Edge detection for Roof Images using Transfer Learning

1st Aneeqa Ahmed
Computer Engineering
Jeju National University
Jeju, South Korea
aneeqaahmed339@jejunu.ac.kr

2nd Yung-Cheol Byun
Computer Engineering
Jeju National University
Jeju, South Korea
ycb@jejunu.ac.kr

3rd Debapriya Hazra
Computer Engineering
Jeju National University
Jeju, South Korea
debapriyah@gmail.com

***Abstract*— Edge Detection in image processing is very important due to large number of applications it offers in variety of fields that extend from medical imaging to text and object detection, security, mapping of roads, real time traffic management, image inpainting, video surveillance and many more. Traditional methods for edge detection mostly rely on gradient filter based algorithms which usually require excessive pre-processing of the images for noise reduction and post-processing of the generated results in order to get fine edges. Moreover, traditional algorithms are not reliable generally because as the noise in images increases their efficiency is affected largely due to increase of mask size which also makes the system computationally expensive. In this paper, we will employ CNN method to detect edges of roof images. Incorporating CNN into edge detection problem makes the whole system simple, fast, and reliable. Moreover, with no more extra training and without additional feature extraction CNN can process input images of any size. This technique employs feature map of the image using Visual Geometry Group (VGG) CNN network followed by application of Roberts, Prewitt, Scharr and Sobel edge operators separately to compute required edges. Interpretations of ground truths were obtained using manual techniques on roof images and for performance comparison, PNSR value of computed results via multiple operators against the ground truths is calculated.**

Keywords—Edge detection, deep learning, convolutional neural network, roof images.

I. INTRODUCTION (*HEADING 1*)

Edges in images can be defined as a boundary between the object and the background which can be located by detecting the discontinuities in image brightness known as edge detection problem in image processing [1]. Edge detection is considered as one of the important phenomenon for extracting useful information from images [2] and it is broadly applied in computer vision and image processing domains. Mohamed A. El-Sayedet.al [3] stated that crucial information in a scene is mostly hidden in its edges, thus edge detection has been a subject of vast interest to the scientists and there have been widespread research efforts in developing a good and reliable edge detection algorithm from many years. However, edges in images can be formed in variety of ways depending upon object's geometry and they may vary in shapes and sizes [3, 4]. Also, efficiency of edge detection algorithms largely depends upon the illumination conditions and noise level of the images. To devise a standard edge detection algorithm that is universally accepted is not an easy task. Traditional approaches for edge detection include gradient or difference based filters, non-maximum suppression or detection of zero crossings by Laplacian of Gaussian etc. [2, 5, 6, 7].

Traditional methods for edge detection are complex and computationally expensive as they mainly depend on computing several image features [1]. Moreover, with the increase of noise level or change of illumination conditions of images the efficiency of traditional algorithms degrades most of the times which make them highly unreliable [5]. Furthermore, many type of filters that form the basis for traditional edge detection algorithms can only detect edges in particular directions, hence it demands use of multiple filters together to detect edges in every direction. These limitations of traditional algorithms motivated researchers to incorporate machine learning techniques into edge detection problem.

In this paper, we will make use of convolutional layers of VGG CNN network [8] for computing the image feature map. Employing transfer learning from CNN to solve edge detection problem is inspired from the work of Eppel [9]. Basically, CNN is very simple and fast as it does not require any extra feature processing to detect the edges [1]. Moreover, with CNN the task of deciding and extracting the concerned features is not to be programmed manually rather CNN itself does the job and it is more efficient than traditional manual approaches. Also, unlike traditional methods CNN is strong enough to learn directly from the images and due to these capabilities, CNN can process any input image irrespective of its size without the need of extra training [1]. Besides, it is very simple to integrate CNN into other computer vision and image processing systems so by incorporating CNN into edge detection problem, people can easily gel in this network to other applications that may require edge detection as a first step for some disease diagnosis or for image segmentation etc. [1].

II. VGG MODEL

Deep Neural Networks (DNNs) have emerged as a noticeable approach in the field of computer vision and intelligent computing [10]. Researchers and engineers are motivated to design more powerful and accurate algorithms because of their increasing utilities and emerging applicability in variety of tasks. It is the architecture of neural networks that determines the accuracy of the algorithms as well it defines resource utilization. Therefore, the success of these algorithms mainly depends on the careful designing of neural network architecture.

VGG network as shown in Figure.1 is a deep learning (DL) model which was launched in 2014 by Simonyan and Zisserman in 2014 ImageNet Large Scale Visual Recognition Challenge (ILSVRC) [8]. The architecture of VGG network depicted that network depth is very crucial and important parameter in order to achieve better performance and accuracy. VGG network is well-known for its simple architecture in view point of depth which increases progressively by adding new convolution layers using merely 3x3 convolution filters in all layers. This network comprises of 16 convolution layers as shown in Fig. 1 and has been trained on 4 GPUs for 2-3 weeks. VGG network has achieved remarkable accuracy in 2014 ILSVRC and team VGG network has secured top two positions in the localization and classification tasks. Along with that, VGG network is fundamentally applicable to other image processing and computer vision applications as a standard feature extractor. Weight configurations of VGG network are publically available and presently researchers are widely using this network to extract feature from images with a greater accuracy as compared to previous DL models.

The rest of the paper is organized as follows: Section II describes the proposed system for edge detection. Section III presents experimental results and gives performance comparison. Section IV summarizes conclusion with a summary of future goals.

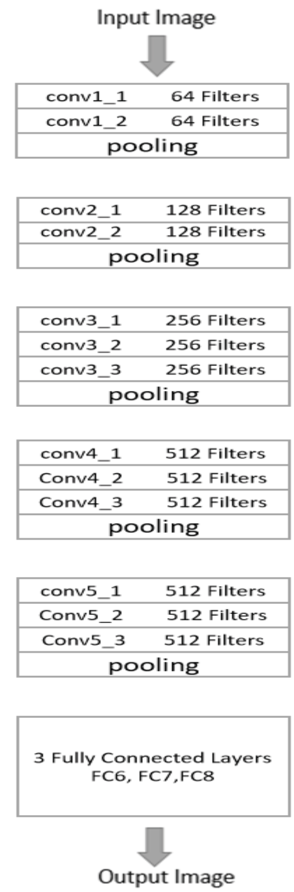


Fig. 1. VGG Network Architecure

III. EDGE DETECTION METHOD

In this section, we have briefly discussed about the dataset that we have used in this paper, which is followed by detailed explanation of the technique being employed in this paper. Broadly, the methodology comprises of four important steps to detect the edges which are feature map computation via convolution neural network, gradient map computation, averaging all filters gradient maps, averaging all layers' gradient maps.

A. Dataset

In this paper, we have used dataset of roof images as seen from satellite for the purpose of edge detection. The dataset used is downloaded from a renowned web repository of datasets [11] and is publically available. It contains 42760 images of roof tops that include various types and shapes of roofs like cylindrical, square, rectangle, pointed etc. The dataset also comprises of roof images of varying illumination conditions and resolutions. Using this dataset allowed us to test our system with roof top images of variable illumination conditions with different kinds of roof shapes which increased system robustness.

B. Methodology

The available dataset only composed of roof images but it lacked the ground truths for edge detection. We have employed manual techniques to generate the ground truths from our datasets for few images only and these ground truths have been used during performance analysis this method. However to generate manual ground truths for all 42760 images is a tedious task. Ground truths are mandatory requirement if we want to specifically train a CNN for direct edge detection. Keeping this in view, we have incorporated an indirect method that is CNN transfer learning methodology to achieve our goal. Transfer learning allows us to use a model that is trained for one task to be reused for another allied task. We have made use of a pre-trained strong convolutional neural network namely VGG network and computed all the required features automatically from VGG network. Later this automatically computed strong feature map is being fed to Roberts, Prewitt, Scharr and Sobel edge detectors separately and gradients maps for every filter in each layer of VGG network is computed which are ultimately averaged to give the total gradient maps for each layer of the network. The overall methodology is being explained in the following steps:

- At the first step feature map of the input image is computed. To determine the required feature map, this approach employs the convolution layers of VGG network.
- Subsequently, Roberts, Prewitt, Scharr and Sobel operators are applied separately to compute gradient map for each filter in a given convolution layer.
- Then the average of gradient maps of all filters in a given layer is computed which gives the total gradient map for the specific layer.
- At this point, overall average gradient map is computed for all layers of network in order to get the finalized edge map of the image. Finally, the edge map is being resized to original input image size to visualize the resultant edges.

Fig. 2 shows the block diagram of the system.

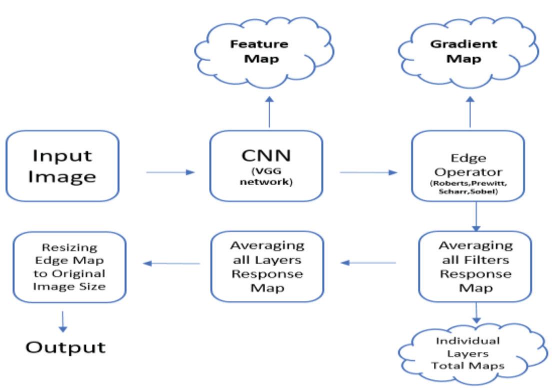


Fig. 2. Block Diagram of Edge Detection System

In order to investigate the response and of each layer of the network to the final outcome we have made few observations. After comparing the "step d-all layers averaged result" to the individual gradient maps of every layer in "step c" we have observed the final result for every type of image is quite degraded as compared to the manual ground truths. The reason of this type of outcome is because edge information is mainly extracted in lower layers and higher layers mainly focus towards learning of high level features like recognizing the shapes etc. The individual total gradient maps for the specific layer in "step c" have shown very promising results for edge detection in the lower layers of the network, while results gradually degrade as we move towards the higher layers of the network.

All these investigations of individual response of each layer in this approach lead us to slightly modify our methodology. Instead of computing "overall average gradient map for all layers in step d", we just picked total gradient map of all filters in layer 1 of VGG network as our final outcome for every type of image and then we resized this outcome to original input image size to visualize the resultant edges. Figure.3 shows the upgraded methodology that we followed in this work.

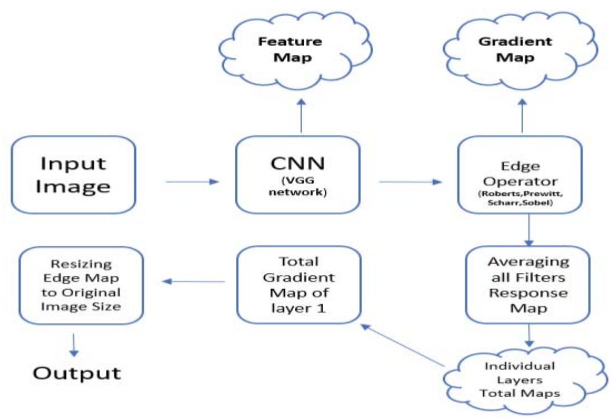


Fig. 3 Upgraded Block diagram of Edge Detection System

IV. Experimental Results

We have evaluated our CNN based edge detection system on the dataset of roof images provided by kaggle [11]. The results of different stages corresponding to Figure.3 are being shown in this section. At the first step, the image is being input to the system and feature map is computed using VGG CNN network.

After the computation of feature map, Roberts, Prewitt, Scharr and Sobel operators are being employed separately to determine the gradient map of every filter in each layer of VGG network. Figure.4 displays resultant outcomes of edges for only few filters in different layers of the network with Sobel edge detector due to limitation of space. The first image in every column is the original image, followed by grey scale outcome and the last one is the resultant image in binary level.

As there are many filters in each layer of VGG network as shown already in Fig. 1, so in order to get the total response map of an entire layer, gradient maps of all filters within a layer are averaged. Figure. 5 shows the complete edge maps for different layers.

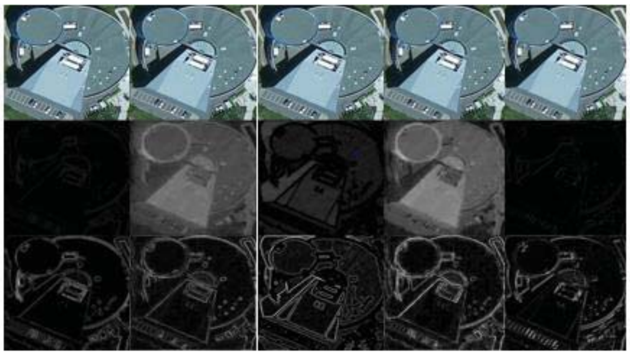


Fig 4. Gradient Map at Filter Level for different layers using Sobel operator

From Fig. 5 it can be observed easily that gradient maps of high level layers are much degraded as we discussed earlier in section 2 while discussing the individual response of each layer, so when we took the average of gradient maps for all the layers to generate a single outcome we got very degraded results Fig. 6 (a) shows the overall Averaged Gradient Map of all layers with Sobel Operator and Fig. 6 (b) shows the overall Averaged Gradient Map of only first 4 layers with Sobel operator. From Fig. 5 (a), (b) we have judged that if we only consider the lowest layer that is layer 1 outcome in the final results of edge detection, we can improve our results.

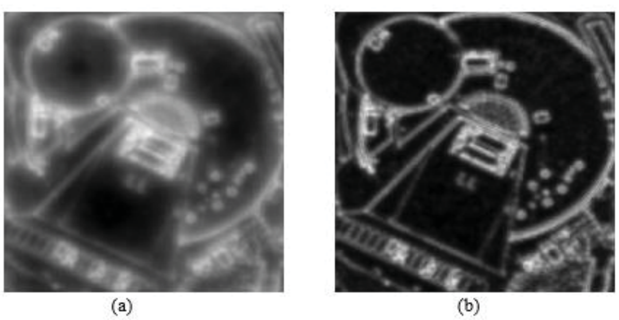


Fig 6. (a) Overall Averaged Gradient Map of all layers with Sobel Operator, (b) Overall Averaged Gradient Map of only first 4 layers with Sobel operator

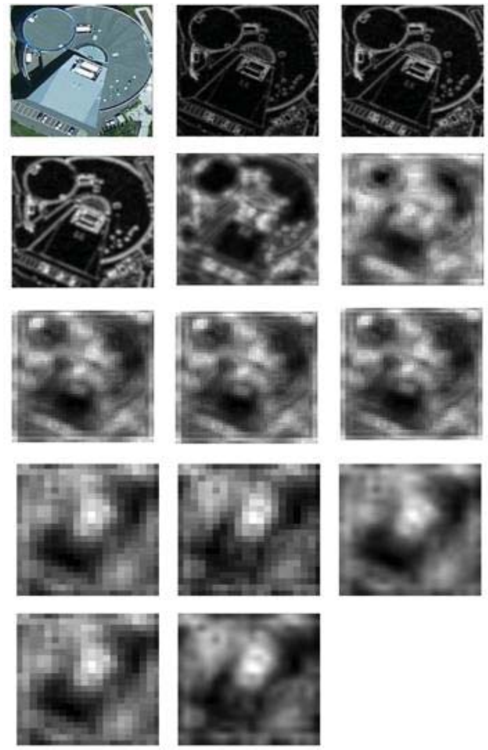


Fig. 5 Individual Gradient Maps for all layers using Sobel operator
1st Row (L to R): Input Image, Individual layer response of conv 1_1, Individual layer response of conv 1_2
2nd Row (L to R): Individual layer response of conv 2_1, Individual layer response of conv2_2, Individual layer response of conv 3_1
3rd Row (L to R): Individual layer response of conv 3_2, Individual layer response of conv3_3, Individual layer response of conv 4_1
4th Row (L to R): Individual layer response of conv 4_2, Individual layer response of conv4_3, Individual layer response of conv 5_1
5th Row (L to R Individual layer response of conv 5_2, Individual layer response of conv5_3

So, following the upgraded methodology as depicted in Fig. 3 we picked the total gradient map of layer 1 of VGG network as our resultant outcome for edges which is finally resized to the original size of the input image. Fig. 7 shows the final outcome of our edge detection system with Sobel operator.

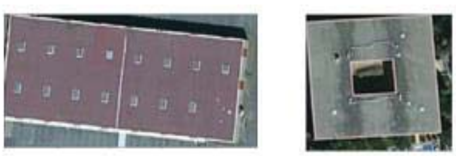
Fig 7. Final Outcome VGG followed by Sobel

Similarly, Fig. 8 shows output of edge detection system with Roberts, Prewitt and Scharr operator for the same input image.

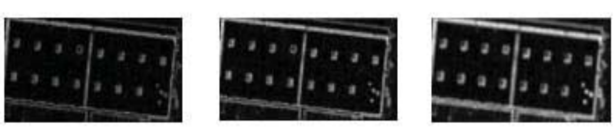
(a) (b) (c)

Fig. 8(a) Final Outcome-VGG followed by Roberts, (b) Final Outcome-VGG followed by Prewitt, (c) Final Outcome-VGG followed by Scharr

In Fig. 9 we have shown total gradient maps of first three layers with Prewitt and Sobel operators using this methodology for different type of roofs, also these images are of variable size and different illumination conditions. From Fig. 9 it can be observed that this technique performs equally good for images of any size and even with bad resolution, which is one of the major advantages of introducing CNN into edge detection problem over traditional methods.



(a) Input Test Images



(b) Gradient Maps of first 3 layers with Prewitt Operator

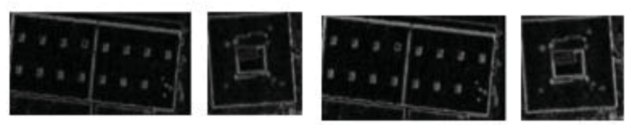
(c) Gradient Maps of first 3 layers with Sobel Operator

(d) Gradient Maps of first 3 layers with Prewitt Operator

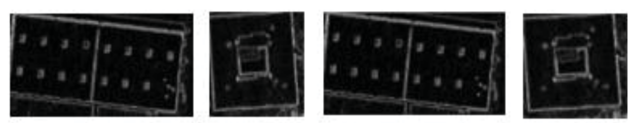
(e) Gradient Maps of first 3 layers with Sobel Operator

Fig. 9 Depiction of Gradient maps of first three layers for various test image

Likewise, Figure. 10 shows final outcomes with Prewitt, Roberts, Scharr and Sobel operators for different size of roof images.



(a) Final outcome with Robert's operator (b) Final outcome with Prewitt operator



(c) Final outcome with Scharr operator (d) Final outcome with Sobel operator

Fig. 10 Final outcome of test images with different operators for variable size images

To further check the effectiveness of this technique a comparison of this technique is done with the method that computes the edges without employing VGG network. Figure. 11 shows the caparison results for Robets and Sobel operator in both cases. This paper chose these two operators to show comparison since Roberts is the simplest operator and Sobel is the most sophisticated one among the used operators. The comparison for the rest two operators will fall in the comparison limit of these two operators.

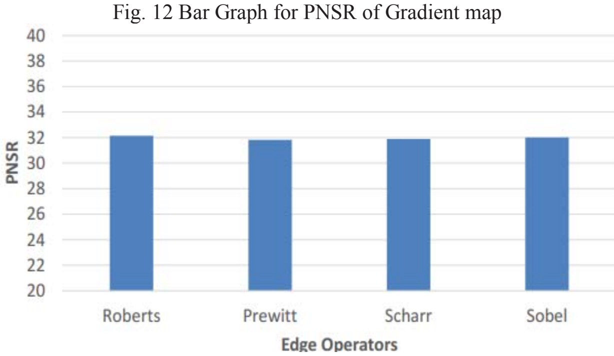
(a) Outcome with Roberts without VGG (b) Outcome with Roberts with VGG

(c) Outcome with Sobel without VGG (d) Outcome with Sobel with VGG

Fig. 11 Comparison Figure with and without VGG application

To evaluate the performance of the edge detection technique PSNR (peak signal-to-noise ratio) is being employed. To compute the PNSR value, manual methods have been employed to generate the corresponding ground truths for roof edges. Figure. 12 shows a bar graph of different PNSR values for the outcomes of the gradient maps computed via Roberts, Prewitt, Scharr and Sobel operator on the feature map of VGG network separately.

Fig. 12 Bar Graph for PNSR of Gradient map



PNSR values for every case are very close to each other. We have experimented with different size and different resolution of images and computed PNSR values but still we got comparable results that are minor difference in PNSR values. This observation made it clear that basically this technique is independent of the edge operator and its performance solely depends upon the computed feature map of the VGG CNN network. Basically, it is the feature map of CNN network that plays the vital role in deciding the final outcome of the proposed technique.

CONCLUSION

A powerful algorithm is presented for the detection of edges in images using Convolution Neural Networks. This technique can detect edges of roof tops irrespective of the shape of the roof or size of the image. CNN method is employed in this paper to detect the edges because unlike traditional approaches for edge detection its efficiency does not degrade with the increase of noise in images. Traditional edge detection approaches are computationally expensive and non-reliable in case of noise variant image dataset. CNN, on the other hand is very simple, does not require manual feature extraction which makes it fast and it can process images of variable size without any additional training. We have employed VGG convolutional neural network for the computation of features which is followed by the implementation of edge detectors to find out the gradient map of the image. The outcomes of this technique depicted better performance for every type of roof images irrespective of size, location and type of roof as compared to classical methods.

In future, we are focused to improve this technique by training another CNN network using ground truths and would compare the performance accuracy of that technique with current method that used pre-trained CNN model. Succeeding goal would be the implementation of Generative Adversarial Nets (GANs) for edge restoration in images in order to detect the edges more efficiently. Edge restoration application will be extremely beneficial prior to edge detection as it would immensely increase the performance of the system by the restoration of any missed edges in roof images due to presence of a tree, shadow or any obstacle etc.

ACKNOWLEDGEMENT

This research was financially supported by the Ministry of SMEs and Startups(MSS), Korea, under the “Regional Specialized Industry Development Program(R&D or non-R&D, Project number)” supervised by the Korea Institute for Advancement of Technology (KIAT).

REFERENCES

[1] Ruohui Wang, “Edge Detection Using Convolutional Neural Network”, International Symposium on Neural Networks, Advances in Neural Networks – ISNN, 20 Dec. 2016.

[2] T. Peli and D. Malah, “A study of edge detection algorithms”, IEEE transaction on Computer Graphics and Image Processing. vol. 20, no. 1, pp. 1–21, Sept. 1982.

[3] Mohamed A. El-Sayed,Yarub A. Estaitia, Mohamed A. Khafagy , “Automated Edge Detection Using Convolutional Neural Network”, (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 4, No. 10, 2013.

[4] Vincent Torre, Tomaso A. Poggio, "On Edge Detection", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol . PAMI-8, No. 2, MARCH 1986.

[5] D.Marr, E.Hildreth, "Theory of Edge Detection", Proc. Of Royal Society Landon, B(207):187-217,1980.

[6] L. Canny, "A computational approach to edge detection", IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 8 no. 1, pp. 679-698, 1986.

[7] F. Rosenblatt, "The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain", Cornell Aeronautical Laboratory, Psychological Review, v65, No. 6, pp. 386-408, 1958.

[8] Karen Simonyan, Andrew Zisserman' "Very Deep Convolutional Networks for Large-Scale Image Recognition", International Conference on Learning Representations (ICLR) 2015.

[9] https://github.com/sagieppel/Using-sobel-edge-detector-on-activation-map-of-convolutional-neural networks/commits?author=sagieppel,

[10] Alom, Md. Zahangir & M. Taha, Tarek & Yakopcic, Christopher & Westberg, Stefan & Hasan, Mahmudul & C Van Esesn, Brian & Awwal, Abdul & Asari, Vijayan., "The History Began from AlexNet: A Comprehensive Survey on Deep Learning Approaches", Mrach 2018.

[11] https://www.kaggle.com/canonwu/roof-images