. Sci.*, vol. 9, no. 3, pp. 261–278, 2015, doi: 10.9734/bjmcs/2015/14611.

[5] K. K. Patel, A. Kar, S. N. Jha, and M. A. Khan, “Machine vision system: A tool for quality inspection of food and agricultural products,” *J. Food Sci. Technol.*, vol. 49, no. 2, pp. 123–141, 2012, doi: 10.1007/s13197-011-0321-4.

[6] S. S. Rautaray and A. Agrawal, “Vision based hand gesture recognition for human computer interaction: a survey,” *Artif. Intell. Rev.*, vol. 43, no. 1, 2015, doi: 10.1007/s10462-012-9356-9.

[7] F. Zhao, X. Xie, and M. Roach, “Computer Vision Techniques for Transcatheter Intervention,” *IEEE J. Transl. Eng. Heal. Med.*, vol. 3, no. April, 2015, doi: 10.1109/JTEHM.2015.2446988.

[8] H. He, M. Chen, T. Chen, and D. Li, “Matching of remote sensing images with complex background variations via Siamese convolutional neural network,” *Remote Sens.*, vol. 10, no. 2, pp. 1–23, 2018, doi: 10.3390/rs10020355.

[9] L. Ma, Y. Liu, X. Zhang, Y. Ye, G. Yin, and B. A. Johnson, “Deep learning in remote sensing applications: A meta-analysis and review,” *ISPRS J. Photogramm. Remote Sens.*, vol. 152, no. April, pp. 166–177, 2019, doi: 10.1016/j.isprsjprs.2019.04.015.

[10] R. Kemker, C. Salvaggio, and C. Kanan, “Algorithms for semantic segmentation of multispectral remote sensing imagery using deep learning,” *ISPRS J. Photogramm. Remote Sens.*, vol. 145, no. March, pp. 60–77, 2018, doi: 10.1016/j.isprsjprs.2018.04.014.

[11] E. Cuevas and M. Díaz, “A method for estimating view transformations from image correspondences based on the harmony search algorithm,” *Comput. Intell. Neurosci.*, vol. 2015, 2015, doi: 10.1155/2015/434263.

[12] V. Antoniou *et al.*, “Investigating the feasibility of geo-Tagged photographs as sources of land cover input data,” *ISPRS Int. J. Geo-Information*, vol. 5, no. 5, 2016, doi: 10.3390/ijgi5050064.

[13] S. Matiacevich, D. Celis Cofré, P. Silva, J. Enrione, and F. Osorio, “Quality parameters of six cultivars of blueberry using computer vision,” *Int. J. Food Sci.*, vol. 2013, no. December 2015, 2013, doi: 10.1155/2013/419535.

[14] A. T. Zai, S. Bhargava, N. Mesgarani, and S. C. Liu, “Reconstruction of audio waveforms from spike trains of artificial cochlea models,” *Front. Neurosci.*, vol. 9, no. OCT, pp. 1–13, 2015, doi: 10.3389/fnins.2015.00347.

[15] O. W. Liew, P. C. J. Chong, B. Li, and A. K. Asundi, *Signature optical cues: Emerging technologies for monitoring plant health*, vol. 8, no. 5. 2008.

[16] M. Usman and H. Chen, “Recent Trends in Food Intake Monitoring using Wearable Sensors,” pp. 1–27, 2021, [Online]. Available: http://arxiv.org/abs/2101.01378.

[17] H. Wu, Y. Zhou, Q. Luo, and M. A. Basset, “Training feedforward neural networks using symbiotic organisms search algorithm,” *Comput. Intell. Neurosci.*, vol. 2016, no. January, 2016, doi: 10.1155/2016/9063065.

[18] V. Nandakumar *et al.*, “Vorinostat differentially alters 3D nuclear structure of cancer and non-cancerous esophageal cells,” *Sci. Rep.*, vol. 6, 2016, doi: 10.1038/srep30593.