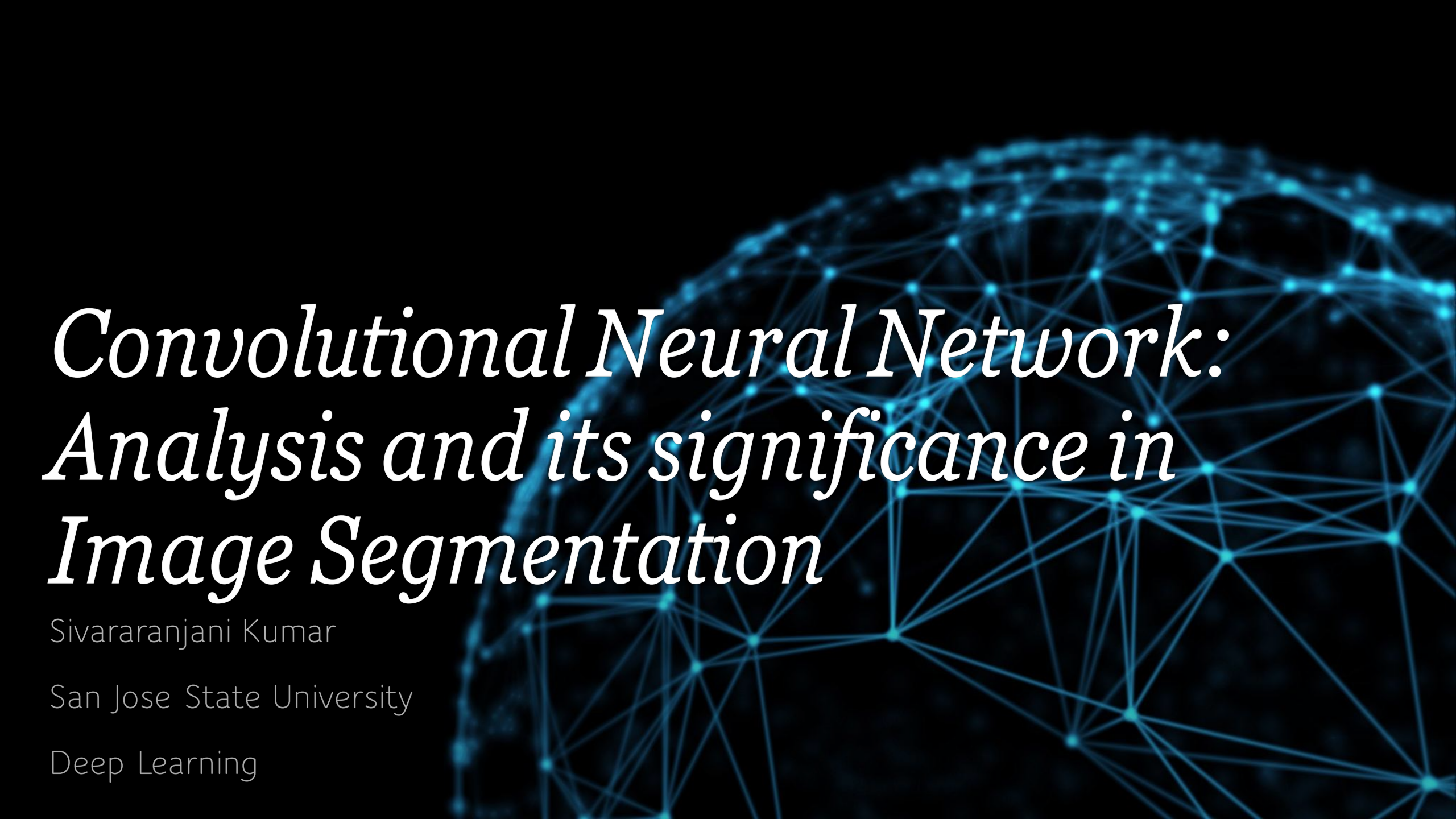


Convolutional Neural Network: Analysis and its significance in Image Segmentation

Sivararanjani Kumar

San Jose State University

Deep Learning

Introduction

CNN is the most widely used deep learning Neural Network for image processing. It has made some impressive achievements for the past few years in many fields, not just limited to computer vision and natural language processing.

We must have heard lot about the applications of CNN in different scenarios, but not from the general perspective. Here, I aim to provide some novel ideas which are proposed recently about CNN and also its significance in image segmentation.

History of CNN

The concept of CNN was not a new one. It all started from the discovery of Hubel and Wiesel which explained that there are simple and complex neurons in the primary visual cortex and the visual processing always starts with simple structures such as oriented edges.

The term ‘Convolution’ was first used by LeCun et al, he constructed a convolutional neural network for a handwritten zip code recognition which is the original version of LeNet.

Convolutional Neural Network

- It is a kind of feedforward neural network that is able to extract features from image data with convolution structures. The architecture of CNN is inspired by visual perception.
- Each neuron is no longer connected to all neurons of the previous layer.
- Reduces the number of parameters and speed up convergence.
- A group of connections can share the same weights.
- Down-sampling dimensionality reduction.

Convolutional Neural Network Model

Padding: When setting a convolution kernel with a certain size, we will lose information in the border. Hence, padding is required to enlarge the input size with zeros.

Stride: To control the density of convolving, stride is employed. Larger the stride, lower the density.

Feature maps: Convolution is a pivotal step for feature extraction. The outputs of convolution can be called as feature maps.

Pooling: It is used to obviate redundancy. We have two types one is max pooling and average pooling.

Classic CNN Models

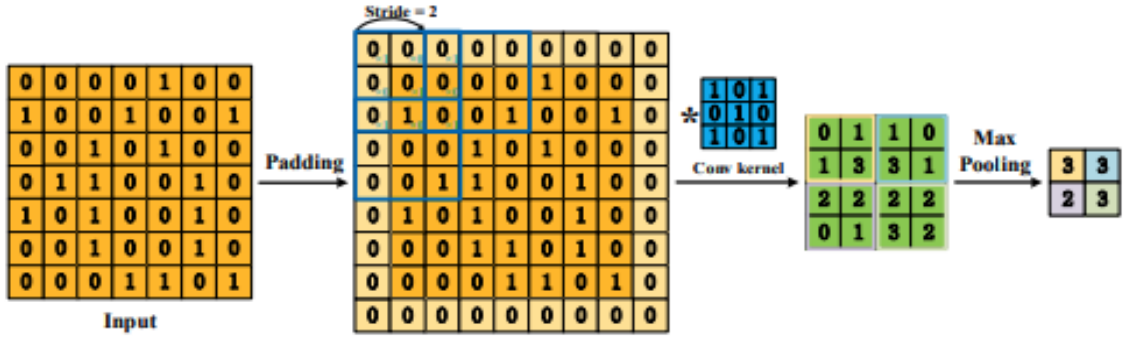
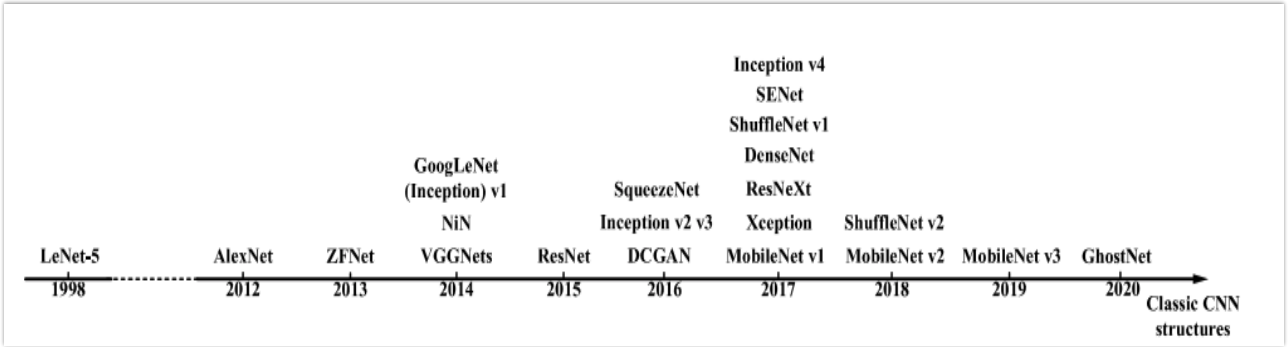


Fig. 1. Procedure of a two-dimensional CNN

There are some pre-trained CNN models that can be used for processing image data, instead of building our own model from scratch which are time consuming

Pre-Trained Models



There are some pre-trained CNN models that can be used for processing image data, instead of building our own model from scratch which are time consuming.

Applications of CNN

- CNN is able to harness a massive amount of data to achieve a promising result. Hence, there are lots of applications that come up. It can be used not only in the processing of two-dimensional images but also in one-dimensional and multi-dimensional scenarios. Some of the widely used applications are
- Image classification
- Object detection
- Image segmentation
- Face recognition

Image Segmentation

As humans, if given a picture of any object we will be able to identify it. For instance, when we want to cross the road at intersection, we typically look left and right and make our decision. It's because our brain is able to analyze a kind of vehicle that comes, or it could be any object in a given scenario.

The question is, can machines do that??? The answer was 'no' till few years back. But with the advancements in computer vision, things have changed too fast and too soon. Below is an example of segmented image verses natural image

Image segmentation Types

Image segmentation is basically the ability of computer vision model to detect objects and be able to determine its shapes and can also predict the direction the objects that can move. It builds upon the idea of object detection.

Image segmentation is of two types:

- Semantic segmentation
- Instance segmentation and
- The combination of these two can be known as panoptic segmentation

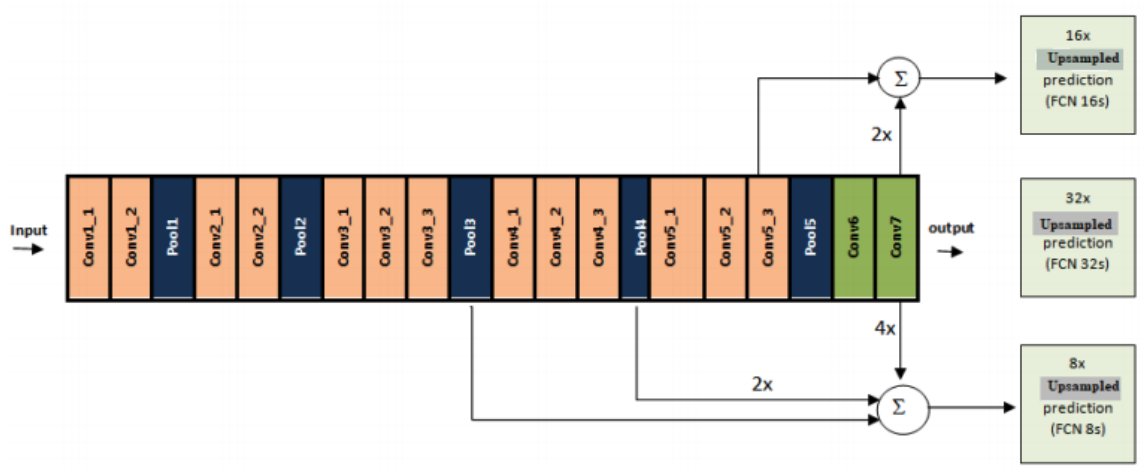
Semantic Segmentation

- Semantic segmentation describes the process of associating each pixel of an image with a class label
- There are different ways to extract *segmented images*,
- Multi-scale CNN have been used for scene labeling and achieve state-of-the-art results in the Sift flow.
- R-CNN used selective search algorithm to extract region proposals first and then applied upon each proposal.
- R-CNN achieved record result over second order pooling which was a leading hand-engineered semantic segmentation system.

State-of-the-art Semantic Segmentation Models

Among different CNN based semantic segmentation models, Fully Convolutional Network (FCN) gained the maximum attention and an FCN based semantic segmentation model trend has emerged.

- DeepLab, Deconvnet, U-Net, SegNet, etc,.

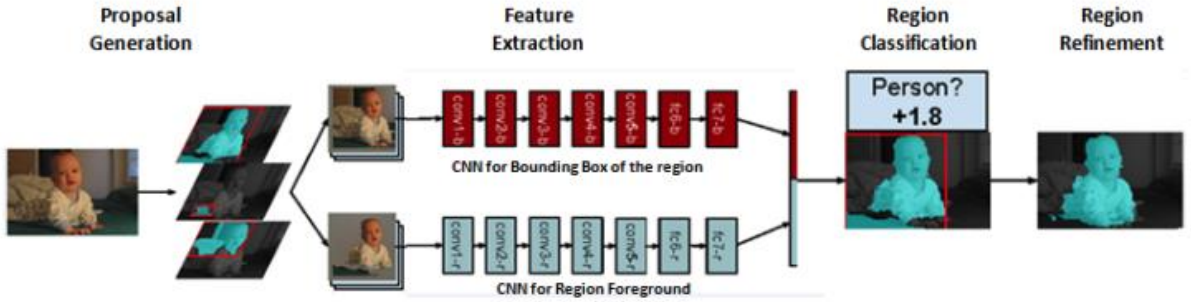


Instance Segmentation

- Unlike semantic segmentation, *instance segmentation* masks each instance of an object contained in an image independently.
- In object detection, researchers use the bounding box to detect each object instance of an image with a label for classification.
- Instance segmentation put this task one step forward and put a segmentation mask for each instance.
- Instance segmentation task only adds a segmentation mask to the output of object detection task. That is why most of the CNN based instance segmentation models have used different CNN based object detection models to produce better segmentation accuracy and to reduce test time.

State-of-the-art Instance Segmentation Models

- Simultaneous Detection and Segmentation (SDS) model consists of 4 steps for instance segmentation. The steps are proposal generation, feature extraction, region classification, and region refinement respectively
- Then each region proposals are fed into two CNN based sibling networks.
- DeepMask, Multi-task Network Cascades (MNC), InstanceFCN, Mask R-CNN, etc.



Panoptic Segmentation

- In *Panoptic segmentation*, we need to associate all the pixels in the image with a semantic label for classification and also identify the instances of a particular class.
- The output of a panoptic segmentation model will contain two channels: one for pixel's label (semantic segmentation) and another for predicting each pixel instance (instance segmentation).
- On the other hand, fine-tuning of hyper-parameters, data pre-processing methods, choice of the loss function and optimization function, etc. are also play an important role in the success of a model that we choose

Conclusion

We have discussed about an overall view of convolutional neural network, its history, analysis, applications and prospects. We have also talked about image segmentation using CNN. The advantages of convolutional neural networks, such as local connection, weight sharing, and down-sampling dimensionality reduction, have been widely deployed in both research and industry projects.

Even though convolutions possess many benefits and have been widely used for many applications, we reckon that it can be refined further in terms of model size, security, and easy hyperparameters selection.

Thank You