Transfer Learning

Transfer Learning has become popular in the field of Image classification and natural language processing

Here we take a pre-trained model and we try to re-train it for the new problem.

I have used a mobilenet pre-trained model from Google's TensorFlow hub and I used that pre-trained model to classify our flowers dataset

Previously it used to take many epochs to train the complete model and achieve high accuracy.

If we use a pre-trained model it takes only 2-5 epochs to get a good accuracy.

Transfer Learning saves a lot of computation power.

Bcz many times these pre-trained models that we can get from TensorFlow hub, they are trained on millions of images

If u try to train that model on ur computer, it might take days or even months.

But all we are doing is we are taking that pre-trained model, getting all the weights and everything, and u kind of change only the last layer or last few layers for ur new problem

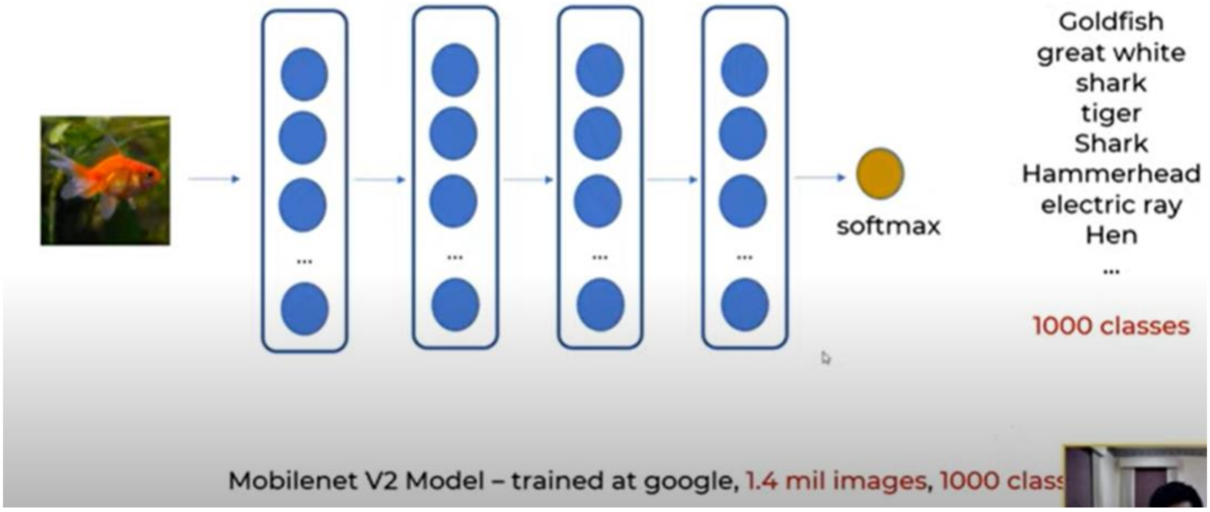
U can get super high accuracy with this kind of approach.

Wikipedia's definition of Transfer Learning:
Transfer learning is a research problem in Machine Learning that focuses on storing knowledge gained while solving one problem and applying it to a different but related problem.

For ex: Knowledge gained while learning to recognize cars could also apply when trying to recognize trucks.

Bcz the basic features for example the tires, the steering wheel and some of the components between cars and trucks will be still similar.

So, u can use this knowledge of this visual world to transfer that knowledge into solving a different problem.



We will take Google's trained mobilenet v2 model which is trained on 1.4 million images and total 1000 classes.

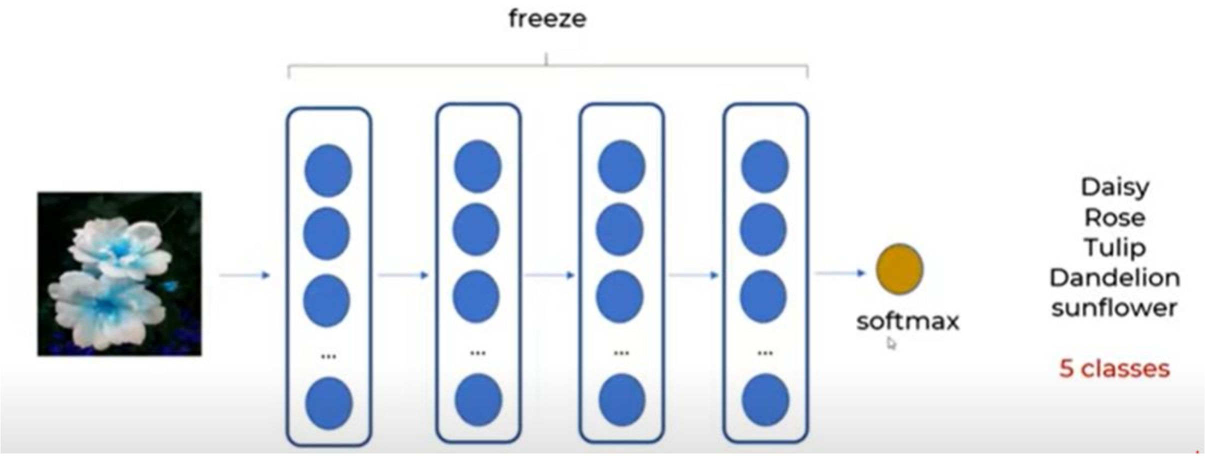
So this is a deep learning model that is trained at Google, it would have taken long time and lot of computational resources

U can see 1.4 million images is huge and the output is 1000 classes

These classes are diverse like we have goldfish, shark, some animals, some hammerhead, military uniform

When this model was trained it would have input layer, hidden layers in between and in the output layer we have SoftMax function, classifying into 1000 categories.

In transfer learning what happens is we freeze all the layers except the last layer.



We will freeze these layers and then we will use this model to classify flowers, which could be one of the five flower types shown here

So when we freeze these layers what happens is the model weights don’t change.

So when we perform training the weights in these frozen layers will not change.

So we will pass a flower during training phase, and using the weights we will get a feature vector and then we use SoftMax to classify into five classes instead of 1000.

So we are generating the feature vector using these frozen layers.

So during the training none of the weights changes.

Omitting the last layer is very common approach, in some approaches we only freeze 3 layers or 2 layers and remaining layers go through the usual neural network training.

We will write a python code to use mobile net v2 model and then use it to classify the flowers