April 03, 2019

PIC2PRINT - DEEP LEARNING BASED ON IMAGE RECOMMENDER FOR PRINTING

HP INC

Pic2Print – Deep Learning Based Image Recommender for Printing

Abstract

There are more than 2.5 billion smartphone users in the world but only 3% of the total prints are coming from mobile. U.S. photo printing market alone in 2016 was estimated at $10.1 billion. We have developed a solution named Pic2Print, which uses machine learning to build a model which automatically identifies and recommends the best pictures to print from a user's smartphone feed.

Problem statement

There is a misconception that the photo printing market is declining. Ongoing market analysis of both the consumer retail and professional market segments strongly indicates that it is, in fact, growing. More than a trillion pictures were captured using smartphones in 2018. Since smartphone users capture a lot of images, finding the best picture for printing is not easy. If a solution could automatically recommend which of the captured pictures are ideal for printing based on the aesthetic qualities of the images and the user's past behavior and, it will make the printing process easier for the customer and would significantly increase mobile printing.

Our solution

Promoting mobile photo printing and making the process simpler for the user is what our solution aims for. Sometimes among hundreds of images, it becomes difficult to choose one image which would be a perfect candidate for printing. For scoring the images, we have used a deep learning algorithm called Neural Image Assessment (NIMA), which computes a score based on aesthetic quality. Based on the score calculated by the NIMA module, our solution recommends a photograph with the highest score which the user could print. Also, our solution aims to increase the aesthetic score of the photograph by tuning a few parameters like brightness, noise, contrast etc. such that user gets an enhanced picture to print. Based on the picture selected by the user i.e. from the original or the fixed image, our model will relearn so that each user gets his own customized model. Fig below shows the flow of our solution:

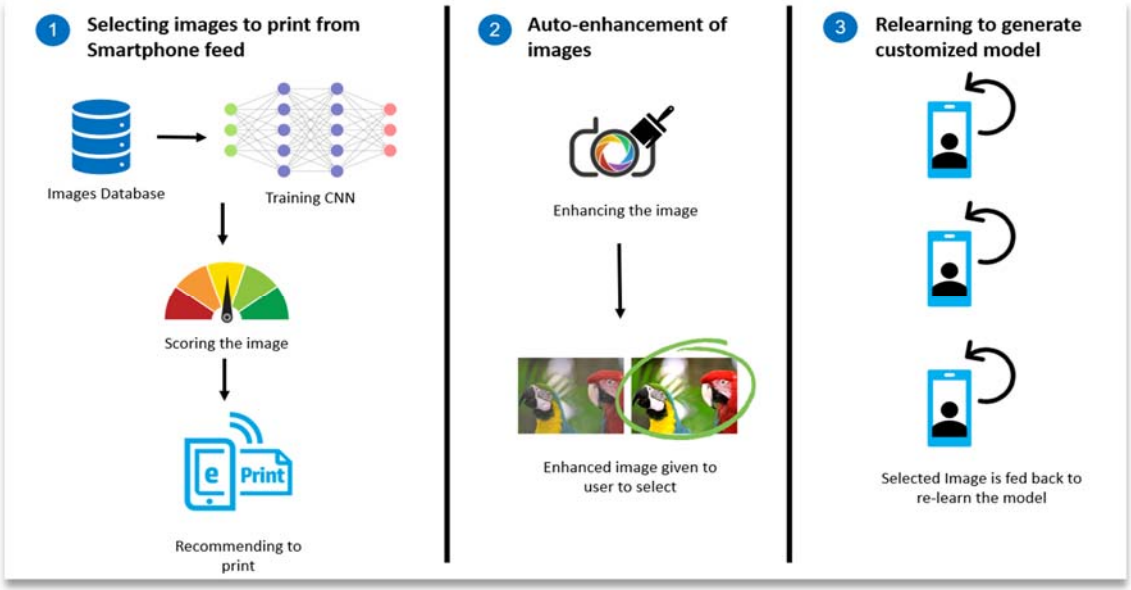

Figure 1 Details of our solution

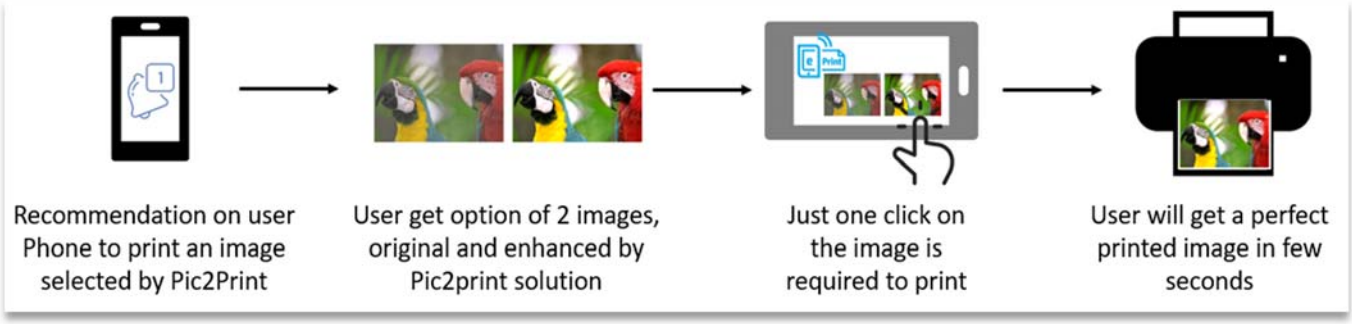

Figure 2 Our solution from the user's perspective

Input Data:

We have used the Atomic Visual Action (AVA) dataset as the input to train our model. It contains around 255,000 images which were rated by photographers. The images are annotated with a score on the scale of 1-10 where 10 is the highest aesthetic score for the image. The aesthetic quality of the image can be represented by the mean score and standard deviation.

Model Training:

We used the MobileNet algorithm which is basically for mobile and embedded vision applications. It is an efficient deep convolutional neural network which is based on depthwise separable convolutions. The last layer in the network was replaced by fully-connect layer, followed by soft-max activation. Figure on the right shows the model architecture.

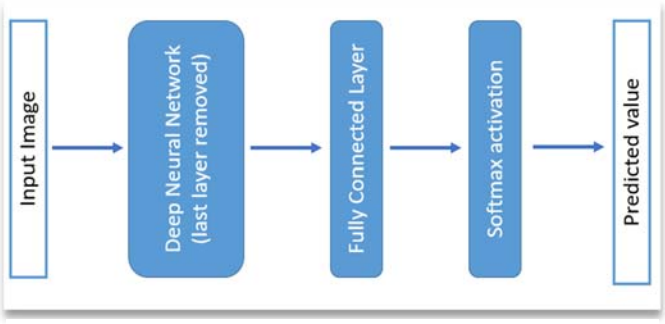

Figure 3 Model Architecture

The model predicts the distribution of ratings for a given image. Distribution of the rated images can be expressed as a probability mass function $|p| = [p_{s1}, p_{s2}, \ldots\ldots, p_{sn}]$ where $s_1 = 1$ and $s_n = 10$. The probability of a quality score can be represented by $\sum_{i=1}^{n} p_{si} = 1$. Given the distribution of the rating as p,

Mean Quality Score $\mu = \sum_{i=1}^{n} s_i \times p_{s_i}$

Standard Deviation Score $\sigma = \left(\sum_{i=1}^{n} (s_i - \mu)^2 \times p_{s_i}\right)^{1/2}$

Recommendation and Enhancement of the Image:

Once we have calculated the score, picture with the highest score will be recommended to the user to print. When the user opens the app to print the image, we provide the option to enhance the image i.e. we will try to increase the score of the image. Some parameters such as denoising the image, contrast enhancement etc. can be tuned to improve the perceptual quality of the picture. The multi-layer Laplacian technique was used to improve the local and global contrast of the image, which controls the quantity of detail, brightness, and shadow of the picture. Also, Turbo denoising technique was used to clean the picture, which uses the domain transform as its core filter.
After increasing the score of the picture, it is recommended to the user along with the original image. The user can choose any image he likes and print that. Here, the user can automatically get the best image with the increased quality without putting any effort which will encourage the user to print more images.

Relearning:

Definition of a good picture varies from user to user, for example, one user can like brighter picture other can like images with high contrast. So, we plan to put the model on the user phone which will relearn based on the image user has chosen to print. By doing this every user can have their own customized image recommender.

Evidence the solution works

Below are some examples of smartphone captured pics passed through our model and their scores.

Mean Score – 6.61

Mean Score - 6.03

Mean Score– 5.13

Also, we were able to enhance the aesthetic quality of the picture using the method explained above.

Mean - 4.83, SD - ±6.16

Mean – 6.69, SD - ±0.29

Competitive approaches

There are features in solutions like Google Photos which creates an assortment of the user's pictures. But these assortments are usually a selection from the user's picture feed without any knowledge of the user's preferences or aesthetic quality. We couldn't find any solution which automatically finds and recommends the best user captured images for printing.

Disclosed by Pavleen Brar, Shameed Sait and Madhusoodhana Rao, HP Inc.