Image Aesthetics

WHAT DEFINES AN IMAGE ?

Image Aesthetics

★ Image aesthetic evaluation aims to classify photos into high quality or low quality from the perspective of human.

★ Many visual features have been explored under this formulation (handcrafted features), ranging from low-level image statistics, such as edge distributions and color histograms, to high-level photographic rules, such as the rule of thirds and golden ratio

A SUBJECTIVE CONJECTURE

Beautiful Capture

“BEAUTY IS REALLY IN THE EYE OF THE BEHOLDER”

- While everyone has different tastes, there are universally accepted norms when it comes to beauty – things which everyone pretty much agrees are beautiful, like sunsets or sunrises over the mountains or the ocean.

A SUBJECTIVE CONJECTURE

Beautiful Capture

TECHNIQUES OF IMAGE CAPTURING

- Image aesthetics can be affected by the different usages of **lighting**, **contrast** , and image **composition**

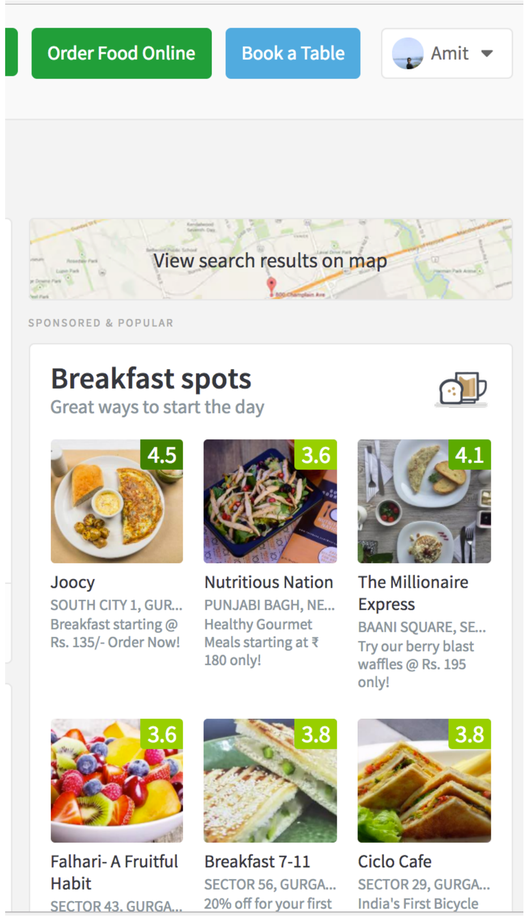
FACTORS LIKE LIGHTING,CONTRAST,ANGLE DETERMINE THE AESTHETICS

The Two Images

Significance of Aesthetic Assessment

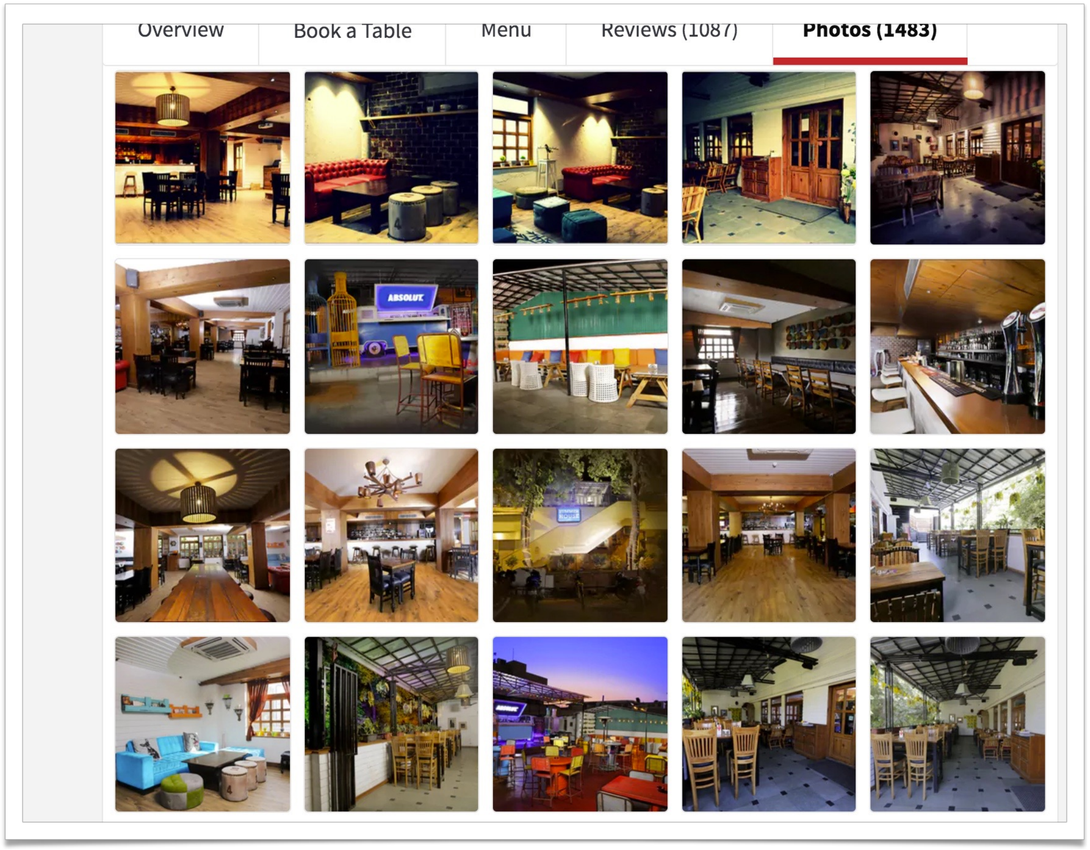
- With the ever increasing user generated content, biggest challenge is to showcase high quality images to users.

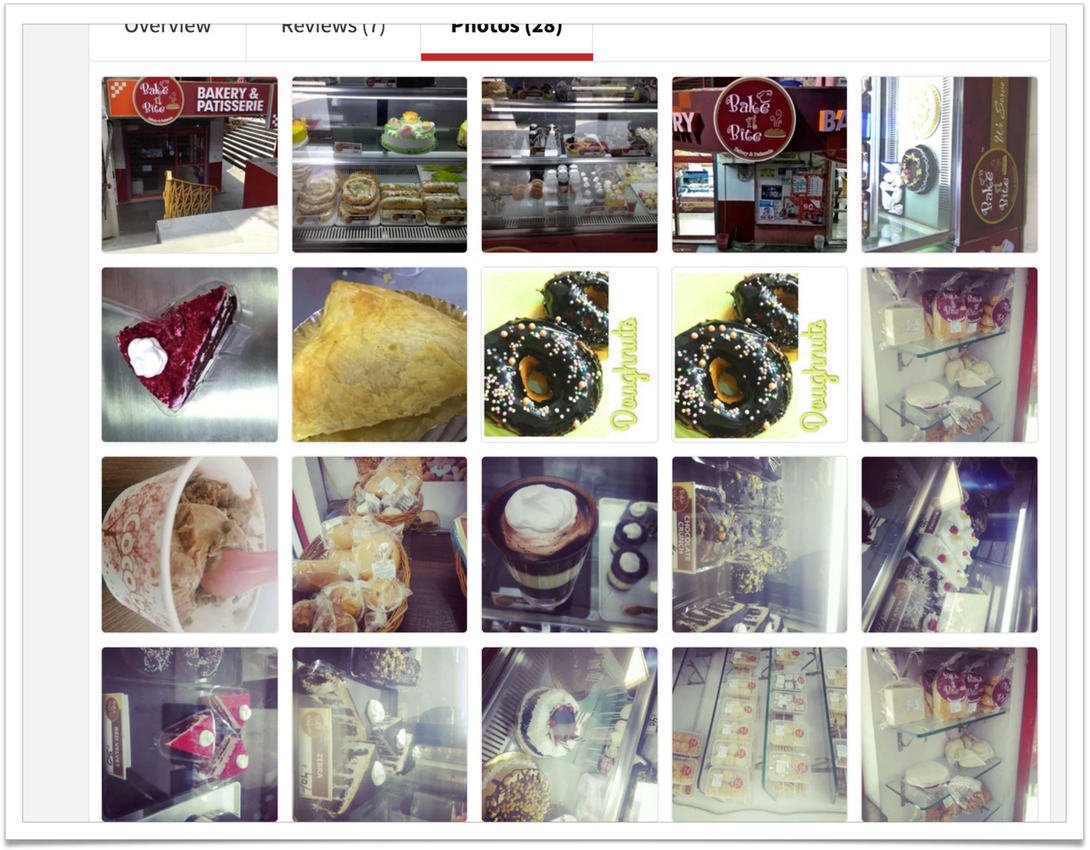
- Zomato measures the performance of restaurant for ad sales with click through rate (CTR) as one of the parameter. Restaurants with High Quality Display Image have greater CTR.



PHOTOS DO HAVE A STORY

Photo Gallery

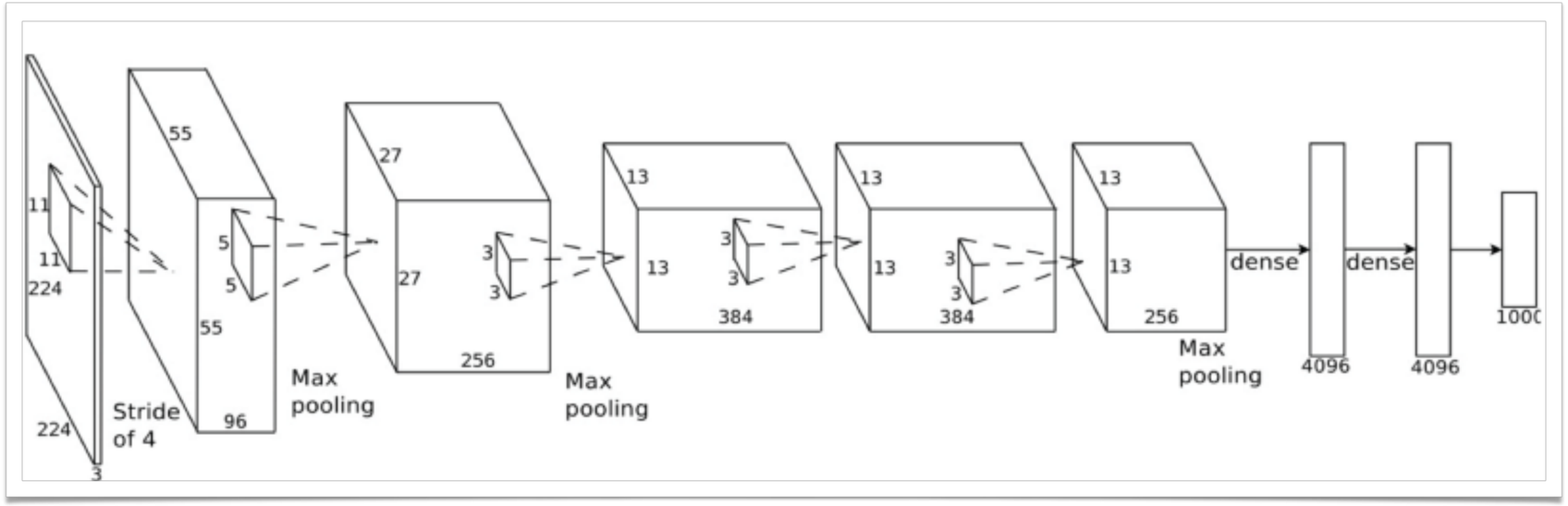




Let’s solve Image Aesthetics with the most disruptive technology over the last five years.

STATE OF THE ART IN MANY FIELDS NOW

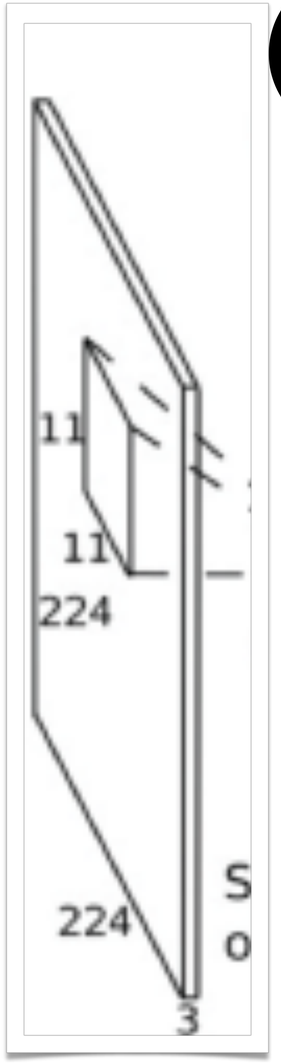
Deep Learning



DEEP MODEL INPUT LAYER

Fixed Size Input Constraint

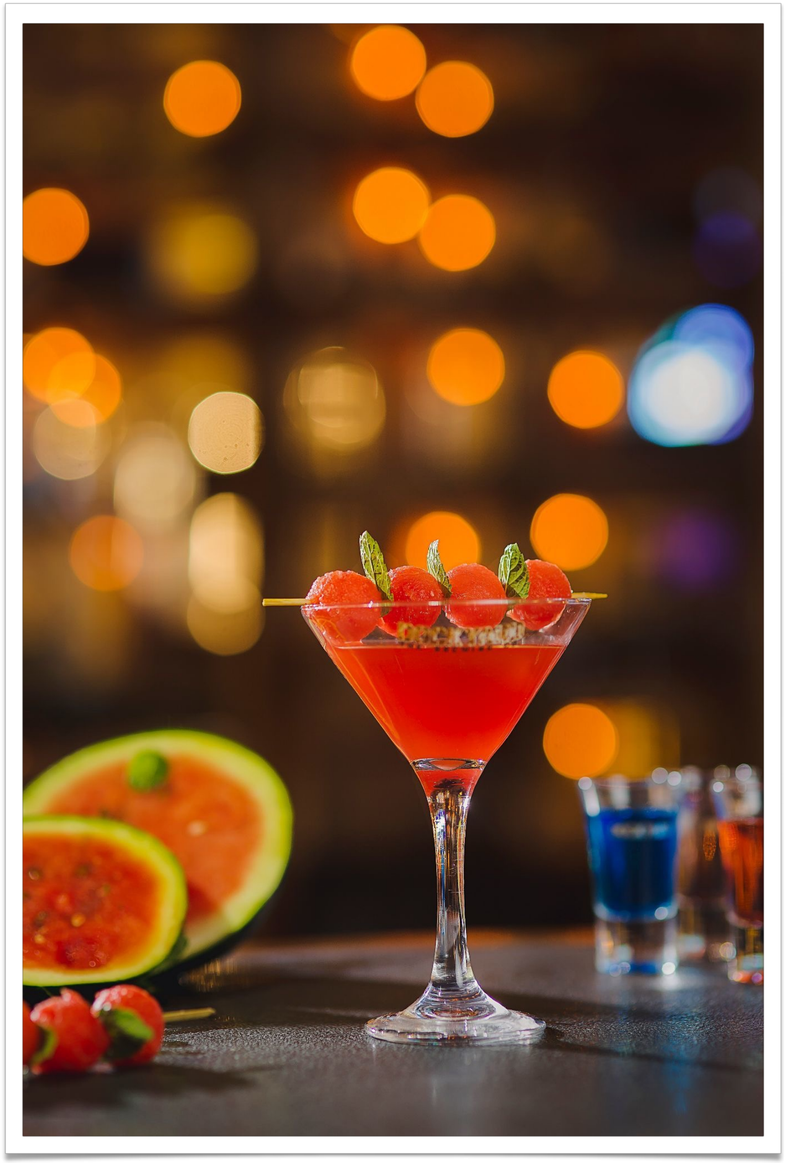
- This constraint of fixed size input in Deep Network compromises the aesthetics of original image. input images need to be transformed via cropping, scaling, or padding, which often damages image composition, reduces image resolution, or causes image distortion, thus compromising the aesthetics of the original images.



Input Layer of Network

AMAZING IN THEIR ORIGINAL ASPECT RATIO

Original Image



RESIZING LOSES THE ORIGINAL AESTHETICS OF IMAGE

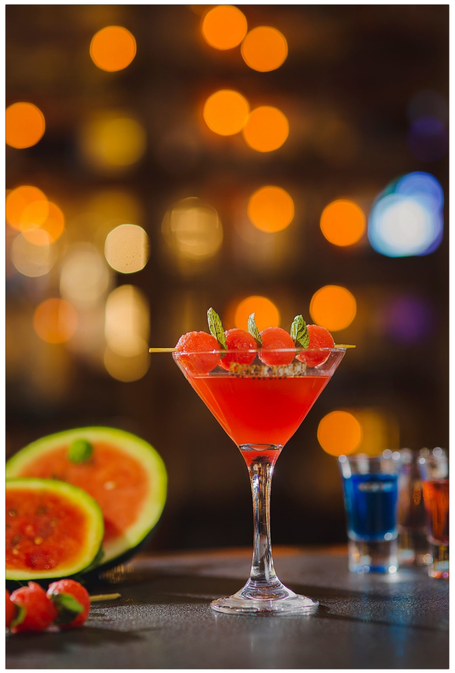
Transformed Image

ORIGINAL

RESIZED

RESIZING LOSES THE ORIGINAL AESTHETICS OF IMAGE

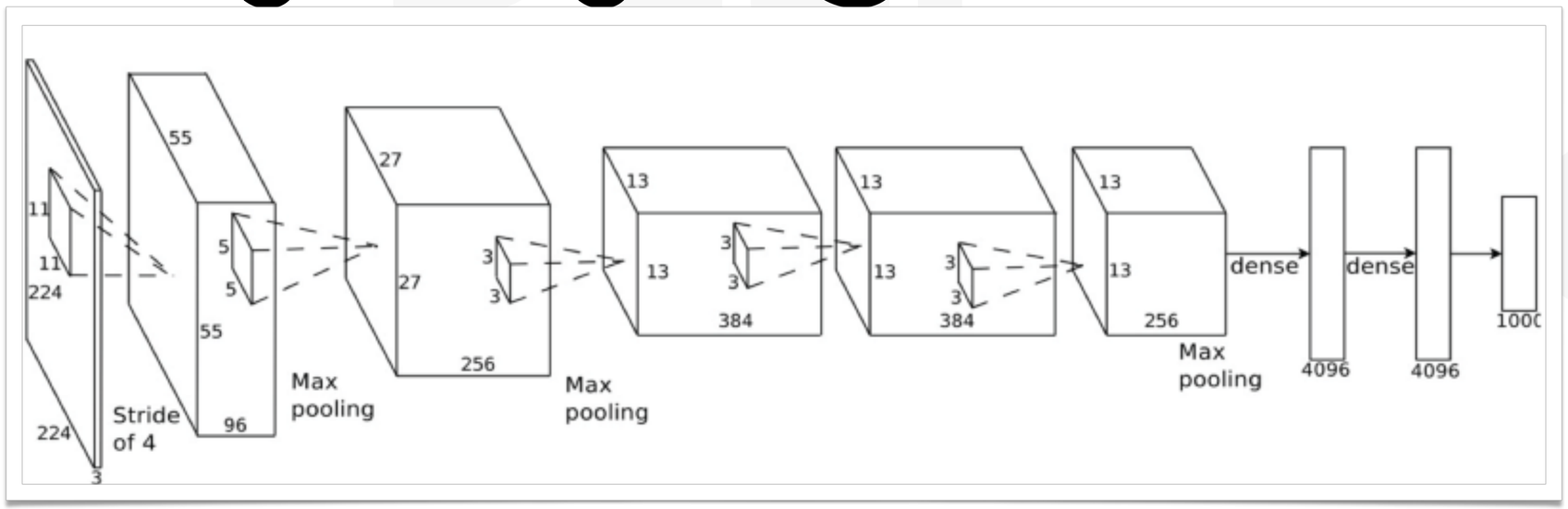
Transformed Image



ORIGINAL

RESIZED

Demystifying the Network



- The first few layers of networks are either convolution or pooling layer, followed by fully connected layers.
- The fixed length input is only the constraint of fully-connected layer. Be it convolution or pooling, both perform operations on input, with the spatial information.

Spatial Pyramid Pooling

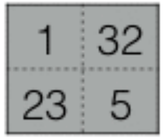
- Network with another pooling strategy, ‘spatial pyramid pooling’ just before the fully connected layer, makes it capable of taking input of any size and generate fixed length representation.

OUTPUT IS PROPORTIONAL TO INPUT

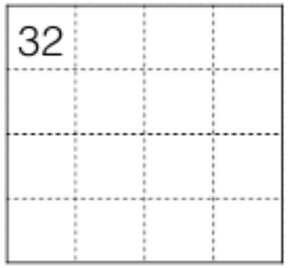
Max Pooling

Kernel size is fixed

Input



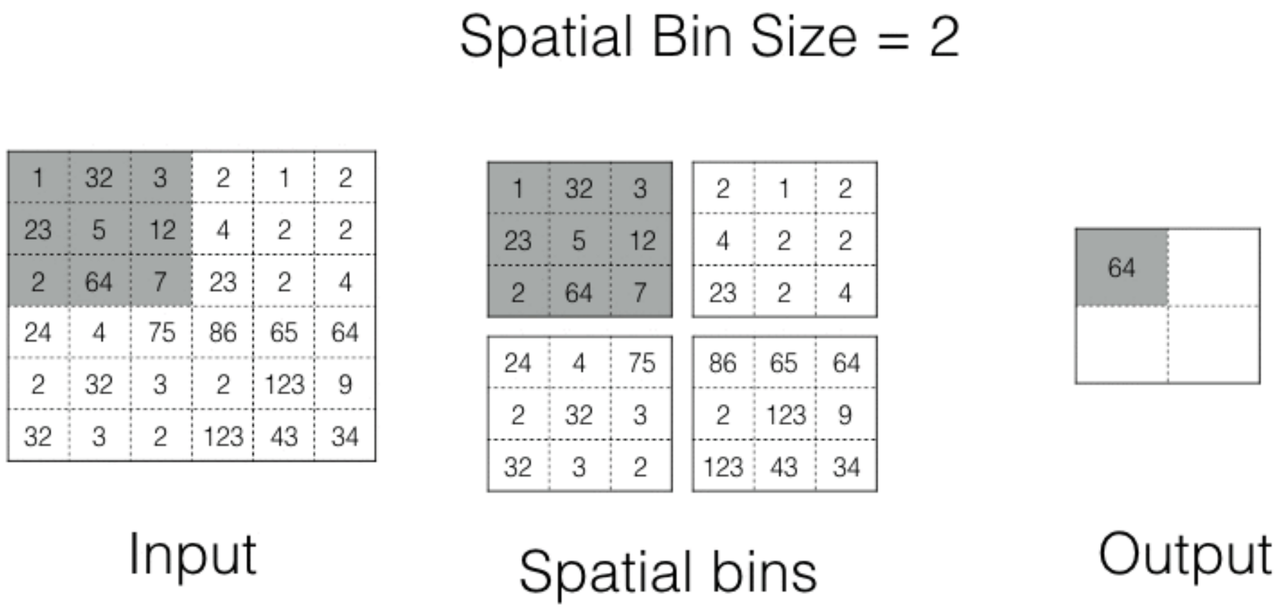
Max Pool Filter



Output

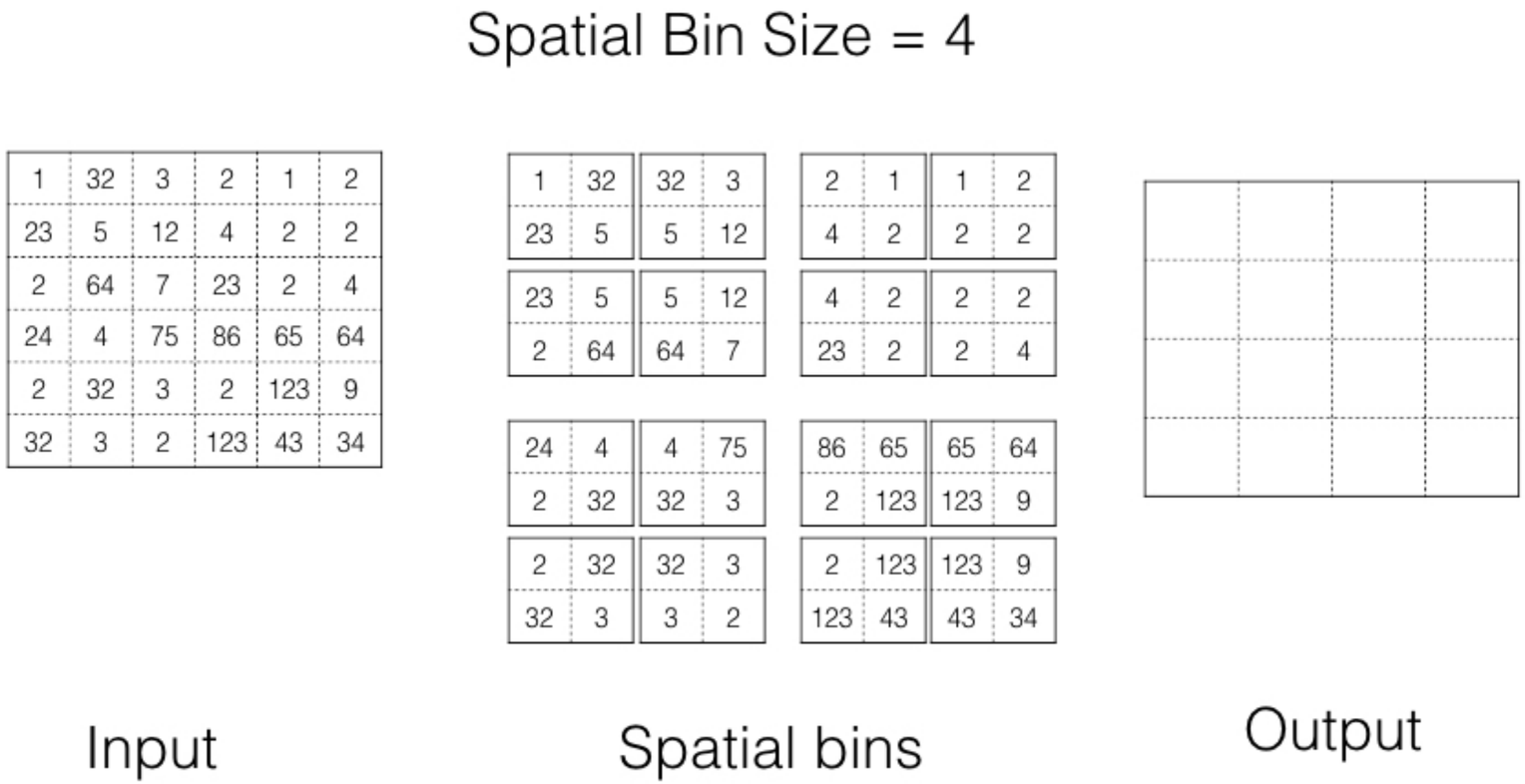
Spatial Pyramid Pooling

Output dimension (2, 2)

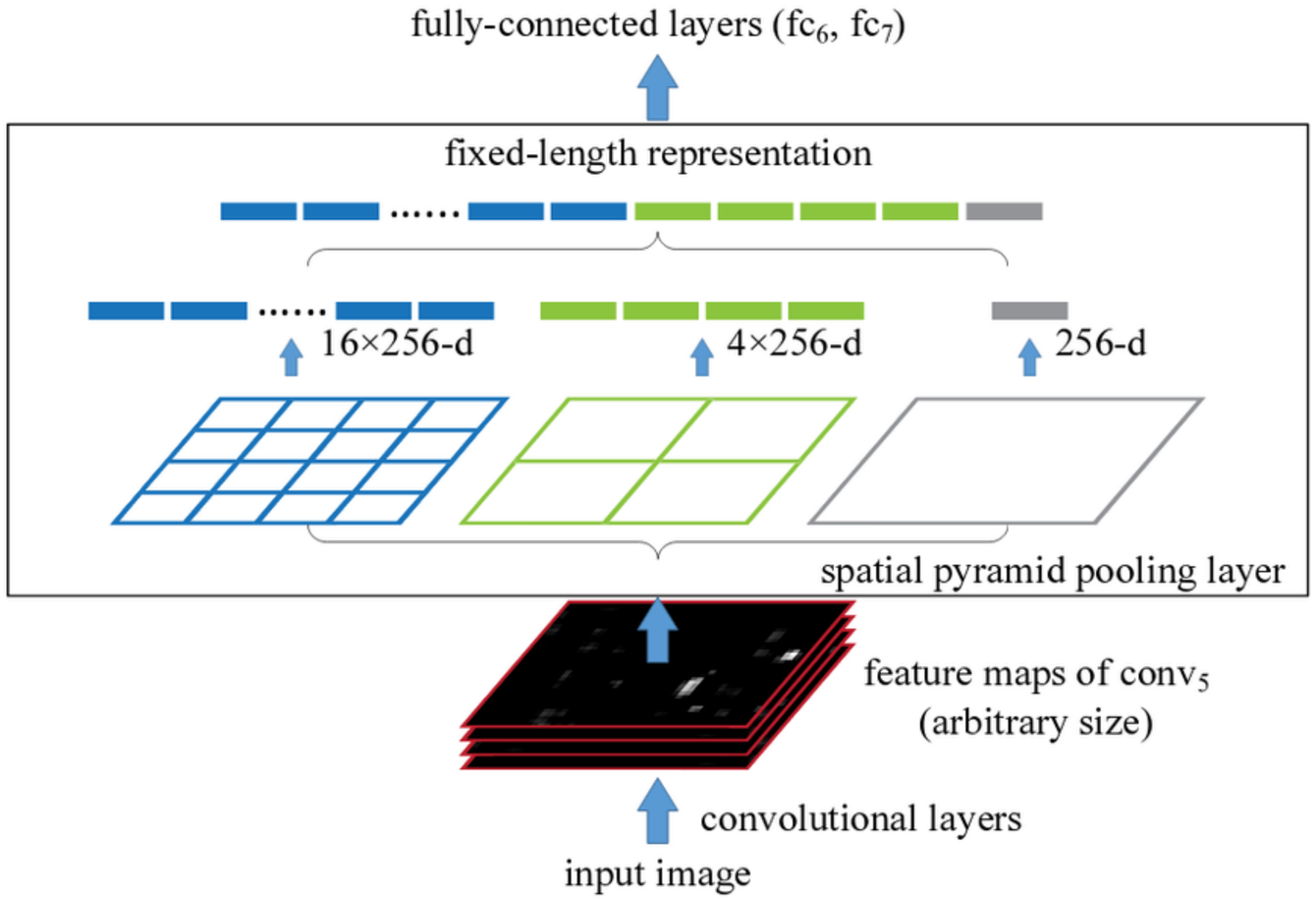


Spatial Pyramid Pooling

Output dimension (4, 4)

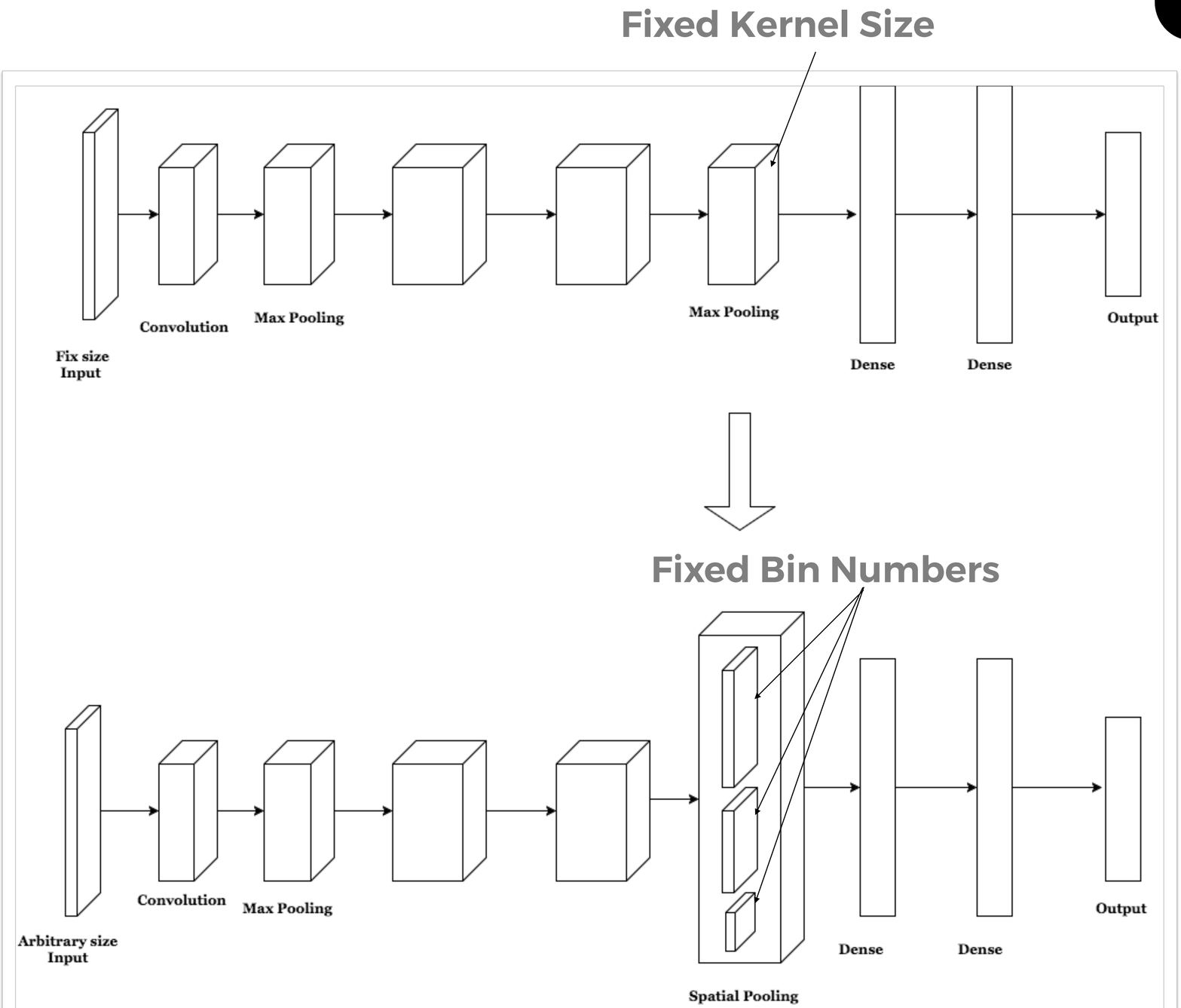


Spatial Pyramid Pooling



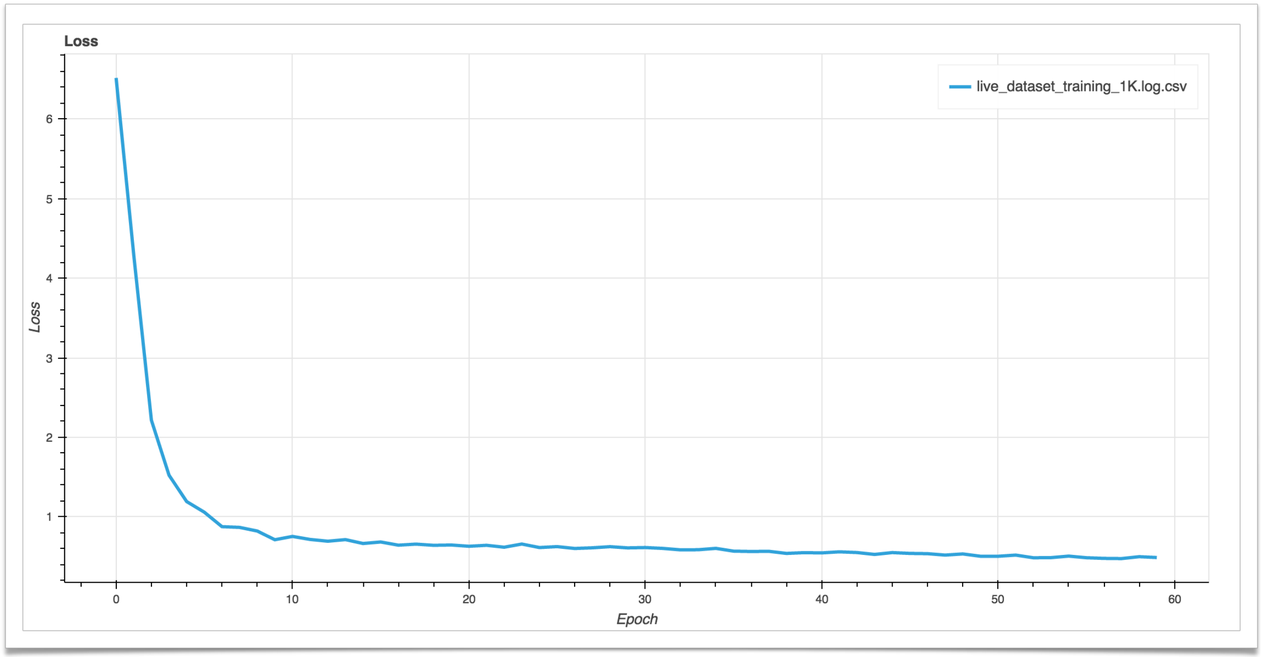
CNN

SPP-Net

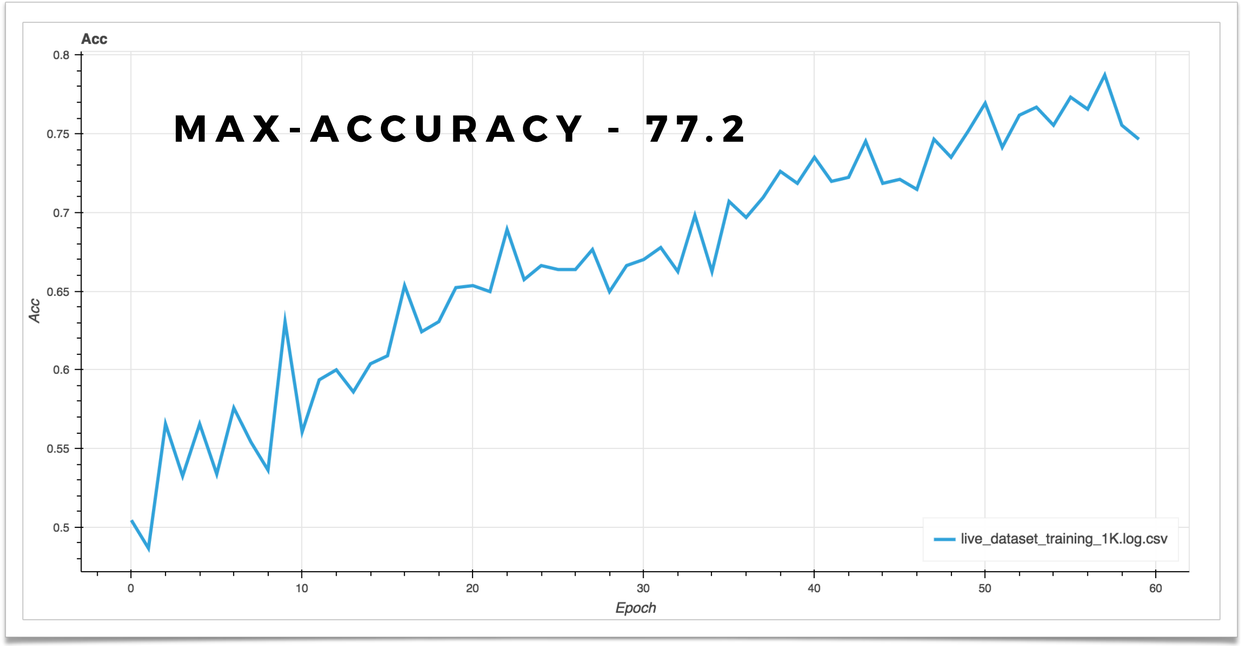

Fixed Kernel Size
Fix size Input
Convolution
Max Pooling
Max Pooling
Dense
Dense
Output
Fixed Bin Numbers
Arbitrary size Input
Convolution
Max Pooling
Spatial Pooling
Dense
Dense
Output

Training a SPP-net

ZOMATO DATASET USED



TRAINING LOSS



TRAINING ACCURACY

Interesting Prediction Results

NO REGION OF INTEREST

Spp-Score=0.46

Interesting Prediction Results

REGION OF INTEREST IS PRESENT

Spp-Score=0.94

Take Aways

SCALE INVARIANT LEARNING

It’s often neglected, but plays a major role in classification learning.

FIXED SIZE INPUT

Allows to explore CNN without the constraint of fixed size representation of input.

REFERENCES

- Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun, “Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition”

- Alex Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton, ImageNet Classification with Deep Convolutional Neural Networks 2012

- Long Mai, Hailin Jin, Feing Liu, Composition-Preserving Deep Photo Aesthetics Assessment

Thank You

Find me here

https://www.github.com/yardstick17/

https://www.linkedin.com/in/yardstick17/

amit_kushwaha@outlook.com