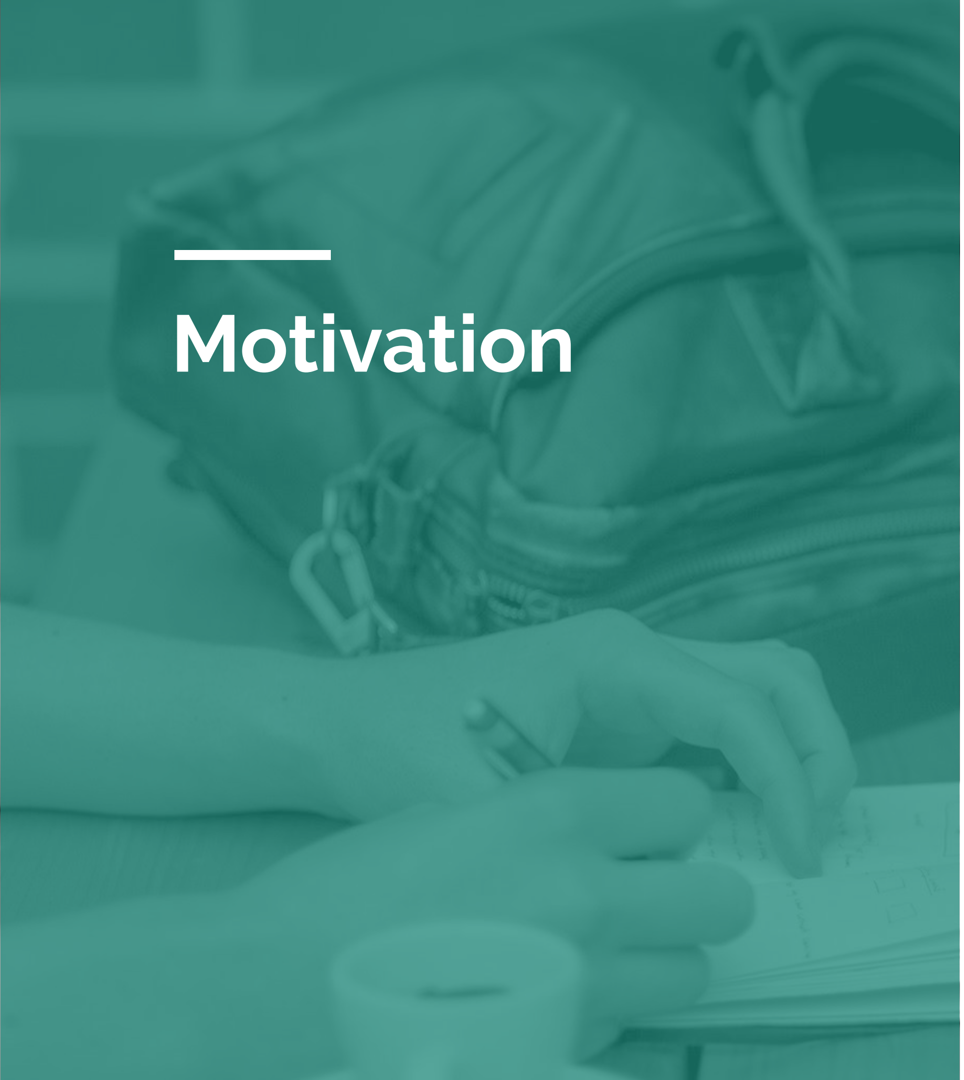
Improved Image Pre-Processing for Sharpened Object Detection in Image

CAPSTONE PROJECT

Aditya Firoda
-16BCE2184

Defense Research and Development Organisation

Under the Guidance of -
Dr. Jayakumar K. (Associate Professor)
Mr. Manish Choudhary (Scientist - D)



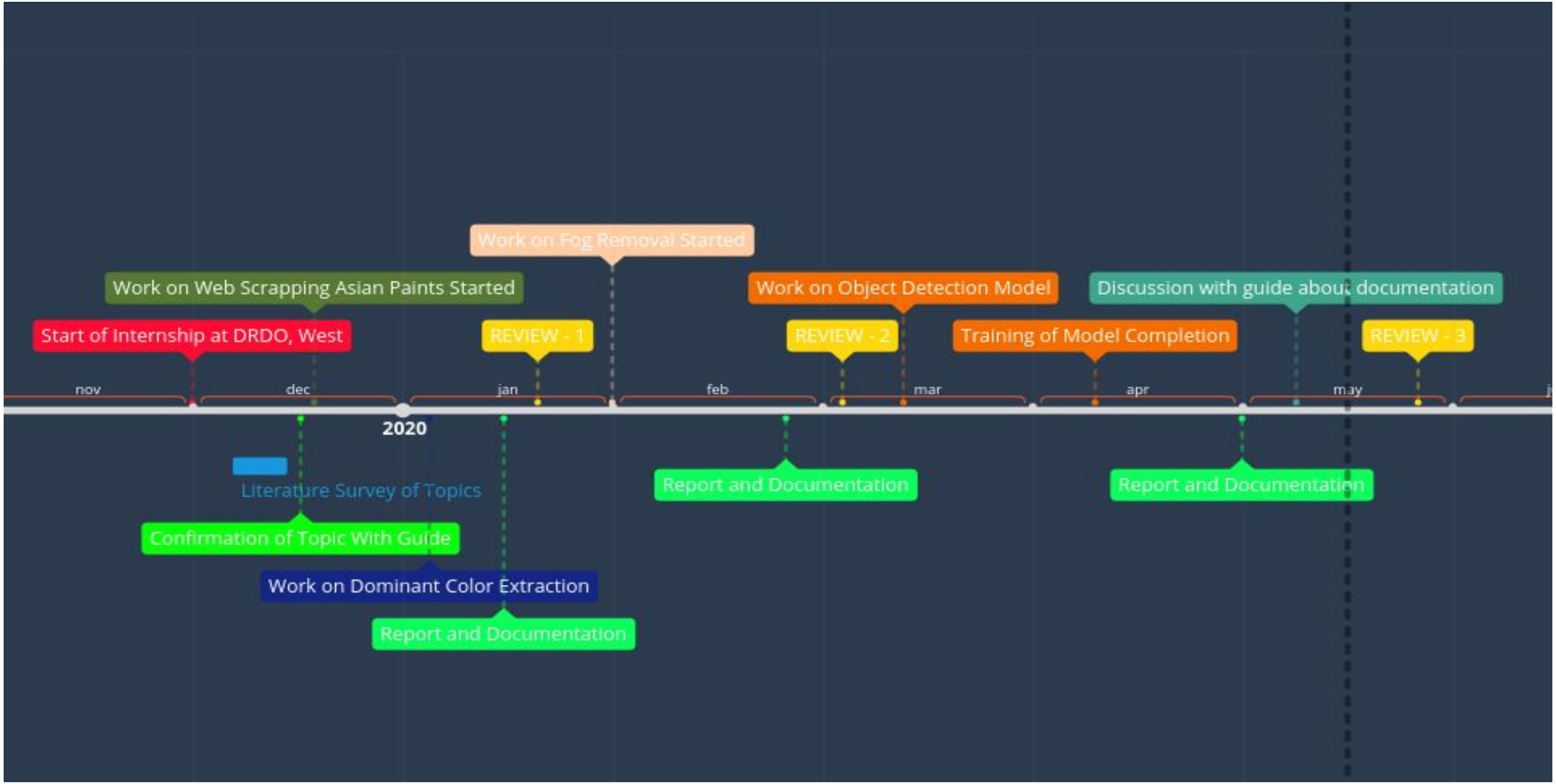
Motivation

While there is a lot of topographical data available for the DRDO by the Indian satellites but most are not usable due to the image quality and various weather and geographic conditions across India. These problems such as fog during winters or image distortion due to shadows in the image which decrease the efficiency of the computer vision algorithms can be reduced by employing certain image Pre-Processing pipelines which can help in improving the features of the image.

Aim of the Project

To create a model where there is Pre-Processing of the image done to improve features, these Pre-Processing are removal of fog, shadows and finding the dominant color to create image from basic color palette and further after the Preprocessing using the image to train the best convolutional neural network which detects the different vehicles across the different images with high level of confidence. While creating a model from scratch is not required only that find the best existing model which has least losses and trains faster on images.

Timeline



Work on Fog Removal Started
Work on Web Scrapping Asian Paints Started
Work on Object Detection Model
Discussion with guide about documentation
Start of Internship at DRDO, West
REVIEW - 1
REVIEW - 2
Training of Model Completion
REVIEW - 3
nov
dec
jan
feb
mar
apr
may
2020
Literature Survey of Topics
Report and Documentation
Report and Documentation
Confirmation of Topic With Guide
Work on Dominant Color Extraction
Report and Documentation

Literature Survey

1)Object Recognition in Images using Convolutional Neural Network
Authors - Mr.Sudharshan Duth P and Ms.Swathi Raj Dept. of Computer Science Amrita School of Arts and Sciences Mysuru, Karanataka, India

Advantages	Disadvantages
The training part ensures minimum human Interaction which helps in better results	The CNN is prone to high error rate and must be run for multiple epochs for better results which is computationally expensive and at same time still has some error rate.

Literature Survey

2)Deep Residual Learning for Image Recognition
Authors - Kaiming He Xiangyu Zhang Shaoqing Ren Jian Sun Microsoft Research

Advantages	Disadvantages
improves the accuracy of the model and eases the problem of vanishing and losing gradients suffered by CNN due to increase in depth of the deep neural networks.	This model is extremely compute intensive in natures which is bad for scaling.

Project Components

Parsing the Asian Paints Library for the entire color palette offered by them

Pre-processing the image for removal of fog

Find the dominant colors from the image using hierarchical quantization

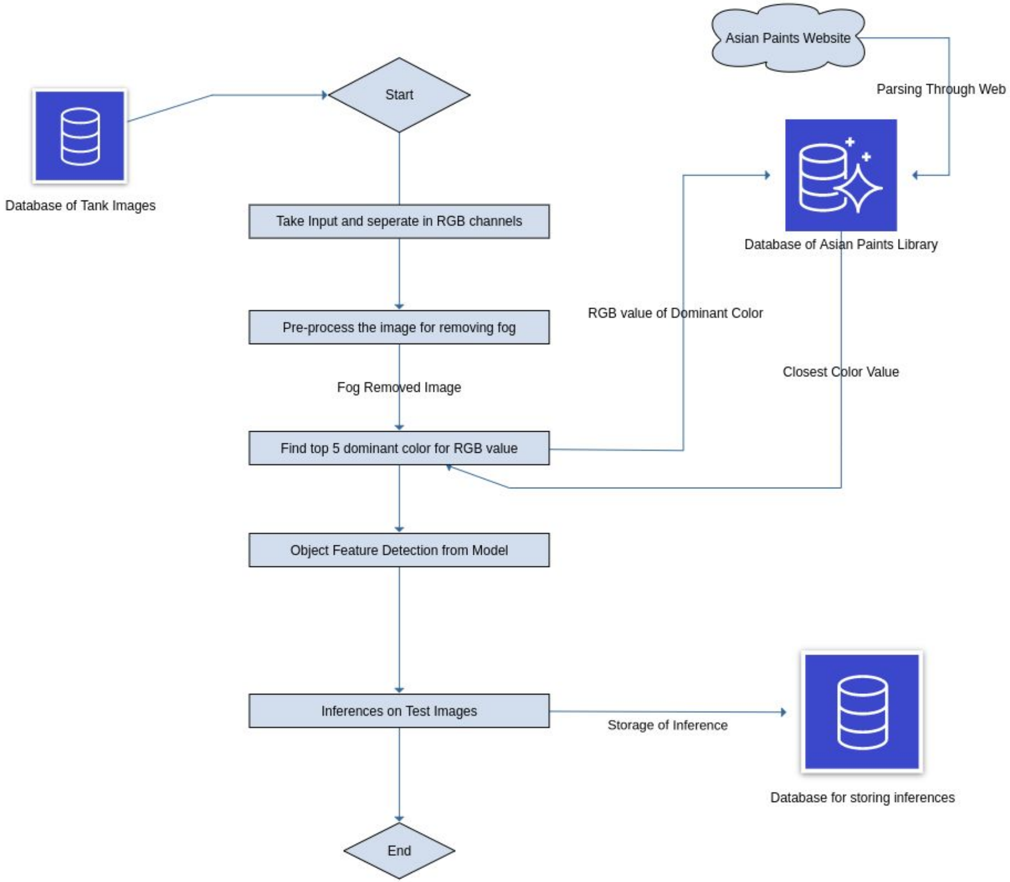
Creating a Object Detection model to get inference of objects in the image to classify them as armoured vehicles

Project Goals

- **Time Reduction** - Reducing the time taken to detect subjects across images, as a human takes much more time per image and also cannot process multiple images concurrently.
- **Reliability** - System should be able to detect the subjects at high accuracy and precision as it might form the part of defense measures deployed by the country.

Project Goals

- → **Scalability** - The system should be scalable in nature that it is able to identify different types of subjects should the requirements of the system change and can work much better if provided with parallel systems.
- → **Cost Reduction** - The system should be cost efficient in a manner that it should use the resources provided to the system as judiciously as possible.
- → **More efficient and accurate** - The system should provide results with high accuracy and precision compared to the existing models.


Asian Paints Website
Parsing Through Web
Start
Database of Tank Images
Take Input and seperate in RGB channels
Database of Asian Paints Library
RGB value of Dominant Color
Pre-process the image for removing fog
Closest Color Value
Fog Removed Image
Find top 5 dominant color for RGB value
Object Feature Detection from Model
Inferences on Test Images
Storage of Inference
Database for storing inferences
End

Overview

Parsing Color Palette Library

Parsing Color Pallete Library

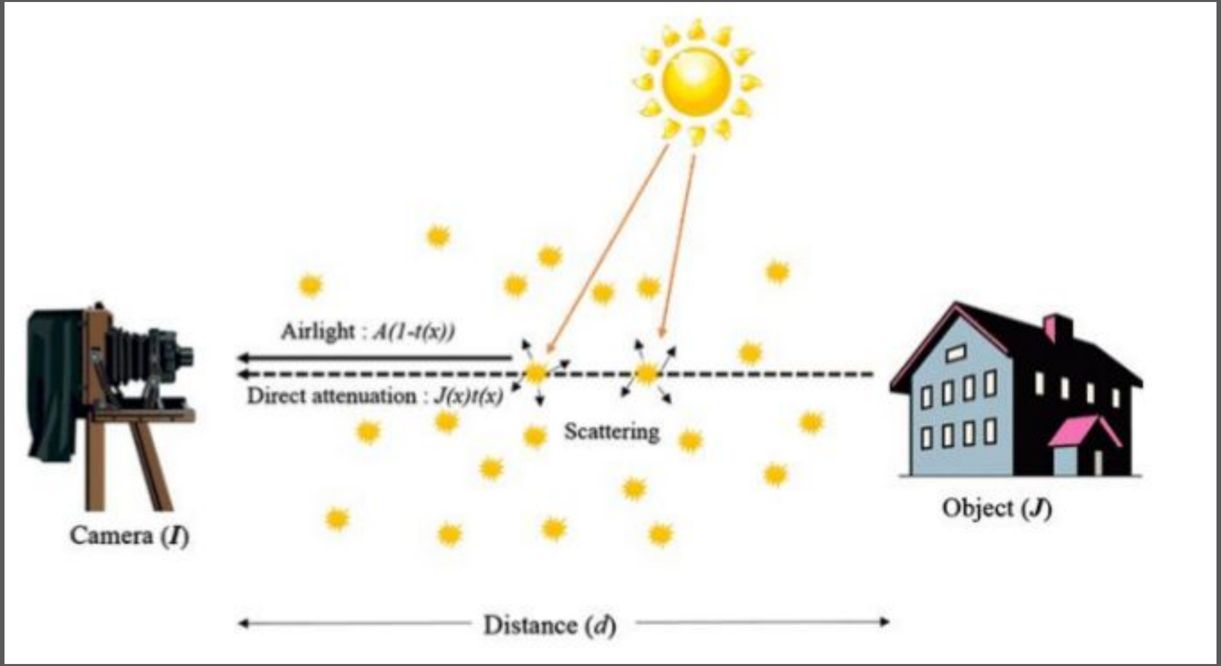
The parsing was done using python library, to get the following column fields there were 1800 values -

- ➔ Color URL
- ➔ Shade Code
- ➔ Shade Name
- ➔ Color Family
- ➔ Red Value
- ➔ Green Value
- ➔ Blue Value

Removal of Fog Based on Dark Channel Prior

This is a new method that is based on the acute observation of images in natural settings and based on the observation the de-hazing of the image is done which leads to much better, clearer images. It arrives from the point that most pictures which capture outdoor settings have intensity value of at least one of RGB channels to be around zero, within the boundaries of the local window being considered.

Removal of Fog



Removal of Fog

$$I(x) = J(x)e^{-\beta d(x)} + A\left(1 - e^{-\beta d(x)}\right)$$

$$J(x) = \frac{I(x) - A}{t(x)} + A.$$

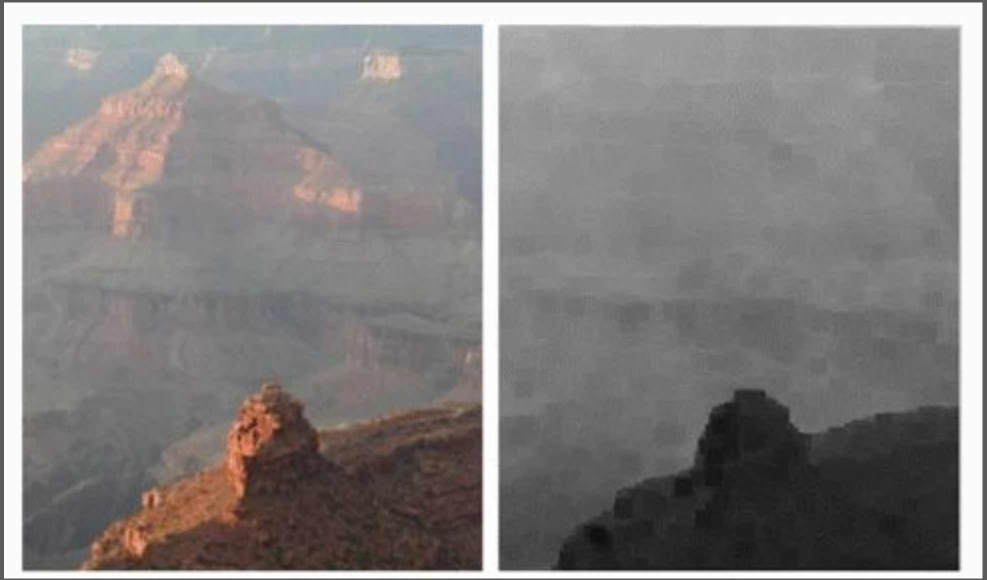
This is attributed due to the following reasons -

- Shadows, The shadows cast due to different objects in the natural world such as hills, trees etc.
- Extremely bright colors used by the different natural objects such as red rose, Dark colored animals etc.
- There are multiple dark objects and faces, such as cars, animals, rocky mountains etc.

Which confirms the observation that in all the images

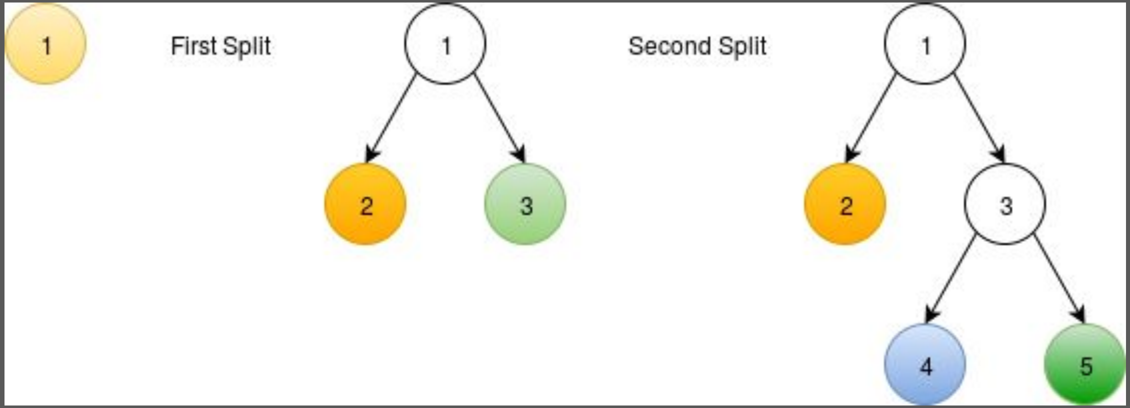
J^{dark} ---> 0

Also when this is compared with images that do contain the fog it causes the dark channel to give values which will have values far above zero. And this in turn helps us in discovering the density of haze at each point in the image.



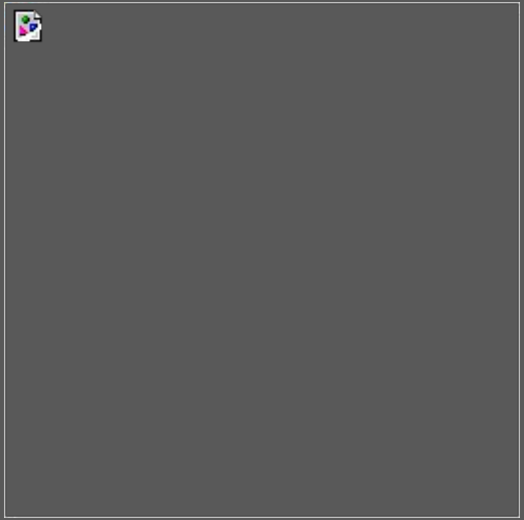
Dominant Color Extraction

The idea of hierarchical quantization is that every image contains a millions of colors but the displays on which they are projected only have 256 shades of RGB, so if we dial down the number of colors to 4-5 colors we will be able to get the dominant colors of the image.



Dominant Color Extraction

We now take the help of eigenvectors, these help us in identifying the plane where the covariance is maximum. But for our purposes we will be using a unit magnitude vector for direction and storing the magnitude separately as eigenvalues.



Residual Net Architecture

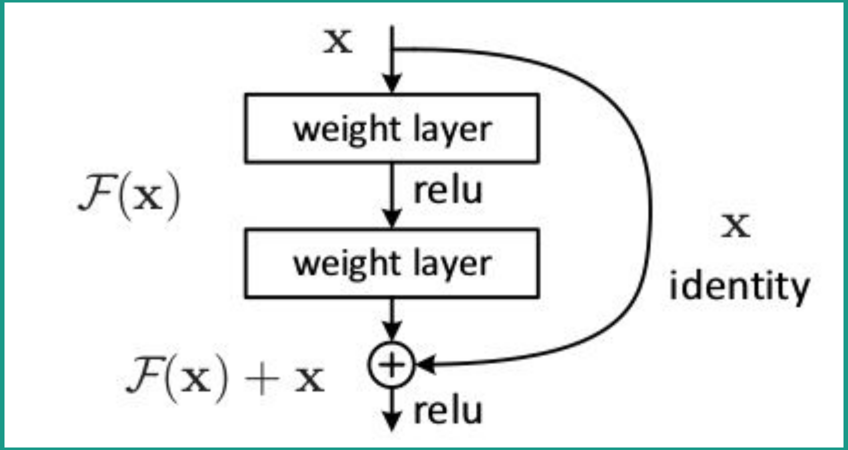
In the beginning when the research on deep neural networks was being conducted, there was a belief that as we increase the depth of the networks we will be able to gain more accuracy as there are more parameters to change and hence the model will get better at prediction. But as models such as VGG16 and others were stacked with more layers the model in actual reality started performing poorly as compared to models with less layers.

Residual Net Architecture

This was later found out that with the increase in number of layers other problems start creeping in, which causes the model to perform so badly that it does not work well even with the training data set.

To overcome this problem Kaiming He et al.

Came with the concept of skip connections



Residual Net Architecture

A [l] = x (Output of layer l)

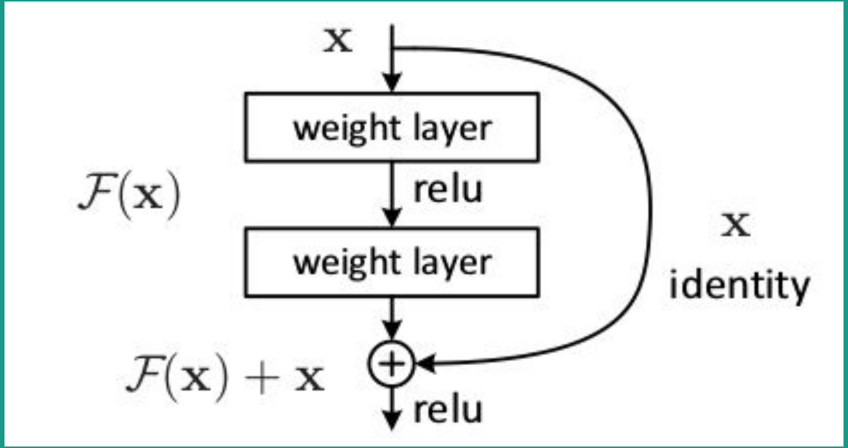
Z [l+1] = W [l+1] .A [l] + B [l+1]

A [l+1] = relu(Z [l+1])

Z [l+2] = W [l+2] .A [l+1] + B [l+2]

A [l+2] = relu(W [l+2] .A [l+1] + B [l+2] + A [l])

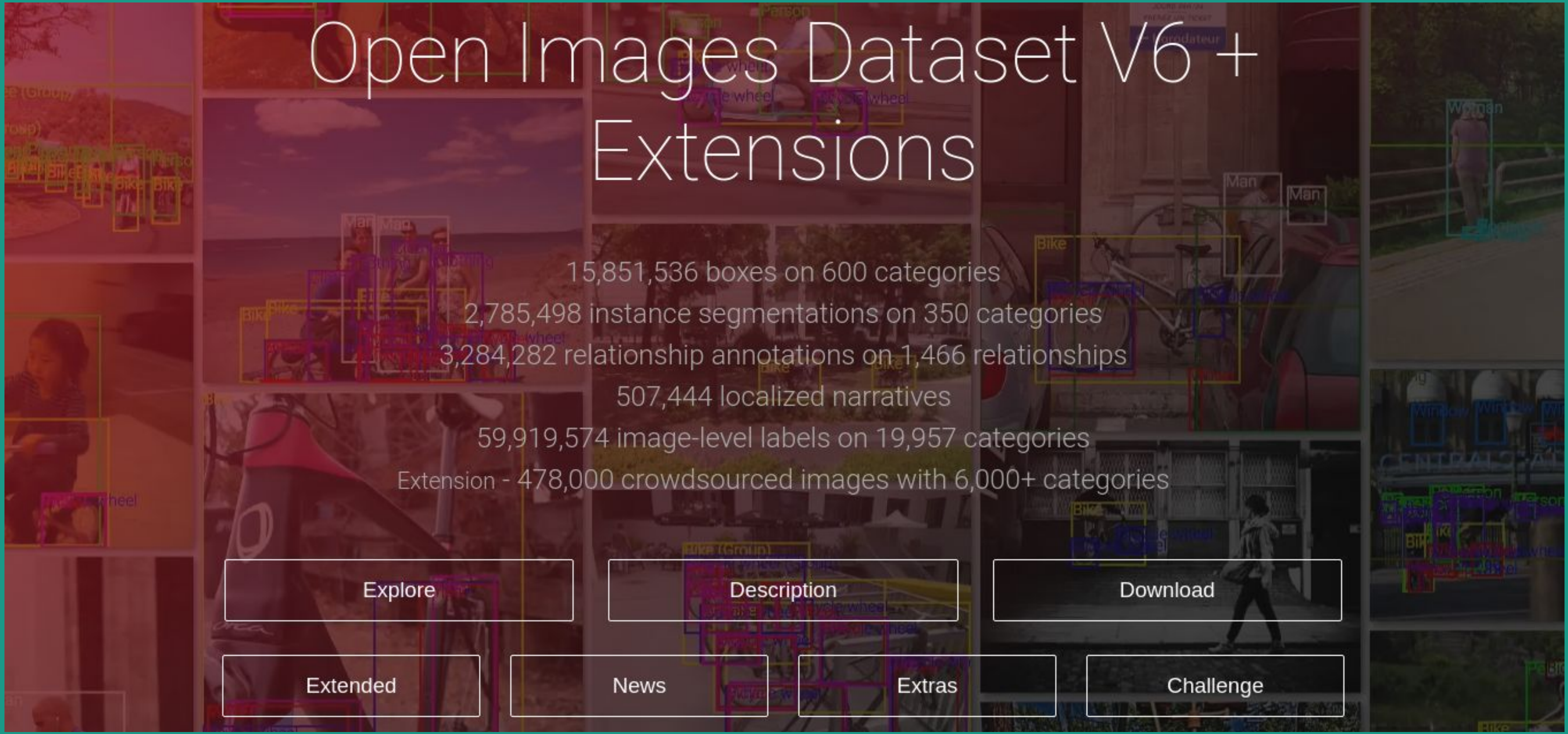
first layer pooling output to the third layer linear output and then take relu of that and pass that as input to the next layer.



Dataset

This is an open source database project created by google for use by everyone trying to find relevant dataset to train their models. This dataset currently contains in its latest version around 9 million images. These images also have their annotation with them that means they can be use directly to train the model without having the user to create labels for each image using different tools such as hyperlabel etc.

Dataset

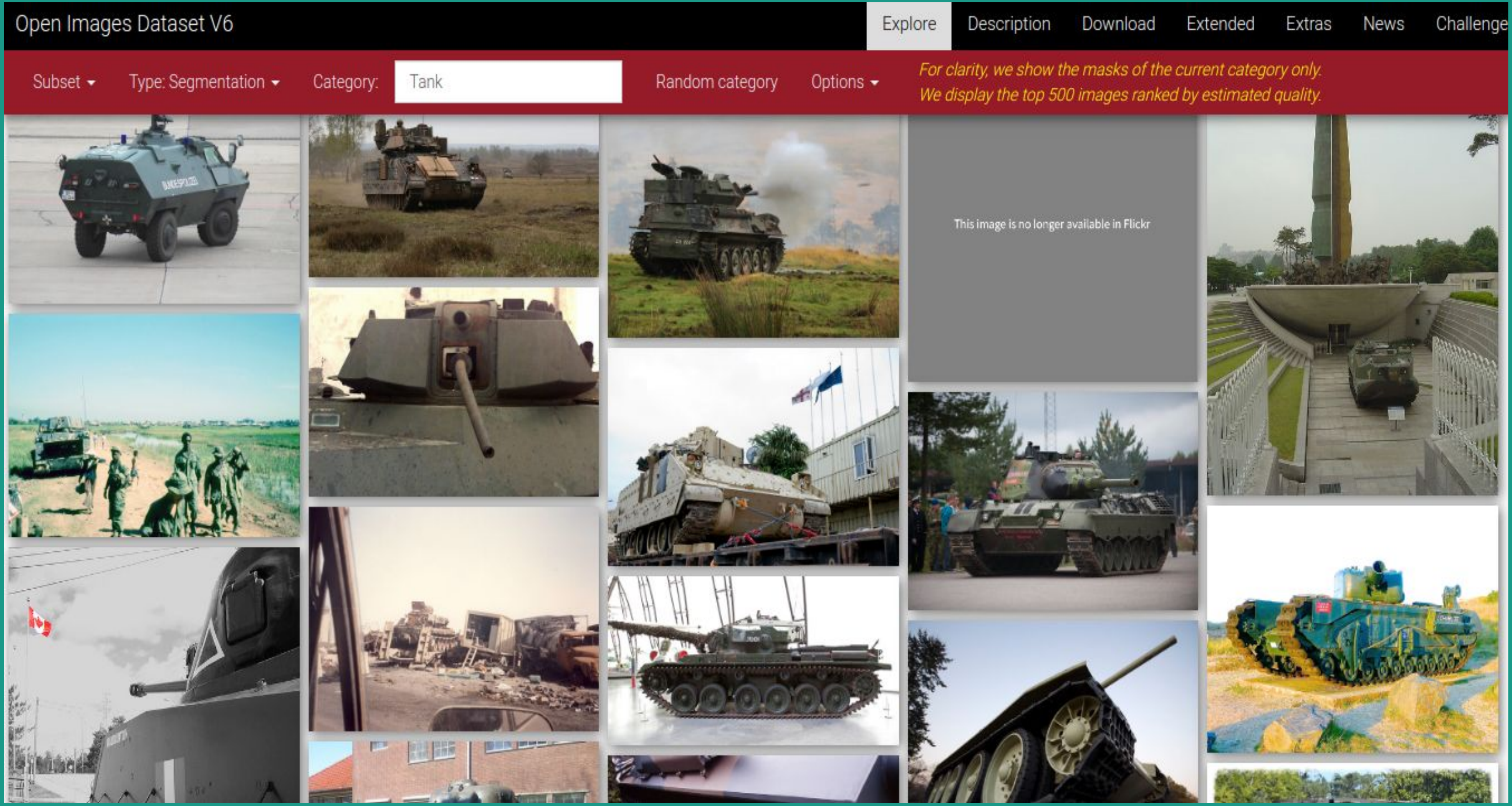

Open Images Dataset V6 + Extensions
15,851,536 boxes on 600 categories
2,785,498 instance segmentations on 350 categories
3,284,282 relationship annotations on 1,466 relationships
507,444 localized narratives
59,919,574 image-level labels on 19,957 categories
Extension - 478,000 crowdsourced images with 6,000+ categories
Explore
Description
Download
Extended
News
Extras
Challenge

Object Detection

	Train	Validation	Test	#Classes
Images	1,743,042	41,620	125,436	-
Boxes	14,610,229	204,621	625,282	600

Image Classification

	Train	Validation	Test	#Classes
Images	9,011,219	41,620	125,436	-
Machine-Generated Labels	78,977,695	512,093	1,545,835	7,870
Human-Verified Labels	27,894,289	551,390	1,667,399	19,794


Open Images Dataset V6
Explore
Description
Download
Extended
Extras
News
Challenge
Subset
Type: Segmentation
Category:
Tank
Random category
Options
For clarity, we show the masks of the current category only.
We display the top 500 images ranked by estimated quality.
This image is no longer available in Flickr

OIDv4_Toolkit

While there are millions of images and around 19000 classes on the dataset if somebody wants only a subset of the all images for their use there is no proper way to download the subset other than downloading the entire annotations and then comparing the classes of each image to get the desired results.

References

[1]Color Segmentation Using an Eigen Color Representation

Authors: Alaa E. Abdel-Hakim and Aly A. Farag Computer Vision and Image Processing Laboratory (CVIP Lab.) University of Louisville, Louisville/40292, KY USA

[2] Object Recognition in Images using Convolutional Neural Network

Authors: Mr.Sudharshan Duth P and Ms.Swathi Raj Dept. of Computer Science Amrita School of Arts and Sciences Mysuru, Karanataka, India

[3] towardsdatascience.com

[4] https://github.com/AtriSaxena/OIDv4_to_VOC.git

THANK YOU