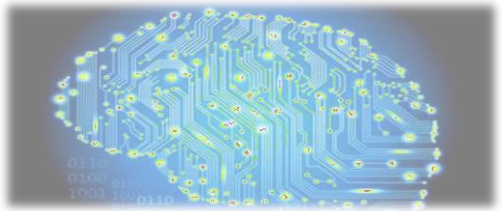
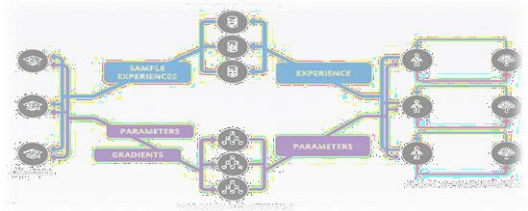
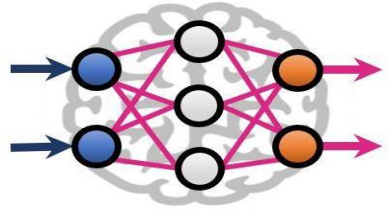
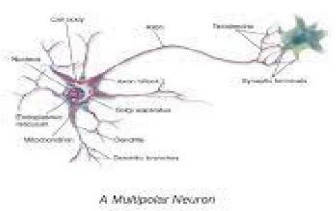
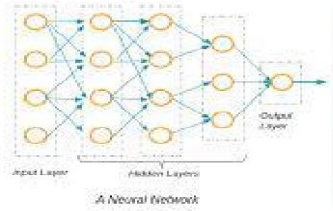





Fully Convolutional Networks for Semantic Segmentation

Sujeong Han,
0001entropy@gmail.com
2021. 12. 14.

Abstract

Convolutional networks are powerful visual models that yield hierarchies of features. We show that convolutional networks by themselves, trained end-to-end, pixelsto-pixels, exceed the state-of-the-art in semantic segmentation. Our key insight is to build “fully convolutional” networks that take input of arbitrary size and produce correspondingly-sized output with efficient inference and learning. We define and detail the space of fully convolutional networks, explain their application to spatially dense prediction tasks, and draw connections to prior models.

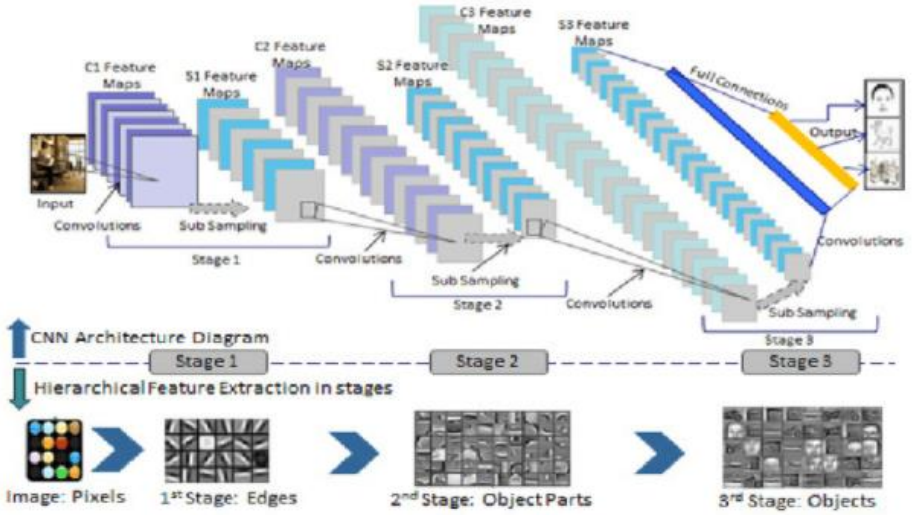
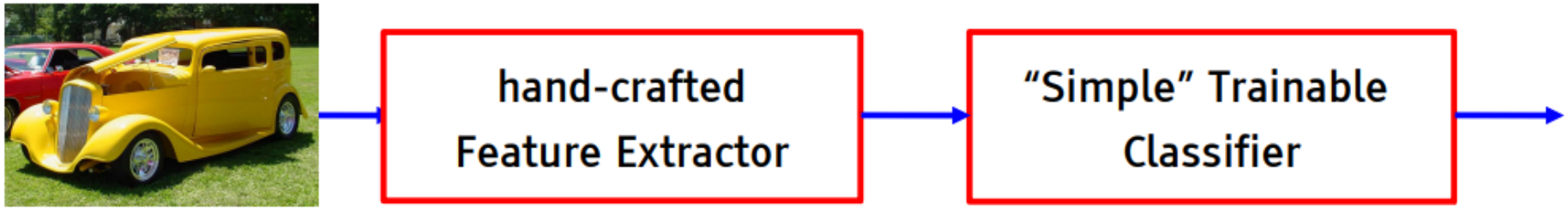


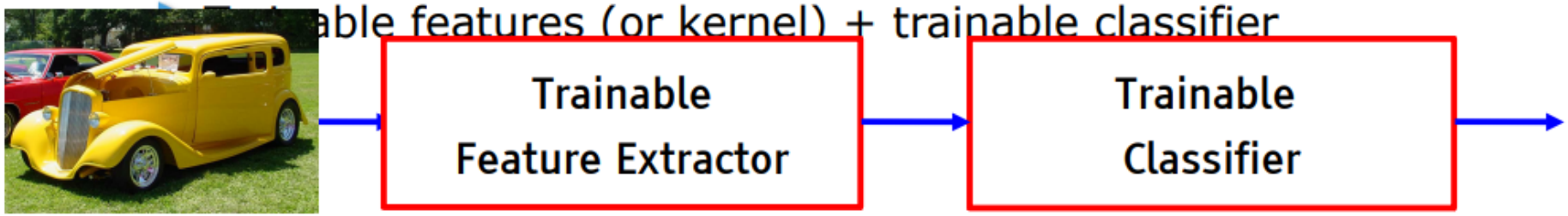
그림1 이미지 출처: https://www.researchgate.net/figure/Learning-hierarchy-of-visual-features-in-CNN-architecture_fig1_281607765

Introduction

- The traditional model of pattern recognition (since the late 50's)
  - Fixed/engineered features (or fixed kernel) + trainable classifier

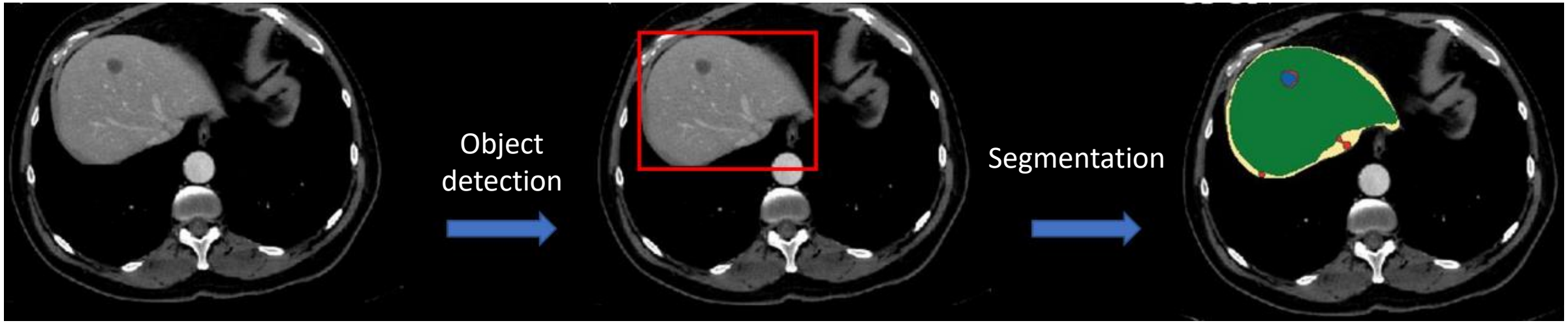


- End-to-end learning / Feature learning / Deep learning
  - able features (or kernel) + trainable classifier



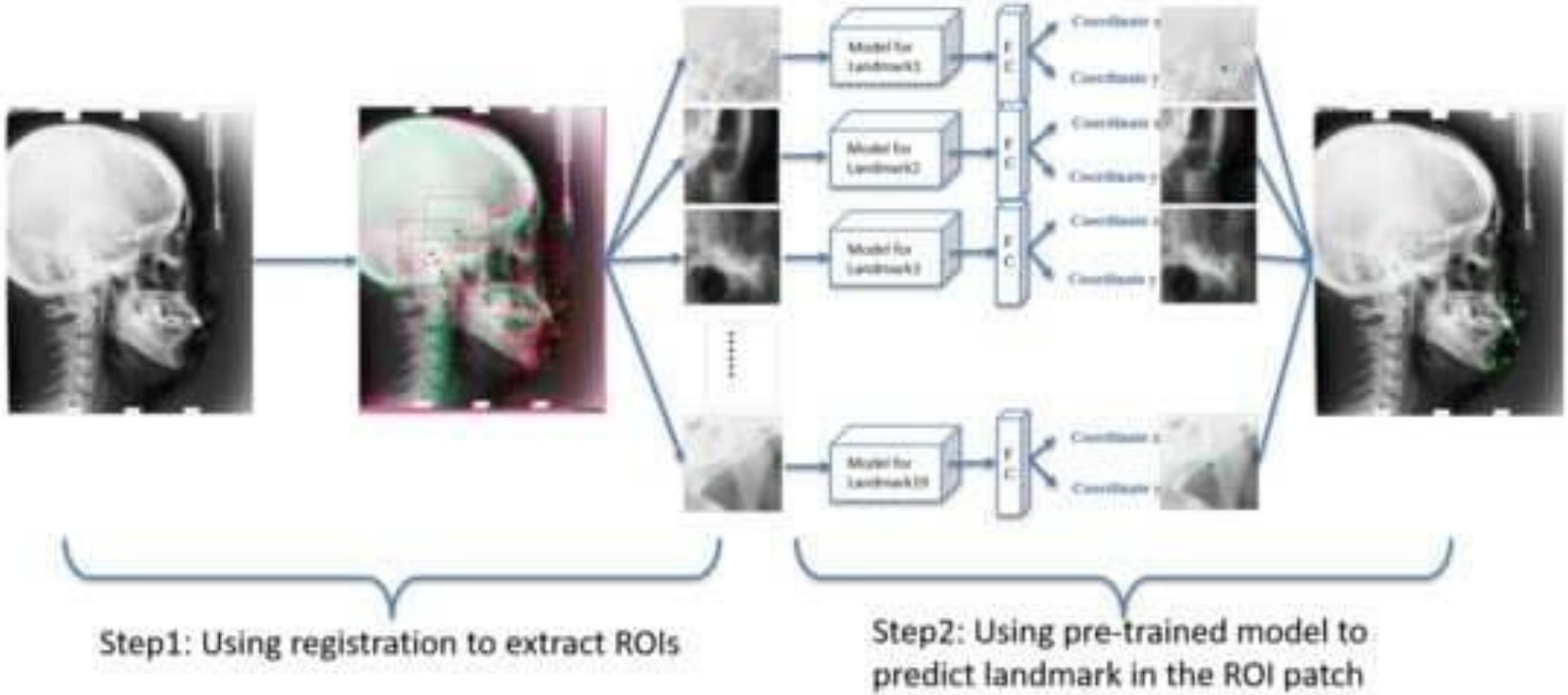
출처 : stackoverflow.com/questions/40552928/what-is-meant-by-end-to-end-fashion-convolutional-neural-networkcnn

Introduction



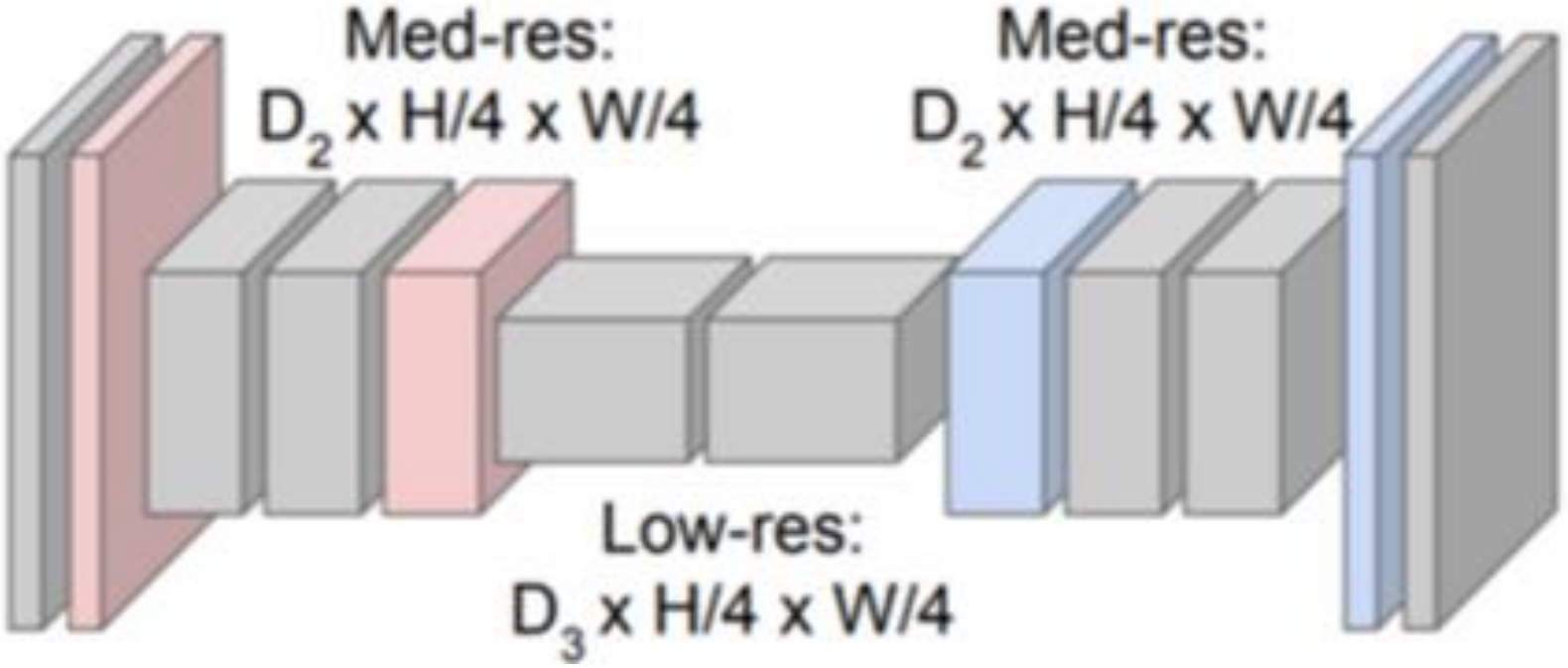
출처 https://www.programmersought.com/article/21122704833/

Introduction



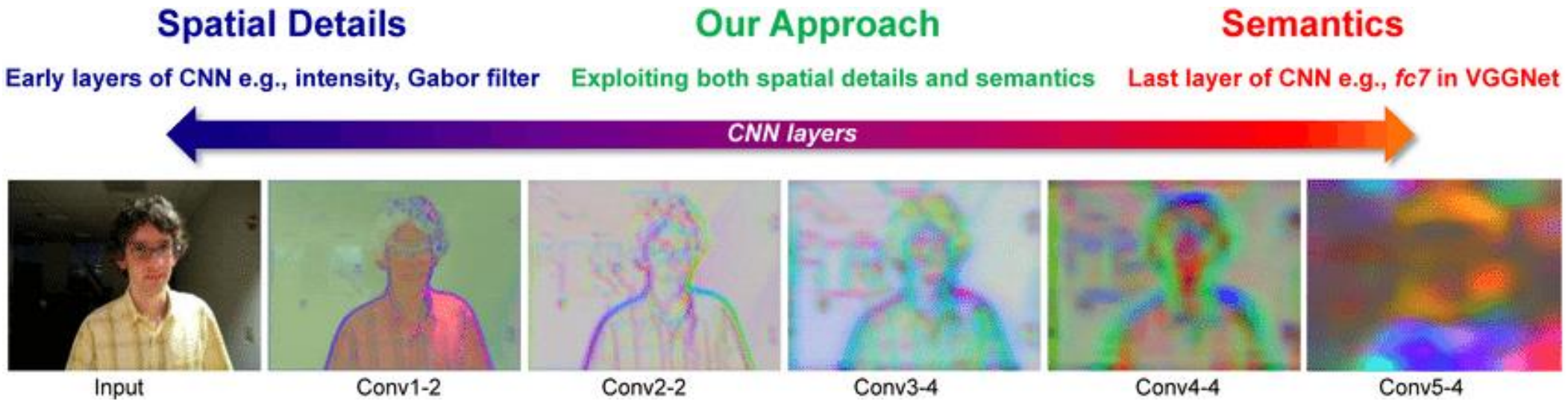
Introduction

Fully convolutional networks <-"An FCN naturally operates on an input of any size, and produces an output of corresponding (possibly resampled) spatial dimensions."



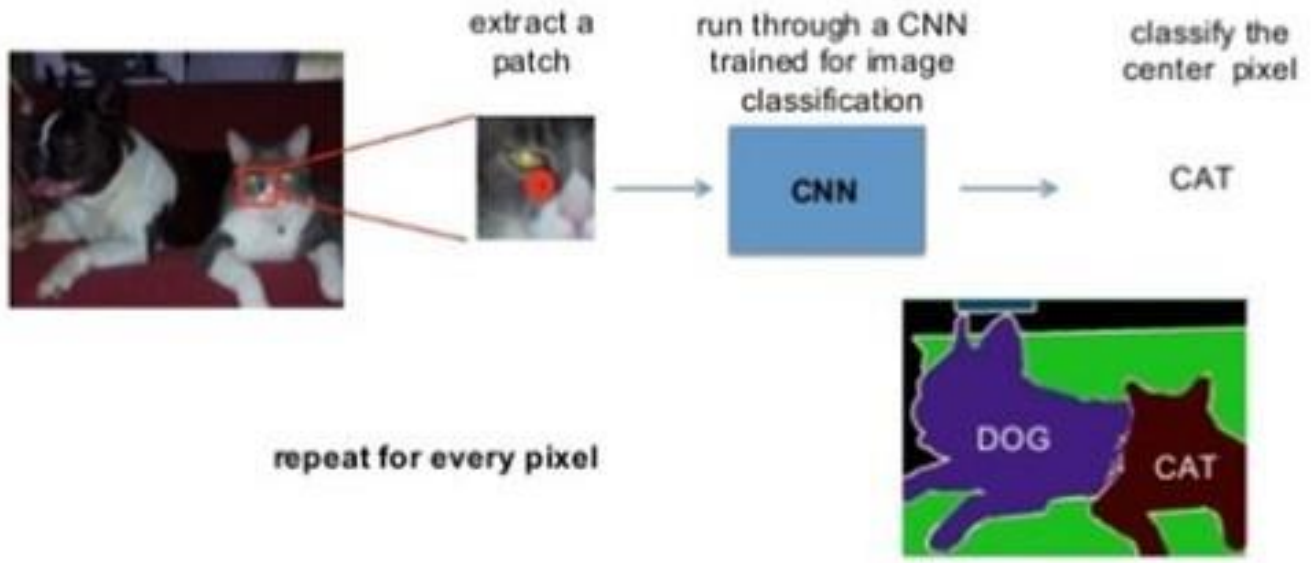
Introduction

We then define a skip architecture that combines semantic information from a deep, coarse layer with appearance information from a shallow, fine layer to produce accurate and detailed segmentations



Introduction

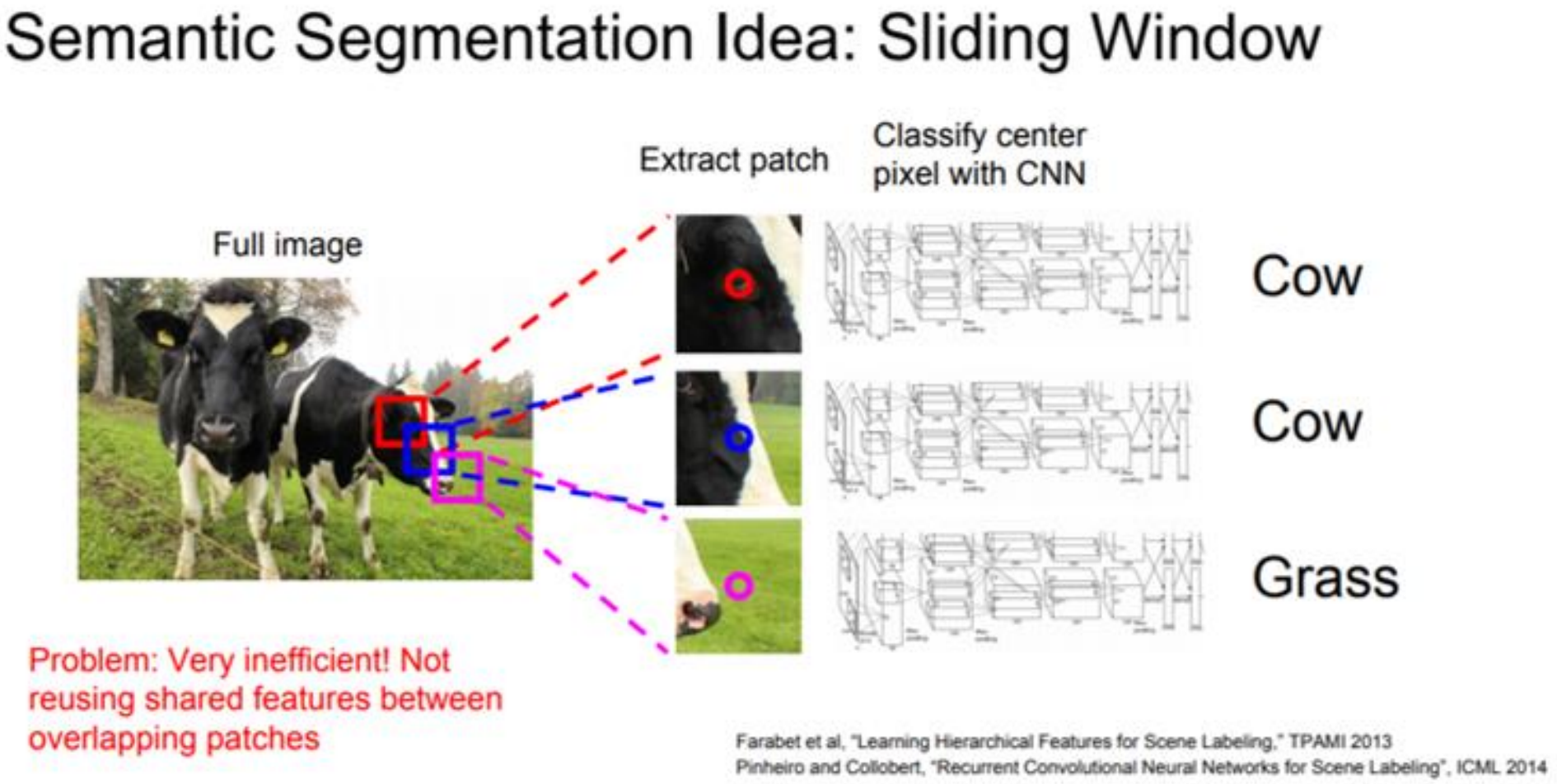
Patchwise learning



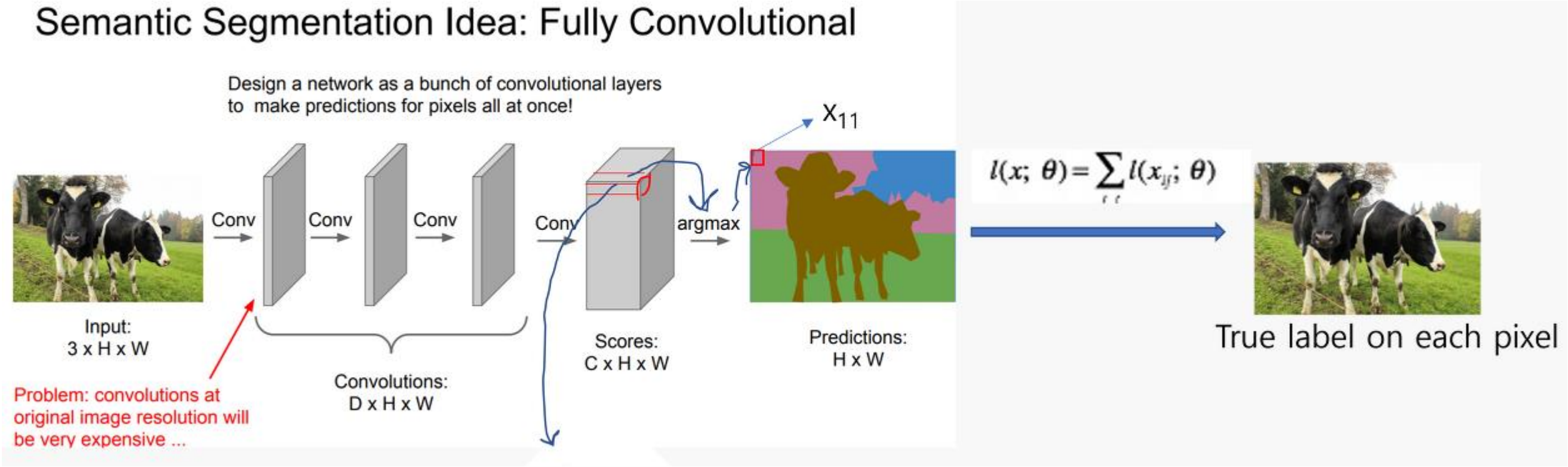
Slide window 방식

Introduction

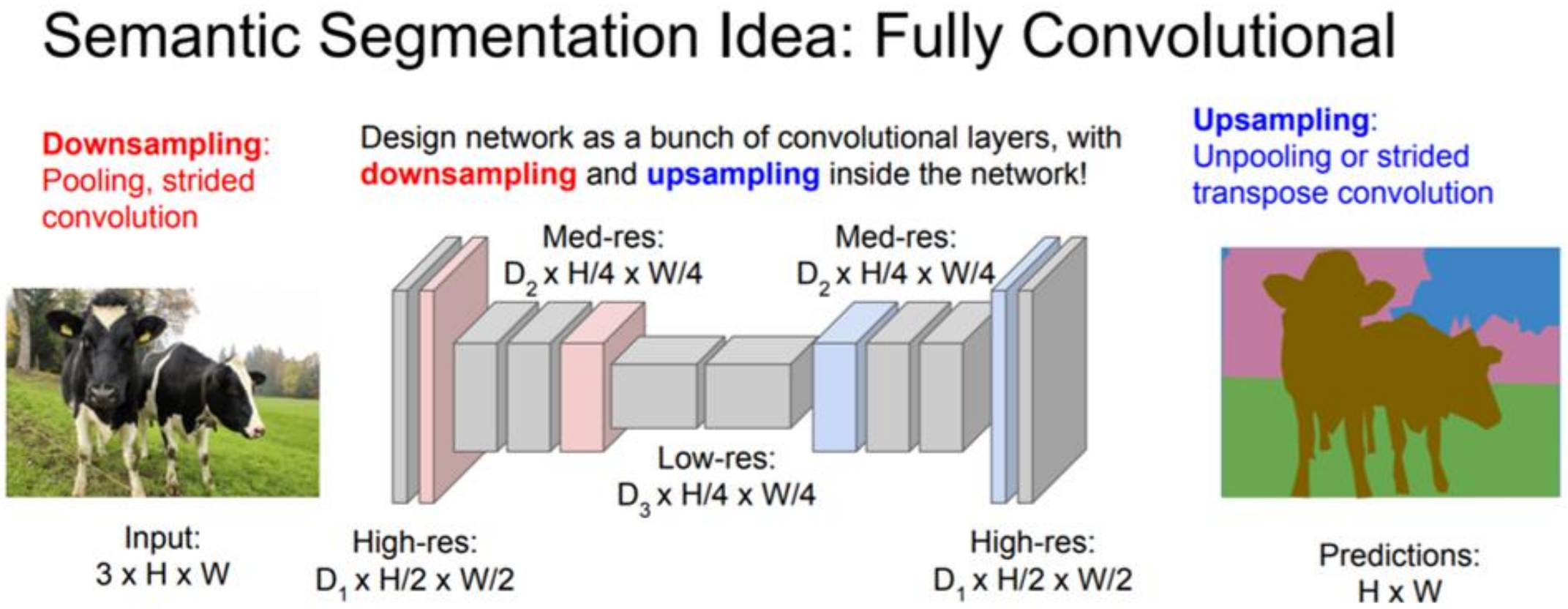
Patch learning Problem



Introduction



Introduction



Introduction

lacks the efficiency of fully convolutional training. Our approach does **not make use of pre- and post-processing complications**, including **superpixels**, **proposals**, or **post-hoc refinement** by random fields or local classifiers . Our model transfers recent success in classification to dense prediction by reinterpreting classification nets as fully convolutional and **fine-tuning from their learned representations**. In contrast, previous works have applied small convnets without supervised pre-training.

Semantic segmentation faces an **inherent tension** between **semantics** and **location**: global information resolves what while local information resolves where. Deep feature hierarchies encode location and semantics in a nonlinear 1 local-to-global pyramid. We define a skip architecture to take advantage of this feature spectrum that **combines deep, coarse, semantic information and shallow, fine, appearance information in Section .**

Thank you