Vision Anomaly Detection Using Self-Gated Rectified Linear Unit

Israt Jahan[1,2], Md. Osman Ali[1,3], Md. Habibur Rahman[1], and Yeong Min Jang[1]
[1]Department of Electronics Engineering, Kookmin University, Seoul 02707, South Korea
[2]Department of Electrical and Electronic Engineering, Daffodil International University, Dhaka 1341, Bangladesh
[3]Deptment of Electrical and Electronic Engineering, Noakhali Science and Technology University, Noakhali 3814, Bangladesh
Email: israt@kookmin.ac.kr; osman@kookmin.ac.kr; rahman.habibur@ieee.org; yjang@kookmin.ac.kr

***Abstract*—In the area of image processing and computer vision, visual anomaly detection is a critical and difficult task. For anomaly detection in surface image data, a customized neural network incorporating self-gated rectified linear unit (SGReLU) was designed, and the SGReLU-based model excelled other activation function-based models with a top-20 average test accuracy of 99.84%. The computational time needed for the operation is 10533 s for 20 epochs and the top-20 average test loss is 0.0125 using SGReLU, both of them were comparatively less than other activation functions.**

***Index Terms*—vision anomaly detection, self-gated rectified linear unit, computer vision.**

I. Introduction

Visual anomaly detection, often known as anomaly detection in images is essential in terms of both theoretical and empirical work [1], [2]. When it comes to recognizing visual anomalies, deep learning networks outperform machine learning algorithms. To detect surface irregularities or fissures, the Keras functional API is utilized to generate sophisticated models in a flexible pattern.

Activation function, an essential part of the neural network, has a vital role in image processing. Different activation functions such as rectified linear unit (ReLU) [3], [4], Leaky ReLU (LReLU) [5], swish [6], scaled exponential linear unit (SELU) [7], exponential linear unit (ELU) [8] and self-gated rectified linear unit (SGReLU) [9] are mainly used in the dense layers of neural networks. Previously SGReLU was used mainly in the case of image classification. In this paper, SGReLU is used for vision anomaly detection and the comparison with other activation functions in the case of surface anomaly detection has been also investigated.

The contributions of this paper are as follows:

• To design a customized neural network with convolution layers, maxpooling layers, dense layers, dropout layers, and activation layers.

• To compare the performance of different activation functions used in the activation layer of the customized neural network.

• The SGReLU function outperforms other activation functions in the case of top-20 average test accuracy.

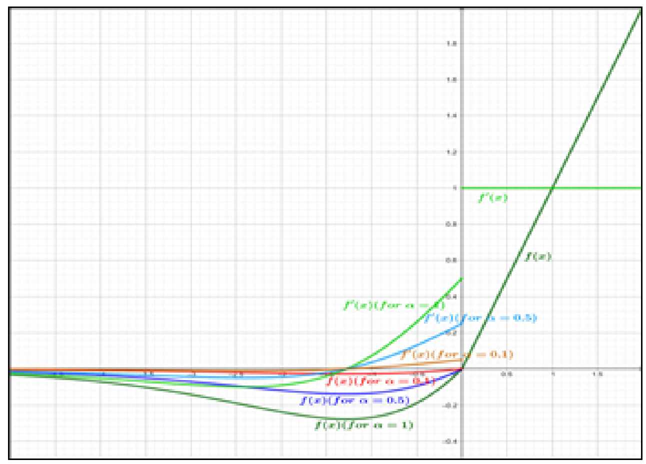

Fig. 1: SGReLU function and its first derivative for different values of α.

• The SGReLU function has also acheived better performance in terms of test loss and computational time in comparison to other activation functions.

The following sections of this paper are arranged as follows. Section II presents the methodology including the SGReLU function, dataset preparation, and model architecture. In Section III, the obtained results are discussed. Finally, the conclusion, as well as future work, are described in Section IV.

II. Methodology

A. SGReLU Function

The generic nature of activation functions can be expressed as a bivariate function, b(x,g(x)), where x is the unfiltered preactivation, that is the input to the final bivariate function. SGReLU function is a non-monotonic continuous function that maintains the identical binary format. It can be expressed as follows:

$$f\left(x\right)=\begin{cases} x, & x\geq 0 \\ \alpha.x.\sigma\left(x\right), & x<0 \end{cases} \tag{1}$$

where $\sigma(x)$ is sigmoid function, which equals to $\frac{1}{1+e^{-x}}$, and α is a hyper-parameter that ranges from 0.1 to 1 in magnitude.

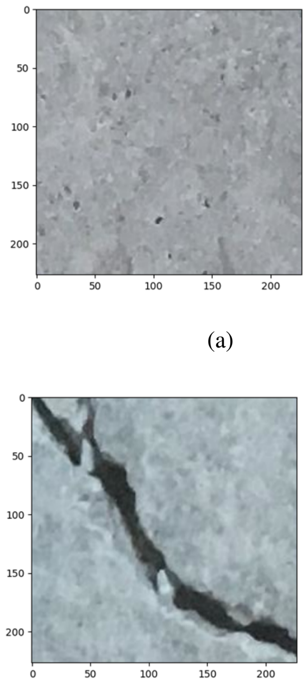

Fig. 2: Surface crack detection (a) normal condition and (b) anomaly condition.

In the positive zone, SGReLU behaves like ReLU in terms of unbounded linear nature. As a result, it solves the issue of saturation as well as has a non-zero gradient, which speeds up the learning process. The absence of the sigmoid function, in contrast to swish, reduces the need for power operations, leading to a faster processing period. The self-gating approach, conversely, is used to make the negative area of the function adaptable with a single input. Consequently, with a huge value of preactivation, x, a little negative bump develops at the start of the negative area that tends to zero. SGReLU provides a negative bump for small negative inputs, which eliminates the neuron death issue. It has negative parts, however, unlike LReLU and PReLU, it is confined below since it tends to zero for huge values of preactivation, providing sparsity in the structure and lowering computational time. As a result, the overfitting issue is eliminated, and the network achieves noise-resilience at the same time. As a result, the network achieves not only non-monotonicity but also self-regularization, making it immune to neuronal death as well as data compatibility. The first derivative of SGReLU function is as follows:

$$f'(x) = \begin{cases} 1, & x \geq 0 \\ \alpha \left[\frac{1}{1+e^{-x}} + \frac{xe^{-x}}{(1+e^{-x})^2} \right], & x < 0 \end{cases} \tag{2}$$

Fig. 1 shows SGReLU function and its derivative for different values of α.

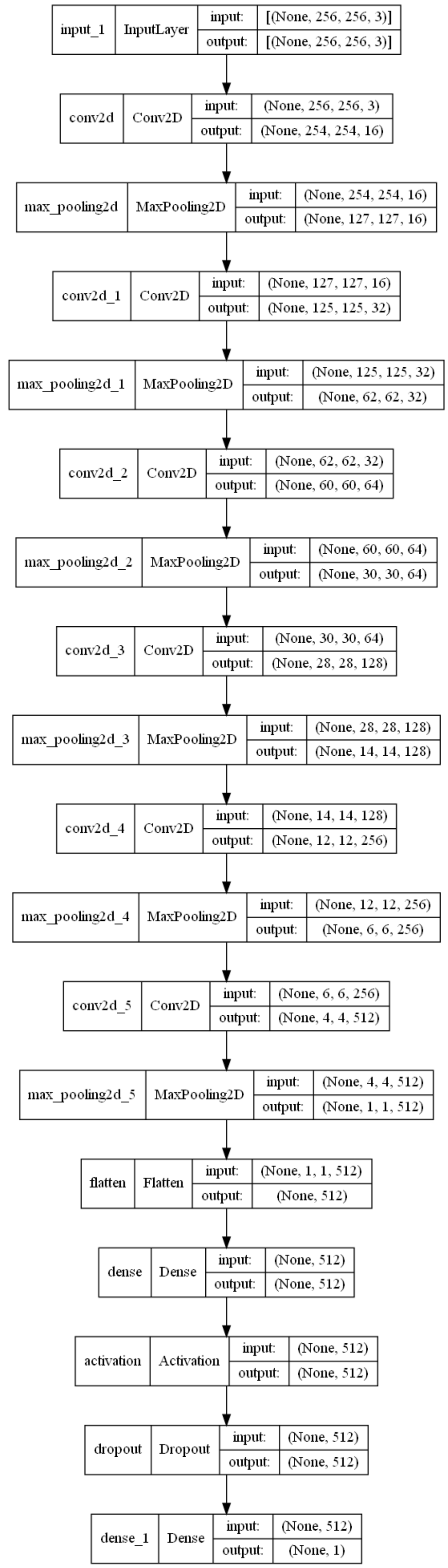

Fig. 3: Overall model architecture for vision anomaly detection.

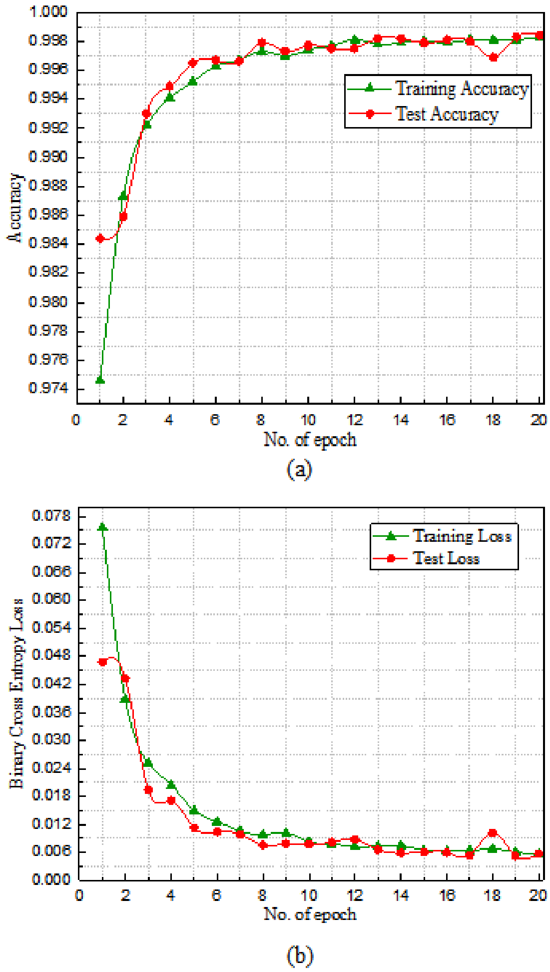

Fig. 4: (a) Training accuracy and test accuracy, and (b) training loss and test loss, curve of the surface crack detection data set using SGReLU.

B. Dataset Preparation

The surface crack detection dataset is used for vision anomaly detection [10]. It contains two types of data, both normal condition and anomaly condition. Normal condition dataset has no cracks in the concrete surface but anomaly condition dataset has cracks in the concrete surface. Both types of datasets contain 20,000 images each with RGB channels and a resolution of 227 × 227 pixels. Fig. 2 shows samples for both normal and anomalous data. Image data generator is used to process the dataset. The rotational range, horizontal flip class mode and color mode are selected 30°C, True, Binary and rgb respectively. Width shift range, zoom range and height shift range, all three are selected as 0.2. The target size of the dataset is (256,256) and $\frac{1}{255.0}$ is used for data scaling.

C. Model Architecture

Fig. 3 shows the overall model architecture for surface anomaly detection. Six conv2D layers and six maxpooling2D layers are used consecutively, then flatten layer, dense layer, activation layer, dropout layer, and another dense layer is used to get the final output result. In the activation layer, six different types of activation functions including SGReLU are used and the results are compared in the next section.

TABLE I: Performance comparison among different activation functions in terms of test accuracy, test loss and computational time.

Activation function	ReLU	LReLU	Swish	ELU	SELU	**SGReLU**
Top- 20 average test accuracy (%)	99.4	99.57	99.52	98.93	99.25	**99.6**
Top- 20 average test loss	0.017	0.014	0.015	0.033	0.025	**0.0125**
Time complexity (s)	10955	11109	11109	11259	11326	**10533**

III. Result and Discussion

In the instance of anomaly detection for the surface crack detection dataset, both training and test accuracy using SGReLU have obtained satisfactory values, as shown in Fig. 4(a). After the 20^{th} epoch, the test accuracy is 99.84%. In the case of the designed SGReLU-based deep learning network, both training and validation losses are presented in Fig. 4(b).

In Table I, SGReLU not only surpasses ReLU in terms of accuracy, but it also significantly reduces binary cross-entropy loss. The total time required by the SGReLU function-based approach is likewise less than that required by ReLU. The value of α was chosen 0.5 for the operation.

The performance comparison of different activation functions in the surface crack detection dataset in terms of test accuracy, test loss and time complexity are provided in Table I. SGReLU has gained 99.6% top-20 average test accuracy and outperformed other activation functions, such as ReLU, LReLU, swish, ELU, and SELU, by 0.2%, 0.03%, 0.08%, 0.67%, and 0.35%, respectively. Furthermore, the top-20 average test loss is 0.0125 s in case of SGReLU which is less than others. The computation complexity in case of SGReLU is 10533 s which is also comparatively less than others.

IV. Conclusion and Future Work

From the experimental results, it can be seen that SGReLU function with a fixed hyper-parameter value of α = 0.5 has outperformed different activation functions in case of vision anomaly detection. SGReLU function has not only achieved the highest accuracy but also has maintained the lowest loss as well as operational time in compared to others. In future, SGReLU with various values of α using hyper-parameter tuning or trainable parameter can be experimented for vision anomaly detection to boost performance.

ACKNOWLEDGMENT

This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2021-2018-0-01396) supervised by the IITP (Institute for Information & Communications Technology Planning & Evaluation).

REFERENCES

[1] J. Yang, R. Xu, Z. Qi, and Y. Shi, "Visual anomaly detection for images: A survey," *arXiv [cs.CV]*, 2021.

[2] X. Xie and M. Mirmehdi, "TEXEMS: texture exemplars for defect detection on random textured surfaces," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 8, pp. 1454–1464, 2007.

[3] V. Nair and G. E. Hinton, "Rectified linear units improve restricted boltzmann machines," in *Proc. International Conference on Machine Learning*, Haifa, Israel, 2010, pp. 807-814.

[4] X. Glorot, A. Bordes, and Y. Bengio, "Deep sparse rectifier neural networks." in *Proc. International Conference on Artificial Intelligence and Statistics Conference*, Ft. Lauderdale, FL, USA, 2011.

[5] A. Apicella, F. Donnarumma, F. Isgrò, and R. Prevete, "A survey on modern trainable activation functions," *Neural Network*, vol. 138, pp. 14–32, 2021.

[6] P. Ramachandran, B. Zoph, and Q. V. Le, "Searching for Activation Functions,"*arXiv [cs.NE]*, 2017.

[7] G. Klambauer, T. Unterthiner, A. Mayr, and S. Hochreiter, "Self-Normalizing Neural Networks," *arXiv [cs.LG]*, 2017.

[8] D.-A. Clevert, T. Unterthiner, and S. Hochreiter, "Fast and accurate deep network learning by exponential linear units (ELUs)," *arXiv [cs.LG]*, 2015.

[9] I. Jahan, M. F. Ahmed, M. O. Ali, and Y. M. Jang, "Self-gated rectified linear unit for performance improvement of deep neural networks," *ICT Express*, 2022.

[10] trinadhbavisetti, "Surface crack detection," *Kaggle.com*, 31-Oct-2021. [Online]. Available: https://www.kaggle.com/trinadhbavisetti/surface-crack-detection/data. [Accessed: 28-Jan-2022].