Deep Learning in Visualization

Abstract:

Computer visualization has seen tremendous breakthroughs as a result of deep learning. This work focuses on leveraging on latest achievements in the field of computer visualization using deep learning technology. Besides popular techniques used in image processing like Convolution Neural Networks (CNN), state-of-the-art techniques like SIFT, SURF and KAZE algorithms are discussed. Further, phenomenal achievements using deeplearning techniques including autonomous cars, Virtual reality, gaming, etc. [1] are explained. Finally, an in-depth analysis is done to understand the working methodology of Image segmentation using the popular CNN deep learning systems; FCN and Mask R-CNN [2].

1.Introduction to Deep Learning

Deep learning has taken the world technology by wave. It is an advanced technology by which computers/machines learn by itself; focuses on using multiple-layered artificial neural networks to extract features out of raw data, in contrast to typical machine learning techniques that rely on manually created features (fig 2).[1,3] As such, it is possible for computers to learn intricate patterns from visualizations, especially 2D images and videos without input from human beings.

Typical layers in deep learning are 7 (based on the idea that human vision has 7 levels), but there are currently 1001 layers. A global view of the system is shown below (Fig 1) with a simple comparison to typical machine learning (Fig.2) [3] which requires human intervention for feature extractions.

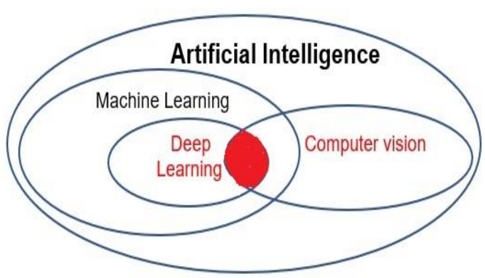


Fig.1

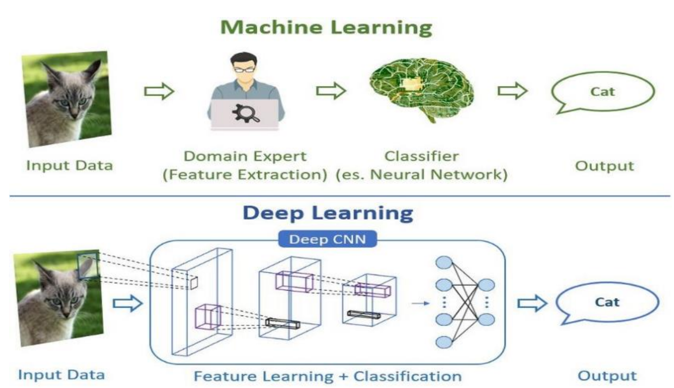


Fig. 2

Computer Vision (CV) allows computers to obtain information from digital images and videos like human vision thereby automating the tasks human visual systems do. Deep learning algorithms have revolutionised the field of computer vision by offering novel and effective answers to a variety of image-related issues that had long been unsolved or only partially addressed.

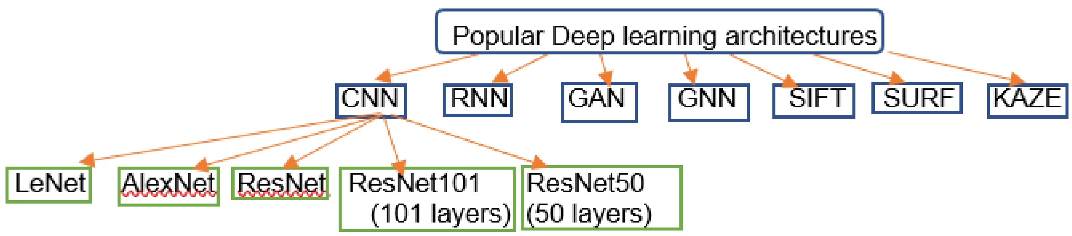
2. Achievements in Computer Visualisation

Deep learning techniques is utilized in a variety of computer visualization applications with impressive outcomes, including computer vision, virtual reality etc. [1,4]

2.1 State of the art

The most popular DeepLearning model for computer vision is the **Convolutional Neural Networks (CNNs),** widely employed for the study of visual data due to their capacity to automatically extract pertinent features off raw data, they enable precise and effective picture analysis. Also De facto benchmark datasets, primarily the **ImageNet, COCO, and MINST** databases, are used to compare developed approaches [2,4].

There are a lot of other architectures now being used mainly for image processing ranging from image classification to image matching using neural networks. These includes numerous techniques from **Recurrent Neural Networks (RNNs) ,**one of the promising model for time series analysis and works well on visuals data with sequential patterns, **Generative Adversarial Networks (GANs**) for generating highly realistic visual data using generator and discriminator networks**, Graph Neural Networks (GNN**) [5] ,a techniques to do inference on data represented by graphs to the **state-of-the-art** techniques like ;
-SIFT (Scale Invariant Feature Transform) algorithms, are well-known technique used for image matching and feature extraction due to its invariance to image transformations like change in scale and rotation [6],
-SURF (Speeded Up Robust Features) algorithms have been successfully used for multi-scale analysis using discrete convolution and boxfilters [7]. This outperforms the art SIFT method being fast and
-KAZE feature, a unique 2D feature detection and description technique can function in a nonlinear scale space unlike Gaussian scale space in SIFT and SURF [8]

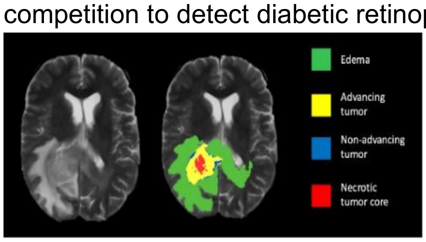


In the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) in 2012, the **AlexNet** computer vision algorithm surpassed its competitors by 10%, which marked a breakthrough in computer vision. In addition to these, another cutting-edge approach is the use of **Variational Autoencoders (VAEs),** generative models that combine deep neural networks and probabilistic modelling, applied in picture synthesis and image editing [4].

2.2 Achievements:

Use of deep learning help to achieve state-of-the-art performance in operations such as image retrieval, classification, object detection, image segmentation, image labelling using Alexnet etc. and automation of processes like **manufacturing and logistics**. It is used in **sports** to determine athletes' poses in real time which help to assess the performance [9]. Further, it has attained great achievements in major fields including:

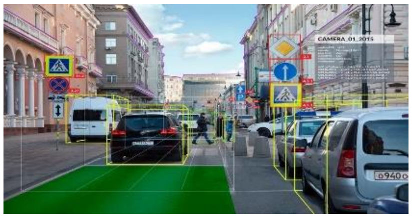
i) **Healthcare**: Medical visualization has also benefited from deep learning, which is used for:
a) segmenting images thereby **diagnosing diseases**, and treatment planning; CNN model won Kaggle competition to detect diabetic retinopathy from images of eye [1,9].



b) **image segmentation** used extensively to segment tumors
c)**image recognition** enabling researchers to distinguish between images from MRI scans that represent cancer and those that do not [1,4,9].

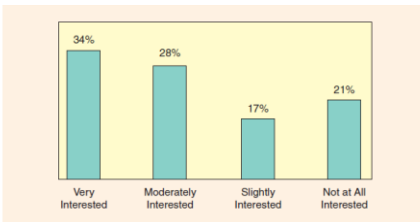
Fig.3 (Source: https://data-flair.training/blogs/image-segmentation-machine-learning/)

ii) **Automotive:** a) Use of use reasonably priced surveillance cameras to construct large-scale **traffic analysis systems** possible to recognize and classify vehicles on the road, parking lot detection and detecting violations like speeding, passing stop-signs etc.



b) A breakthrough is the **autonomous cars** which is the greatest achievement using Deep learning lately ,by which, a vehicle navigates between locations by combining sensors, cameras, radar, and artificial intelligence (AI) without the need for a humandriver and employ CNN for **object detection** and **semantic segmentation** to categorize various road features and to arrive at appropriate conclusions [9,10].

Fig.4 Object detection in Autonomous car (Source: https://becominghuman.ai/computer-vision-applications-in-self-driving-cars-610561e14118)



The figure shows the survey result by Consumer Technology Association in USA, showing the interest in replacing current vehicles with a driverless vehicle. 34% were found to be willing to switch to self-driving cars [11].

Fig.5 (Source: https://ieeexplore.ieee.org/abstract/document/7879382)

iii) **<u>Gaming:</u>** Deep learning is extensively used to create a variety of games including board games, videogames etc. A system developed by DeepMind uses reinforcement learning to play games like pong and Breakout. Google AlphaGo gaming which uses reinforcement learning is the first program which won against professional human player [9,12].

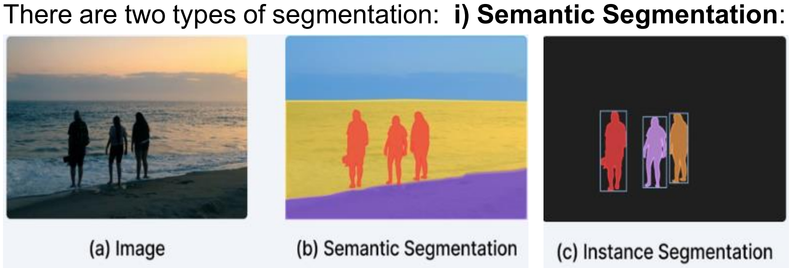
iv)**<u>Augmented Reality</u>:** Deep learning has recently been applied to Augmented Reality, or AR, which superimposes virtual objects in the real world. In AR, object detection, tracking, and registration can be made possible using deep learning algorithms, resulting in immersive and engaging experiences [1,9,12].

Most of the developments use neural networks to perform image segmentation. It is excellent for deriving meaning from the picture. The working of segmentation is explained in the section below.

3. Image segmentation using Deep Learning: Working methodology

Image segmentation, an enhancement to image classification is the process of dissecting an image into smaller parts known as "Image Objects" to extract the area of interest (ROI) from it. It uses localization in addition to classification and groups similar regions of an image to their corresponding class labels[13,14].

There are two types of segmentation: **i) Semantic Segmentation**: It involves categorization of pixels in an image into respective classes.



ii)Instance segmentation: deals with multiple objects as separate entities. Instead of using classes, instance segmentation models group pixels into categories based on "instances". Masks are differentiated using colors and category is shown by bounding boxes. Apart from traditional methods for segmentation like **Edge-based, region-based, cluster-based and threshold -based segmentation, Neural Networks** using CNNs are the new norm in image segmentation. It is widely used for segmentation because of its accuracy and can deal with large datasets [14].

Fig.6 (Source: https://www.v7labs.com/blog/image-segmentation-guide)

The network is trained using a dataset of annotated images like COCO datasets, and each image is labelled with the appropriate segmentation. After being taught, the network can subsequently be utilised to segment new images[13].

Convolution Neural Network (CNN) Principle: It functions as a hierarchical model with a number of layers, and the output is a segmented image.The basic CNN architecture used for image segmentation/classification is as shown below;

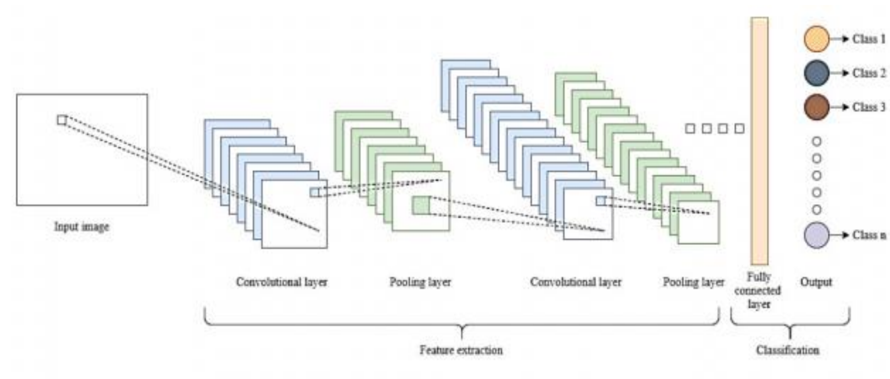
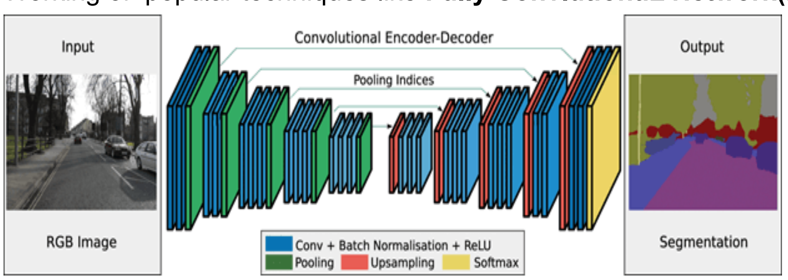


Fig.7(Source:https://vitalflux.com/cnn-basic-architecture-for-classification-segmentation/)

Convolution Layer 1: It takes the input image and filters out low level features maps like edges and the output is passed to activation function like ReLU (Recftified Linear Unit) to introduce non linearity. **Pooling layer**: It reduces the dimensionality of feature maps produced from convolution layer, ie, down-samples it. **Convolution Layer 2:** It takes the output from pooling layer and derive more complex features like patterns .Activation is done using ReLU**. Max Pooling Layer2 :** It further downsamples the output feature maps from which further classification/ segmentation is obtained [2,15,16].

These are achieved using **Tensorflow; Keras or opencv library** in Python which has functions for convolution**(Conv2D)**, pooling **(MaxPooling2D**), Activation function for ReLU etc[14,15,16] .

Working of popular techniques like **Fully ConvlutionaL Network(FCN)** and Mask R-CNN is analysed .



Here is an illustration of how (FCN) performs image segmentation using encoder-decoder structure (Fig.8). CNNs extract features from the region of interest(ROI) defined by the bounding boxes which is given to FCN to carry out segmentation.

(Fig.8 Image segmentation using FCN) [2]

Mask R-CNN uses FCN for pixel wise segmentation.Feature mapping is done using **Resnet101** Architecture ,bounding boxes are detected and ROI is extracted. In order to determine the class and bounding box coordinates, R-CNN uses ROI Pool layer to computes features[13,14,15].The misalignment caused is solved by ROI Align; a fully-connected network that is applied to each RoI which performs masking by predicting a segmentation mask pixel-by-pixel[14,16].

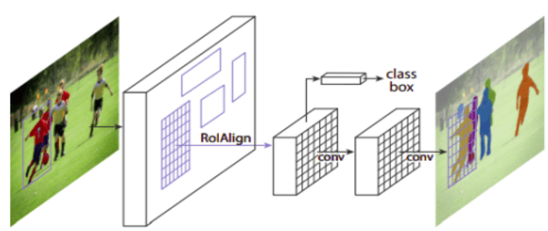


Figure 1. The **Mask R-CNN** framework for instance segmentation.

4. Conclusion:

In a nutshell, advancements in deep learning provides strong tools for automatically analysing and comprehending visual data.With its cutting edge approaches to derive patterns from images and videos, it has found use in a variety of industries, including healthcare, Automotive,Virtual Reality and more. So Computer Vision with deeplearning is here to stay with the potential to revolutionise various applications

5. References:

[1] Chai, J., Zeng, H., Li, A. and Ngai, E.W., 2021. Deep learning in computer vision: A critical review of emerging techniques and application scenarios. Machine Learning with Applications, 6, p.100134.

[2] Kumar, A. (2021). *CNN Basic Architecture for Classification & Segmentation*. [online] Data Analytics. Available at: https://vitalflux.com/cnn-basic-architecture-for-classification-segmentation/.

[3] Deep Learning for Audio Detection and Video Analysis in Railway Applications - Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/figure/Classical-Machine-Learning-vs-Deep-Learning_fig3_355412685/. [accessed 20 Apr 2023]

[4] O'Mahony, N., Campbell, S., Carvalho, A., Harapanahalli, S., Hernandez, G.V., Krpalkova, L., Riordan, D. and Walsh, J., 2020. Deep learning vs. traditional computer vision. In *Advances in Computer Vision: Proceedings of the 2019 Computer Vision Conference (CVC), Volume 1 1* (pp. 128-144). Springer International Publishing.

[5] neptune.ai. (2020). *Graph Neural Network and Some of GNN Applications: Everything You Need to Know*. [online] Available at: https://neptune.ai/blog/graph-neural-network-and-some-of-gnn-applications.

[6] Sachdeva, V.D., Baber, J., Bakhtyar, M., Ullah, I., Noor, W. and Basit, A., 2017. Performance evaluation of SIFT and Convolutional Neural Network for image retrieval. *International Journal of Advanced Computer Science and Applications*, *8*(12).

[7] Oyallon, E. and Rabin, J. (2015). An Analysis of the SURF Method. *Image Processing Online*, [online] 5, pp.176–218. doi: https://doi.org/10.5201/ipol.2015.69.

[8] Robesafe (n.d.). *kaze*. [online] www.robesafe.com. Available at: http://www.robesafe.com/personal/pablo.alcantarilla/kaze.html#:~:text=KAZE%20Features%20is%20a%20novel [Accessed 22 Apr.2023].

[9] Logunova, I. (2022). *Deep Learning Applications for Computer Vision | Serokell*. [online] Serokell Software Development Company. Available at: https://serokell.io/blog/deep-learning-for-computer-vision.

[10] Barla, N. (2021). *Self-Driving Cars with Convolutional Neural Networks (CNN)*. [online] neptune.ai. Available at: https://neptune.ai/blog/self-driving-cars-with-convolutional-neural-networks-cnn.

[11] Markwalter, B. (2017). The Path to Driverless Cars [CTA Insights]. *IEEE Consumer Electronics Magazine*, 6(2), pp.125–126. doi:https://doi.org/10.1109/mce.2016.2640625.

[12] Vedpathak, O. (2019). *Playing Pong from pixels using Reinforcement Learning*. [online] Medium. Available at: https://towardsdatascience.com/intro-to-reinforcement-learning-pong-92a94aa0f84d.

[13] Derrick Mwiti, Katherine (Yi) Li (2020). *Image Segmentation in 2021: Architectures, Losses, Datasets, and Frameworks | Neptune Blog*. [online] neptune.ai. Available at: https://neptune.ai/blog/image-segmentation/.

[14] DataFlair (2020). *Image Segmentation with Machine Learning*. [online] Data Flair. Available at: https://data-flair.training/blogs/image-segmentation-machine-learning/.

[15] Bandyopadhyay, H. (n.d.). *A Gentle Introduction to Image Segmentation for Machine Learning*. [online] www.v7labs.com. Available at: https://www.v7labs.com/blog/image-segmentation-guide.

[16] Potter, R. (2022). *Image Segmentation: The Deep Learning Approach*. [online] indiaai.gov.in. Available at: https://indiaai.gov.in/article/image-segmentation-the-deep-learning-approach.