Image Quality Determination using Artificial Neural Network

Syed Yaser Ahmed Syed

***Abstract*— Legacy image quality determination algorithms are highly depended on the metadata present in a picture. Information such as resolution are easy to manipulate; people can upload any low quality picture taken with a DSLR whose resolutions are usually higher than pictures from a camera phone. With this project, we are aiming to quantify a picture based on picture noise, saturation & blur. We will use this model to score only high quality pictures which will then give us a dataset with which we will train a Network to recognize good pictures. This will be useful for any web application that allows people to upload images, the neural network will help with instant quality determination.**

I. INTRODUCTION

What this project aims to achieve, is to build a reliable image scoring model based on information available in pictures. A good quality picture usually contains the following attributes:Good Contrast, Blur ratio, Noise ratio. Picture quality is usually subjective, but the above attributes are common to all pictures even if they were taken without a DSLR. We are following the same approach DxoMark does with smartphone cameras, but in a limited scale. Because some attributes in a DxoMark pertain more to the capability of a camera rather than the picture itself. Once the quality has been quantified using the score we will convert high quality pictures, which are hand selected, into a numerical dataset that contains the relevant information.

The DxOMark Score consists of three components: Color Depth for Portrait photography, Dynamic Range for Landscape photography Low-light ISO for Sports, wildlife and other action photography; thus covering all aspects of what is considered a good quality image.

Using this dataset we will then be able to train our neural network to create a pluggable image quality determination predictor that can quickly determine if an uploaded picture is good or not.

Note: This model will not be using EXIF data from a picture like the previous models but will extract data directly from an image.

II. DATASET

A. Brief introduction

Yelp dataset consists of all the information that has been submitted by yelp users in their reviews. There are three datasets to choose from - JSON, SQL and photos. For this project we will build a dataset of attributes using the photos part of the dataset.

B. Source

https://www.yelp.com/dataset/

III. QUANTIZING IMAGES

Quantizing images involves processing images using algorithms to get a discrete set of values from the incredible amount of data inherently present in an image. To label an image good or bad, we need to get the following information from the image.

A. Saturation

Saturation is the difference between brightness and colorfulness of a picture. In statistical terms, it is the combination of averages whose coefficients add up to zero, or the difference between two means. Beautiful pictures have medium saturation so that the object is in focus without artificial color boost.

B. Blur ratio

Blur is when an image cannot be seen clearly because of smear or poor focus of an object. To measure blur You simply take a single channel of an image and convolve it with the a 3x3 kernel, then take a variance of the response.

C. Noise ratio

Noise is usually high if the lens does not get enough light necessary to capture a picture. Noise usually looks like dots on a otherwise good picture and is easier for the human eye to find.

D. Luminance

Luminance is a photometric measure of the luminous intensity per unit area of light travelling in a given direction. It describes the amount of light that passes through, is emitted or reflected from a particular area, and falls within a given solid angle. It depends on the users preference, which it is fed to a neural network to train depending on users choice.

E. Descriptor

Image descriptors are descriptions of the visual features of the contents in images, videos, or algorithms or applications that produce such descriptions. They describe elementary characteristics such as the shape, the color, the texture or the motion, among others. Descriptors are helpful in training network depending on the type of photos that need to be classified. For example, we used the yelp photos dataset which contains pictures of food. Descriptor will help specify what type of pictures is currently being analyzed. In this case, pictures of food. So it will train the network depending on this.

IV. Intent

Using the quantization methods described above the photos dataset were reduced to a numerical dataset whose values are normalized by the source code. A label of either 0 or 1 was calculated based on subjective selection. This process is subjective but image quality can be easily judged just by looking at a picture. We will then train our network using the numerical values obtained. Backpropagation algorithm is best suited for training this model due to its high accuracy. Once trained, our system will be able to predict if any given image is a good picture or not.

A. Logic behind classification

First 3 features in both the figures are in the range of a preferred image, but they are still classified differently. This was the intention; since we have a descriptor that calculates the perfect saturation ration for a picture of food. This range is calculated by the neural network and is updated during training.

Preferred

Blur	SNR	Luminance	Saturation	Descriptors
15.9477087	1.04460159	0.38064286	60.0089878	0.00053007

Fig 1 - Example of a Preferred image and its features

Not Preferred

Blur	SNR	Luminance	Saturation	Descriptors
.47494987	2.4554994	0.6111536	55.050903	1.25E-05

Fig 2 - Example of a rejected image and its features

V. Experiments and Results

A. Finding best training algorithm

The dataset was trained on the following linear classifiers in folds of 5 over 100 epochs each:

- Logistic Regression
- Linear Discriminant Analysis
- K Nearest Neighbors
- Naive Bayes
- Support Vector Machine
- Learning Vector Quantization Version 1 and 2
- Backpropagation algorithm

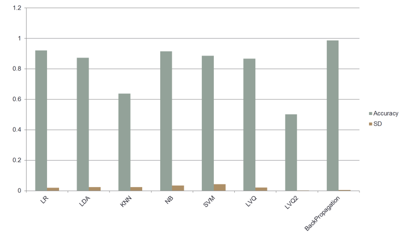

Fig 3 - Comparision of Algorithms

Backpropagation algorithm provides us with the most accurate prediction with the lowest standard deviation as evident from the image above.

B. Tuning Hyper parameters

The dataset was split into 5 folds and trained with 5 hidden layers over a loop of differing learning rates. The following table contains the result of our experiment. For future experiments a learning rate of 0.4 was chosen since it had the most potential with an accuracy of 99.2

Learning rate	Accuracy
0.1	0.424
0.2	0.981
0.3	0.992
0.4	0.992
0.5	0.983
0.6	0.991
0.7	0.98
0.8	0.987
0.9	0.978
1.0	0.981

Fig 4 - Accuracy of different learning rates

C. Result from training on Backpropagation

The next experiment was to check the mean square error of the backpropagation algorithm, so the dataset was again trained over a loop of 10 iterations. Fig 5 shows the results from this experiment. Its evident from the figure that the algorithm yields an average of 0.8 Mean square error.

Accuracy	Mean Square Error
0.98385	0.89595
0.99201	0.89505
0.99067	0.8969
0.99217	0.8951
0.98884	0.8804
0.98284	0.88975
0.99099	0.886
0.98595	0.8983
0.97784	0.87585
0.98511	0.8818

Fig 5 - Accuracy and Mean square error

Fig 6 shows the distribution of MSE over 100 epochs. It clearly shows the error rate decrease exponentially, going down from 82.656 in the 0th epoch to 8.989 in the 100th epoch.

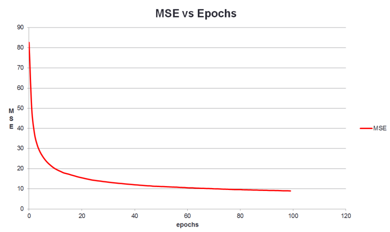

Fig 6 - Mean square error over 100 epochs

VI. Future Improvements

In the current architecture there is a layer between Image collection and the classifier. This layer is where the features are extracted from the images and a new dataset is created from the values. Using Convolutional Neural Network, we can remove this layer and extract the required features directly from the images.

We can use CNN model to classify images based on their subject which will give us a descriptor. We can classify images as preferable or non-preferable by calculating the range of required features based on the descriptor obtained from the CNN model.

Using CNN will also help in reducing the size of the dataset since real world users do not have access to a large collection of images over which the algorithm can be trained to be subjective. Descriptor can be shared over applications since it does not have to be subjective it reduces our reliance on feature extraction layer.

VII. Conclusion

Using this model we were able to train our dataset and get an average accuracy of 99%. It was able to correctly predict whether a new image will be preferred or rejected based on the training set which was curated according to my personal taste. Backpropagation algorithm with a learning rate of 0.4 and 100 epochs was able to achieve such a high accuracy and mean error rate of $\tilde{8}$.

References

[1] Yelp Dataset. (2017). Yelp.com. Retrieved 13 October 2017, from https://www.yelp.com/dataset/challenge

[2] Predicting with a Neural Network — Kaggle. (2017). Kaggle.com. Retrieved 13 October 2017, from https://www.kaggle.com/colladete/predicting-with-a-neural-network

[3] Backpropagation algorithm (2017). Retrieved 14 October 2017, from https://machinelearningmastery.com