UDC 004.4

M. Fedoriaka, K. Melkumian

HYBRID METHOD OF IMAGE PROCESSING BASED ON CONVOLUTIONAL NEURAL NETWORKS

Abstract: The article describes the model of a convolutional neural network designed to improve the resolution of images on mobile devices. The structure of the network differs from the classical approaches by using a combination of neural network processing and one of the more traditional image processing algorithms. The article lists the advantages of using the aforementioned system.

Keywords: neural networks, upscaling, image processing, mobile devices.

Formulation of the problem

Nowadays, mobile photography is becoming more and more popular. Many people choose their smartphone as their main device for taking photos, because it is much more convenient, faster and cheaper than a specialized camera. Unfortunately, high resolution and photo quality are only available to buyers of expensive smartphones. That is why the problem of improving the resolution and sharpness of photos is incredibly relevant.

In the last 4-5 years computing power of smartphones has grown really high. Also, some companies started adding dedicated neural cores to their systems on a chip (SoC) used in smartphones. Apple pioneered this idea in 2017 and since then more and more devices have been shipping with neural cores and features utilizing AI and Machine Learning.

However, the computing power of smartphones, as well as dedicated sub processors designed for machine learning, are still underutilized in mobile applications. In particular, the problem of upscaling (improving the resolution) of photos is still largely unresolved.

To solve this problem, a hybrid method of image processing is proposed, which uses both a convolutional neural network and a classical algorithm for post-processing.

Analysis of recent research

The most famous and successful application with similar functionality is Adobe Lightroom CC. The results of the upscaling function are satisfactory, but far from the results shown by the algorithms used on personal computers. Also, the algorithms of this application are proprietary, so it is impossible to know their features and even the category they belong to.

Due to smartphones being a relatively new category of product and neural cores being introduced in last 4-5 years, there is not much research on the topic of using AI and Machine Learning for mobile applications in general, let alone mobile photography specifically. A lot of scientific work has been done on algorithms used in personal computers, so they will be considered as examples of existing solutions.

Traditional algorithms without the use of machine learning show good results and do not require a large amount of time to select the data sets needed to train the neural network, and, in fact, the training process itself is not needed.

The most common of these algorithms is the Image Sharpening Algorithm Using Unsharp Masking which sharpens the image by blurring to increase the contrast. The basic concept of blurry masking is to first blur the original image, then subtract the blurred image from the original image and, as a final step, add differences to the original image. The linear blur filtering approach is used to enhance a noisy image with a high-pass filter. Blur masks are very suitable for sharpening images. But sharpening too much can also cause the image to lose its natural appearance. This method has two main disadvantages, for example, the contrast in the darker area is enhanced much deeper than in the lightest area. The next problem is that this method also amplifies the effects of noise and digitization. Due to these problems, the image in most cases loses its originality. Of all the sharpening approaches used, this UM approach is the simplest and easiest.

One of the variations of the basic UM scheme is called Adaptive UM. This approach contains an adaptive filter that is used to enhance images whose dynamic range is the same as the one of the CRT monitors. The reason for using the adaptive filter is to avoid overshoot effects and blown out contrast details, emphasizing more on medium contrast details. This approach has the benefit of avoiding noise amplification and blown out details in areas of high contrast.

Another modification of this algorithm uses a depth buffer. This approach is used to improve the perception of images that contain information about depth. As with UM (Unsharp Masking), the difference between the output content of the depth buffer and the copy filtered by the lower frequencies is used to determine information about spatially important areas of the scene. Depending on this data, contrast, color, and other image quality factors improve. The purpose of this approach is to improve the perception of complex scenes by introducing additional depth signals. This allows you to use this algorithm for a variety of scenes, from complex landscapes to documents and mass processing of photos and videos.

The use of wavelet transform allows to further improve the quality of the original result. The main idea of sharpening the image this way is to add a version of the input signal with a high-pass filter to the input signal. Wavelet ratios provide high-frequency image details with multiple resolutions. Using this concept, a wavelet-based approach was used in [2] to sharpen the image. Image sharpening is introduced to improve the contrast and brightness of the image.

Another method uses a combination of both discrete wavelet transform (HAAR) and unsharp masking technique [3]. Wavelet coefficients provide high-frequency image coefficients, such as edge information. The sharpening technique is used here to increase the sharpness of the image. This algorithm uses a correlation between different wavelet coefficients, and high-frequency coefficients are considered image edges.

Final UM variation that is going to be analyzed is Rational Unsharp Masking. This is the Linear UM modified with a few changes to make results look much better. A control term is expressed as a rational function of the input data. The noise amplification and noise sensitivity which are the weakest points of traditional Unsharp Masking can be eliminated because of this parameter. Also, the overblowing of the edges can be reduced as well.

Summarizing the UM algorithms, for this application the variation with the depth buffer would work the best because depth information is easily available on modern smartphones and out of all approaches for getting edge information on the image this one is the easiest for compute on a mobile phone and provides very good results. Second choice would be Wavelet transform variation of UM because of great results, even though the processing costs are higher.

Let's move on to the methods of machine learning. Most deep-learning upscaling models are trained using generative adversarial networks (GANs). One of the limitations of GANs is that they are essentially a lazy approach because their loss function, the critic, learns as part of the process and not specifically for that purpose. Many deep learning methods cannot be applied universally to all types of images, and almost all of them have their drawbacks. For example, a model trained in ultra-high-resolution animals may not be suitable for ultra-high resolution human faces.

An innovative approach to solving this problem is convolutional neural networks, or CNN (Convolutional Neural Networks).

In [4], the authors proposed the architecture of the PanNet neural network, which is used to improve the quality of satellite images. This architecture is based on the well-known ResNet architecture, which is mostly used to recognize objects in a photo, but can be modified to improve image quality.

The authors of [5] outlined another option for using CNN for image processing. It should be noted here that they use an unusual architecture utilizing residual blocks. Other existing image blur reduction studies combine data representation and network blur, but it takes time to extract features from an image, which is why the authors of this article use residual blocks in the encoder for faster convergence and better results.

Hybrid method of image processing based on convolutional neural networks

To solve the problem of image sharpening on mobile devices, it is proposed to use a hybrid method of image processing based on convolutional neural networks. This type of neural network is the most suitable for this task.

Each image is represented as a grid of pixels. The value of each pixel is in the range from 0 to 255, which indicates its brightness.

The proposed model uses an automatic encoder that learns from sharp images by extracting features. First, we train the model by transmitting many clear images to the encoder. Upon completion of the training, it takes the blurred image as input and produces clean functions. Image then enters the decoder to generate a clear image.

First, the input images are taken by the encoder and converted into an array of pixel values. It processes the input image of size h * w and generates an image. To save the original size, we use the padding attribute. It represents the input image in low-dimensional space to restore the blurred image. Then we take all the data extracted from the encoder, after it removes unwanted noise and apply corrections, and then transmit that to the decoder as input. The decoder then takes the corrected data and converts it back to the original image size to obtain the desired sharp image.

After processing by the neural network, the Unsharp Masking algorithm with depth buffer is used to improve the contrast and brightness of the resulting image.

You can see the result of the image processing below. “Before” image is on the left and “after” image is on the right (Fig. 1).

Figure 1.

Even though original image resolution is quite low (300x300 pixels in this case), sharpening effect is very strong. For processing higher resolution images with good sharpening effects, neural network needs to be trained on sharper images to begin with, and that’s what would be even more ideal for smartphone applications, since most smartphones shoot images in resolutions close to 4000x3000 pixels.

Also, it’s worth noting that iOS and Android use different frameworks for ML. iOS uses native Core ML and can use cross-platform ML Kit developed by Google. Android on the other hand can only use ML Kit out of these too. Of course, there are other solutions available, but these are most popular ones. There is no major performance difference you can achieve by going with either one of them, but for most applications ML Kit would be preferable because of its cross-platform nature.

Conclusion

The problem of improving the image quality on mobile devices is important because of the prevalence of smartphones on the market. Mobile photography will only gain popularity, so the problem of photo processing on mobile devices will not lose relevance. The aforementioned approach using a hybrid of convolutional neural networks with traditional post-processing algorithms is very suitable for use on mobile devices due to the simplification of the architecture, which will be show it’s strength when used in less powerful devices, such as smartphones.

REFERENCES

1. Archana J.N. A Review on the Image Sharpening Algorithms Using Unsharp Masking / Archana J.N, Aishwarya .P. // International Journal of Engineering Science and Computing,. – 2016, vol. 6, no. 7, pp. 8729 – 8733, ISSN 2321 3361.

2. Ying L. A wavelet based image sharpening algorithm / L. Ying, N.T Ming, L.B Keat // International Conference on Computer Science and Software Engineering (CSSE 2008). – Wuhen, Hubei, China, 2008. – pp. 1053–1056, DOI: 10.1109/CSSE.2008.1631.

3. Salonika Kansaall. Image Sharpening using Unsharp Masking and Wavelet Transform / Salonika Kansaall, Gurpreet Kaur. // International Journal of Advance Research in Computer Science and Management Studies. – 2014. – no.2, ISSN 2321 7782.

4. Yang J. PanNet: A deep network architecture for pan-sharpening / Junfeng Yang, Xueyang Fu, Yuwen Hu та ін.]. // IEEE International Conference on Computer Vision. – 2017. – pp. 1753–1761, DOI: 10.1109/ICCV.2017.193.

5. Shiva Shankar R. A novel approach for sharpening blur image using convolutional neural networks / R. Shiva Shankar, G Mahesh, K V S S Murthy. // Journal of Critical Reviews. – 2020. – vol. 7, ed. 7. – pp. 139–148, ISSN 2394 5125.