Adaptive Smooth Activation Function for Improved Organ Segmentation and Disease Diagnosis

Koushik Biswas*[1][0000−0002−9818−8966], Debesh Jha[1][0000−0002−8078−6730], Nikhil Kumar Tomar[1], Meghana Karri[1][0000−0002−0440−3785], Amit Reza[2][0000−0001−7934−0259], Gorkem Durak[1], Alpay Medetalibeyoglu[1], Matthew Antalek[1], Yury Velichko[1], Daniela Ladner[1][0000−0001−5526−8272], Amir Borhani, and Ulas Bagci[1][0000−0001−7379−6829]

[1] Machine and Hybrid Intelligence Lab, Northwestern University, USA
{koushik.biswas,debesh.jha,meghana.karri,medetalibeyoglu.alpay,y-velichko, dladner,amir.borhani,ulas.bagci}@northwestern.edu
[2] Space Research Institute (IWF) of Austrian Academy of Sciences, Graz, Austria
amitreza@gmail.com

Abstract. The design of activation functions constitutes a cornerstone for deep learning (DL) applications, exerting a profound influence on the performance and capabilities of neural networks. This influence stems from their ability to introduce non-linearity into the network architecture. By doing so, activation functions empower the network to learn and model intricate data patterns and relationships, surpassing the limitations of linear models. In this study, we propose a new activation function, called Adaptive Smooth Activation Unit ***(ASAU)***, tailored for optimized gradient propagation, thereby enhancing the proficiency of deep networks in medical image analysis. We apply this new activation function to two important and commonly used general tasks in medical image analysis: automatic disease diagnosis and organ segmentation in CT and MRI scans. Our rigorous evaluation on the *RadImageNet* abdominal/pelvis (CT and MRI) demonstrates that our ASAU-integrated classification frameworks achieve a substantial improvement of 4.80% over ReLU based frameworks in classification accuracy for disease detection. Also, the proposed framework on Liver Tumor Segmentation (LiTS) 2017 Benchmarks obtains 1%-to-3% improvement in dice coefficient compared to widely used activations for segmentation tasks. The superior performance and adaptability of ASAU highlight its potential for integration into a wide range of image classification and segmentation tasks. The code is available at https://github.com/koushik313/ASAU.

Keywords: Activation function · Organ segmentation · Classification

*Corresponding author: Koushik Biswas, Email: koushik.biswas@northwestern.edu

1 Introduction

Computer Aided Diagnosis: Automated disease detection from abdominal scans (CT and MRI) has become increasingly important for rapid and accurate diagnoses. This technology plays a pivotal role in early detection of diseases, treatment planning, and prognosis. For instance, abdominal CT scans offer a cost-effective and widely available modality, making them essential for emergency conditions and lesion detection. They provide high-resolution images that are well-suited for identifying acute abdominal pathologies and tumors. Conversely, abdominal MRIs excel in soft tissue contrast, enabling superior tumor type classification without ionizing radiation.

However, the inherent complexity of abdominal anatomy presents a challenge. Numerous organs and structures are densely packed, making the distinction between healthy and pathological tissues difficult, especially in the early stages of disease where subtle differences may arise [21]. This complexity can lead to variability in interpretation among radiologists, particularly for rare or atypical conditions. To address these challenges and improve diagnostic efficiency, deep learning based computer-aided diagnosis (CAD) algorithms are increasingly employed to minimize diagnostic errors.

Organ Segmentation: Organ segmentation serves as a foundational step for numerous tasks in medical image analysis, including CAD systems. Often, the initial stage of a CAD system involves organ segmentation. Here, we will focus on liver segmentation as an illustrative example to clinically motivate our work. However, the proposed method is generic and applicable to diverse organs, pathologies, and even non-medical applications.

Significance of liver diseases: Liver cancer represents the third leading cause of cancer-related mortality globally [1]. Liver volume quantification plays a crucial role in assessing various diseases, such as cirrhosis. Additionally, lesions within the liver can be subsequently identified and quantified. Despite its substantial size, the liver presents a segmentation challenge due to its shape variability, anatomical alterations in diseased states, and close proximity to other abdominal organs.

Smooth activation functions are needed: Current deep learning approaches have shown great promise in enhancing the quality of classification and segmentation tasks in medical imaging [10,23,13]. While the design and selection of appropriate activation functions are fundamental to the success of DL models, existing activation functions are chosen mostly due to efficiency reasons [19,16,6]. Traditional activation functions such as ReLU [19] and its variants (for example, Leaky ReLU [25], PReLU [6]), despite being instrumental to the success of DL, exhibit notable shortcomings. These activation functions are susceptible to information loss in regions with negative inputs and often struggle to capture the subtle nuances crucial for delineating intricate anatomical structures [6,4]. Consequently, these limitations can lead to segmentation inaccuracies, which are highly undesirable in a clinical setting. To address these challenges, the development of more specialized activation functions tailored to the specific needs of medical image analysis becomes paramount. More specifically, we hypothesize

that the smoothness of the activation function is crucial for the practical and theoretical aspects of training DL models. Smoothness (meaning that it is continuously differentiable at all points in its domains) ensures that gradient-based optimization methods work effectively, contributing to the stability, efficiency, and generalization capabilities of the models.

Proposed approach: Here, we present a novel smooth activation function, called Adaptive Smooth Activation Function (ASAU) (Figures 1-2-3), carefully designed to handle the challenges prevalent in deep learning based medical image analysis. This function embodies a methodological shift towards smoother, more continuous transitions, offering refined gradients that promote the intricate learning necessary for high-fidelity classification and segmentation tasks. With the rigorous testing and empirical evaluation on classification and segmentation tasks, our research indicates that applying this ASAU within deep network architectures (CNNs and Transformers) can substantially enhance multiclass disease classification (28 diseases from CT, 26 diseases from MRI) from one of the largest radiology scan benchmark (RadImageNet) and liver segmentation from 201 CT scans. Our comprehensive experiments demonstrate almost 5% improvement in multiclass disease classification accuracy in both CT and MRI and 1-to-3% improvement in liver segmentation tasks from CT scans.

2 Method

We present the Adaptive Smoothing Activation Unit (ASAU) using a smooth approximation of the general maximum function family. The maximum function is defined in equation (1). Popular activation functions like ReLU [19], Leaky ReLU [16], and Parametric ReLU are special cases of the maximum function, and they are not differentiable at the origin. First, we approximate the maximum function using a function that is smooth (continuously differentiable at the whole real line) and then turn it into a differentiable form. We consider the maximum function of two values, $\max(x_1, x_2)$ as:

$$\max(x_1, x_2) = \begin{cases} x_1, & \text{if } x_1 \geq x_2 \\ x_2, & \text{otherwise.} \end{cases} \tag{1}$$

We can rewrite the equation (1) as follows:

$$\max(x_1, x_2) = x_1 + \max(0, x_2 - x_1). \tag{2}$$

Note that the maximum function is not differentiable at the origin, while one preferably needs differentiable functions during backpropagation for uncorrupted information flow. Unlike popular activation functions such as ReLU, Mish [18] is a *smooth* activation function that is handcrafted and fixed. Mish is defined as $xtanh(ln(1 + e^x))$. If we add a non-negative parameter β in Mish, it can approximate the ReLU activation function too. Note that,

$$\text{if } \beta \to \infty, \text{then } xtanh(\alpha \text{SoftPlus}(\beta x)) \approx max(0, x) \tag{3}$$

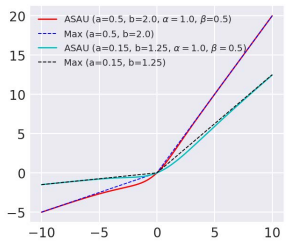

Fig. 1: Approximation of Maximum function using ASAU for varying values of a, b, α, and β. As $\beta \to \infty$, ASAU smoothly approximates the maximum function.

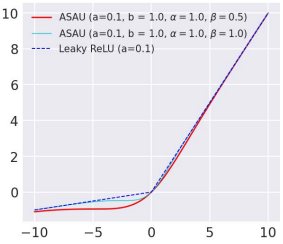

Fig. 2: Approximation of Leaky ReLU using ASAU for different values of a, b, α, and β. As $\beta \to \infty$, ASAU smoothly approximates Leaky ReLU.

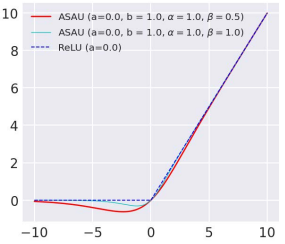

Fig. 3: Approximation of ReLU using ASAU for different values of a, b, α, and β. As $\beta \to \infty$, ASAU smoothly approximates ReLU.

where α is another parameter that controls the smoothness of the function, and SoftPlus is defined as $ln(1 + e^x)$. Maxout is another popular activation function [5], which is a linear combination of the convex functions that generalize the ReLU, Leaky ReLU, or its variants. Replacing the equation (3) in equation (2), we can have

$$\max(x_1, x_2) \approx x_1 + (x_2 - x_1) tanh(\alpha \text{SoftPlus}(\beta(x_2 - x_1)). \tag{4}$$

If we consider $x_1 = ax$ and $x_2 = bx$ we can get an approximation of the Maxout family as

$$f(x; a, b, \alpha, \beta) = \max(ax, bx) \approx ax + (b - a)x tanh(\alpha \text{SoftPlus}(\beta(b - a)x), \tag{5}$$

in particular, we can derive the smooth approximation of the ReLU and Leaky ReLU activation function (by considering $a = 0$ & $b = 1$ or $a = 0.01$ & $b = 1$, respectively). We call the function in equation (5) as the *Adaptive Smooth Activation Unit (ASAU)*. Figure (1) shows how the maxout has been approximated with the ASAU function. Similarly, Figure (2) and Figure (3) demonstrate how the Leaky ReLU and ReLU are approximated with the ASAU function, respectively. The derivative of the ASAU with respect to x is

$$\frac{\partial f}{\partial x} = a + (b - a) \tanh\left(\alpha \ln\left(e^{(b-a)\beta x} + 1\right)\right) + \frac{(b - a)^2 \alpha\beta x e^{(b-a)\beta x} \operatorname{sech}^2\left(\alpha \ln\left(e^{(b-a)\beta x} + 1\right)\right)}{e^{(b-a)\beta x} + 1} \tag{6}$$

Note that the parameters a, b, α, and β can be used as hyperparameters or trainable parameters. In the case of the trainable parameters, we will pass the a, b, α, and β into the backpropagation algorithm. The derivative of a, b, α, and β will be computed as follows:

$$\frac{\partial f}{\partial a} = x - x\tanh\left(\alpha\ln\left(\mathrm{e}^{x\beta(b-a)}+1\right)\right) - \frac{x^2\alpha\beta\left(b-a\right)\mathrm{e}^{x\beta(b-a)}\operatorname{sech}^2\left(\alpha\ln\left(\mathrm{e}^{x\beta(b-a)}+1\right)\right)}{\mathrm{e}^{x\beta(b-a)}+1} \tag{7}$$

$$\frac{\partial f}{\partial b} = x\tanh\left(\alpha\ln\left(\mathrm{e}^{x\beta(b-a)}+1\right)\right) + \frac{x^2\alpha\beta\left(b-a\right)\mathrm{e}^{x\beta(b-a)}\operatorname{sech}^2\left(\alpha\ln\left(\mathrm{e}^{x\beta(b-a)}+1\right)\right)}{\mathrm{e}^{x\beta(b-a)}+1} \tag{8}$$

$$\frac{\partial f}{\partial \alpha} = \left(b-a\right)x\ln\left(\mathrm{e}^{(b-a)x\beta}+1\right)\operatorname{sech}^2\left(\ln\left(\mathrm{e}^{(b-a)x\beta}+1\right)\alpha\right) \tag{9}$$

$$\frac{\partial f}{\partial \beta} = \frac{\left(b-a\right)^2x^2\alpha\mathrm{e}^{(b-a)x\beta}\operatorname{sech}^2\left(\alpha\ln\left(\mathrm{e}^{(b-a)x\beta}+1\right)\right)}{\mathrm{e}^{(b-a)x\beta}+1} \tag{10}$$

By following the proposition by Kidger and Lyons (see below), one can conclude that ASAU is an ideal candidate for an activation function because ASAU is continuously differentiable and is a non-polynomial function.

Proposition 1. (Theorem 1.1 by Kidger and Lyons, 2020 [14]) :- Let $\rho : \mathbb{R} \to \mathbb{R}$ be any continuous function. Let N_n^ρ represent the class of neural networks with activation function ρ, with n neurons in the input layer, one neuron in the output layer, and one hidden layer with an arbitrary number of neurons. Also, let $K \subseteq \mathbb{R}^n$ be compact. Then, N_n^ρ is dense in $C(K)$ if and only if ρ is non-polynomial where C(K) is the space of all continuous functions.

3 Experiments

3.1 Datasets

For the classification tasks, we use *RadImageNet* [17] database, which is an open-access medical imaging database designed to improve transfer learning performance on downstream medical imaging applications and perhaps the largest ever medical imaging dataset so far. From the whole dataset, we experiment on CT and MRI abdominal/pelvis, consisting of 28 disease classes with an average class size of 4994 and a total of 139,825 slices (i.e., the dataset itself is designed to have *slices* per disease although the overall scans are in volumes-3D). 28 disease classes are adrenal pathology, arterial pathology, ascites, bil dil, bladder pathology, bowel abnormality, bowel inflammation, bowel mass degenerative changes, dilated urinary tract, fat-containing tumor, gallbladder pathology, gallstone, intraperitoneal mass, liver lesion, normal osseous, neoplasm, ovarian pathology, pancreatic lesion, post-op, prostate lesion, renal lesion, soft tissue collection, soft tissue mass, splenic lesion, urolithiasis and uterine pathology. We

also experimented on the MRI abdomen/pelvis dataset consisting of 26 distinct (disease) classes with an average class size of 3513 slices and a total number of 91,348 slices. While most of these classes overlap with the CT dataset, the MRI dataset uniquely includes disease classes like enlarged organs and liver disease, which are not in the CT dataset. Conversely, the CT dataset has a specific class for entire abdominal organs.

For the segmentation tasks, we select the Liver Tumor Segmentation Benchmark (LiTS) [2] dataset, which is a multi-center dataset collected from seven clinical centers. It contains 201 CT images of the abdomen. The dataset is completely anonymized, and the images have been reviewed visually to preclude the presence of personal identifiers. The whole dataset is distributed into a training dataset with 130 CT scans (available to the public), and the test dataset has 71 CT scans (blind). Only the training dataset is made publicly available. Thus, we only trained the segmentation networks on the training dataset.

3.2 Implementation details

We used the PyTorch [20] framework for all the segmentation experiments. The networks were configured to train for liver segmentation tasks with a batch size of 16 and a learning rate set to $1e^{-4}$. 500 epochs of training were performed to fine-tune the network parameters adequately with an early stopping patience of 50. To enhance the performance of our network, we used a combination of binary cross-entropy and dice loss, and an Adam optimizer was chosen for parameter updates. The data was split into 80% for training, 10% for validation, and 10% for testing. We resized the image to 256×256 pixels in-plane resolution to optimize the trade-off between training time and model complexity. To avoid bias, we also split the cases into independent training (70 patients), validation (30 patients), and test (30 patients) sets. The volumetric CT scans were processed pseudo-3D (slice-by-slice) to fit into regular computer hardware (GPU). During prepossessing, we extracted the healthy liver masks.

For classification (disease diagnosis) experiments, we consider Tensorflow-Keras [3] framework. We consider ResNet-18 [7] and ResNet-50 [7] as baseline classification networks. The networks are trained with batch size 32, initial learning rate 0.00001 with Adam [15] optimizer and $1e^{-4}$ weight decay rate. The data was split into 80% for training, 10% for validation, and 10% for testing. Results are reported on CT scan image data in Table 1 and MRI image data in Table 2.

3.3 Results and Discussion

Table 1 shows the results of ResNet-50 [8] and ResNet-18 architectures on the CT abdominal/pelvis scan dataset. Here, we examine the efficacy of ReLU, LReLU, and PReLU activation along with our proposed ASAU. For all the experiments, the ASAU-based method outperformed all other experimental results in all the metrics by a significant margin. On the ResNet-50-based architecture, the ASAU integrated method obtains a high improvement of 8.26% in terms of MCC.

Table 1: Impact of Activation Functions on 28 Classes in RadImageNet Abdominal/Pelvis CT Scans. "MCC" refers to Matthew's Correlation Coefficient.

Activation Function	Method	Macro Average			Micro Average			Accuracy	MCC
		Precision	Recall	F1-score	Precision	Recall	F1-score		
ReLU	ResNet-50	35.74	23.16	23.35	67.67	67.67	67.67	67.67	52.84
LReLU	ResNet-50	35.97	23.56	23.65	68.10	68.10	68.10	68.10	53.37
PReLU	ResNet-50	36.30	23.89	23.77	68.77	68.77	68.77	68.77	54.10
ASAU	ResNet-50	**46.29**	**31.30**	**34.15**	**72.47**	**72.47**	**72.47**	**72.47**	**61.10**
ReLU	ResNet-18	33.70	25.83	27.43	70.12	70.12	70.12	70.12	57.10
Leaky ReLU	ResNet-18	33.91	25.98	27.60	70.38	70.38	70.38	70.38	57.52
PReLU	ResNet-18	34.20	26.31	27.99	70.52	70.52	70.52	70.52	57.67
ASAU	ResNet-18	**36.50**	**27.30**	**28.99**	**71.10**	**71.10**	**71.10**	**71.10**	**58.59**
ReLU	MobileNet V2	21.42	12.16	12.85	58.68	58.68	58.68	58.68	35.21
LReLU	MobileNet V2	21.96	12.51	13.24	59.26	59.26	59.26	59.26	35.93
PReLU	MobileNet V2	22.10	12.69	13.39	59.62	59.46	59.46	59.46	36.09
ASAU	MobileNet V2	**28.07**	**16.29**	**18.10**	**62.64**	**62.64**	**62.64**	**62.64**	**43.05**
ReLU	ShuffleNet	24.29	15.37	16.53	62.70	62.70	62.70	62.70	43.19
Leaky ReLU	ShuffleNet	24.50	15.58	16.77	63.09	63.09	63.09	63.09	43.31
PReLU	ShuffleNet	24.67	15.73	16.93	63.31	63.31	63.31	63.31	43.42
ASAU	ShuffleNet	**26.20**	**17.10**	**18.96**	**65.22**	**65.22**	**65.22**	**65.22**	**46.57**

Table 2: Impact of Activation Functions on 26 Classes in RadImageNet Abdominal/Pelvis MRI Scans. "MCC" refers to Matthew's Correlation Coefficient.

Activation Function	Method	Macro Average			Micro Average			Accuracy	MCC
		Precision	Recall	F1-score	Precision	Recall	F1-score		
ReLU	ResNet-50	40.43	24.83	28.04	86.86	86.86	86.86	86.86	62.92
LReLU	ResNet-50	40.56	24.99	28.56	87.10	87.10	87.10	87.10	63.25
PReLU	ResNet-50	41.10	25.20	28.81	87.25	87.25	87.25	87.25	63.47
ASAU	ResNet-50	**44.46**	**33.23**	**36.17**	**89.20**	**89.20**	**89.20**	**89.20**	**69.75**
ReLU	ResNet-18	35.15	26.66	28.82	87.96	87.96	87.96	87.96	66.65
LReLU	ResNet-18	35.67	26.87	28.99	88.12	88.12	88.12	88.12	66.72
PReLU	ResNet-18	35.91	26.80	29.10	88.27	88.27	88.27	88.27	66.81
ASAU	ResNet-18	**38.58**	**27.58**	**29.67**	**88.62**	**88.62**	**88.62**	**88.62**	**67.32**
ReLU	MobileNet V2	10.64	10.68	10.30	82.64	82.64	82.64	82.64	47.12
LReLU	MobileNet V2	11.05	11.21	10.89	82.99	82.99	82.99	82.99	47.89
PReLU	MobileNet V2	11.21	11.35	11.02	83.15	83.15	83.15	83.15	48.04
ASAU	MobileNet V2	**17.88**	**14.92**	**15.21**	**84.73**	**84.73**	**84.73**	**84.73**	**55.32**
ReLU	ShuffleNet	19.58	14.56	14.89	84.47	84.47	84.47	84.47	55.14
LReLU	ShuffleNet	19.67	14.77	14.95	84.70	84.70	84.70	84.70	55.36
PReLU	ShuffleNet	**19.79**	14.89	14.99	85.30	85.30	85.30	85.30	55.70
ASAU	ShuffleNet	19.78	**14.95**	**15.30**	**85.35**	**85.35**	**85.35**	**85.35**	**55.76**

Table 3: Performance comparison of different activation functions on liver segmentation benchmark (LiTS) dataset.

Activation Function	UNet [22]				DoubleUNet [10]				ColonSegNet [9]			
	mDSC	mIoU	Rec.	Prec.	mDSC	mIoU	Rec.	Prec.	mDSC	mIoU	Rec.	Prec.
ReLU	82.06	73.40	77.82	91.10	86.24	77.89	80.68	95.00	80.87	71.71	**80.07**	87.50
LReLU	82.47	73.98	78.46	91.17	86.27	78.21	80.20	95.67	80.84	71.69	79.31	88.10
PReLU	82.71	73.91	78.77	91.21	86.39	78.09	80.39	95.59	78.91	69.85	77.89	86.27
ASAU	**83.62**	**75.08**	**80.59**	**93.29**	**86.88**	**78.73**	**80.77**	**96.58**	**83.03**	**74.18**	78.79	**91.17**

Table 2 shows the MRI abdominal/pelvis dataset results. Again, here, ASU-integrated ResNet50 obtained an MCC score of 69.75%, which is 6.83% better than ReLU based method.

Tables 3, 4, and 5 summarize LiTs datasets' results. Here, we have compared the performance of eight state-of-the-art medical image segmentation methods

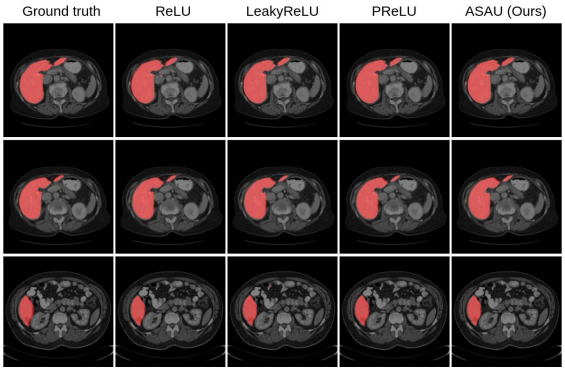

Fig. 4: Qualitative results comparison for liver segmentation results (LiTS) on the ResUNet++ model with various activation functions. It can be observed that ASAU produces more accurate segmentation maps in all three cases.

Table 4: Additional segmentation experiments continued from Table 3.

Activation Function	TransNetR [13]				TransResUNet [23]				ResUNet++ [11]			
	mDSC	mIoU	Rec.	Prec.	mDSC	mIoU	Rec.	Prec.	mDSC	mIoU	Rec.	Prec.
ReLU	86.11	77.95	80.16	96.34	**86.38**	**78.23**	**80.85**	95.77	77.03	68.19	79.90	83.57
LReLU	86.30	78.37	**80.37**	96.67	86.18	77.82	79.80	96.59	75.03	66.35	71.64	**84.05**
PReLU	86.39	78.29	80.32	96.52	85.68	77.30	79.46	96.12	74.39	66.02	74.38	81.99
ASAU	**86.41**	**78.49**	80.30	**96.65**	86.35	78.15	79.89	**96.67**	**78.12**	**69.56**	**81.50**	83.81

Table 5: Additional segmentation experiments continued from Table 3.

Activation Function	NanoNet-A [12]				UNext [24]			
	mDSC	mIoU	Rec.	Prec.	mDSC	mIoU	Rec.	Prec.
ReLU	75.05	66.53	73.33	83.25	80.27	71.31	78.84	87.82
LReLU	74.05	65.02	73.36	84.23	80.92	72.15	76.84	**90.85**
PReLU	74.84	66.08	72.47	85.24	79.08	69.95	77.38	88.55
ASAU	**76.57**	**67.30**	**75.88**	**84.61**	**81.98**	**73.03**	**79.01**	90.45

(UNet [22], DoubleUNet [10], ColonSegNet [9], UNext [24], TransNetR [13], TransResNet [23], ResUNet++ [11] and NanoNet-A [12]) with four different activation functions. From the results, it can be observed that ASAU has improvement from 1% to 3% as compared to the ReLU activation function. DoubleUNet, along with ASAU, has set a new baseline for liver segmentation tasks with a DSC of 86.88%, mIoU of 78.73%, recall of 80.77%, and precision of 96.58%.

The results for liver segmentation (LiTS) on the ResUNet++ model with different activation functions are displayed in Figure 4. It can be observed that ASAU performs better in segmentation accuracy compared to the SOTA baseline activation functions, with fine details captured, while ReLU-based methods started to ignore such details. ASAU is advantageous for both classification and segmentation algorithms because it is smooth and maintains a stable and differentiable gradient, which is essential for backpropagation algorithms using gradient descent.

4 Conclusion

This work presents a novel activation function termed the Adaptive Smooth Maximum Unit (ASAU), specifically designed to elevate the performance of medical image analysis tasks through more stable and uncorrupted information flow during backpropagation, encompassing both disease classification and segmentation. The remarkable outcomes both in multiclass disease diagnosis and segmentation achieved by our framework on established benchmarks like RadImageNet and LiTS datasets convincingly demonstrate its potential for broader applicability across an array of medical image analysis tasks. Moving forward, our research efforts will be directed towards more effectively leveraging three-dimensional spatial information within the framework. This advancement aims to further refine the diagnostic tools available to clinicians and holds promise for extending the applicability of ASAU to encompass a comprehensive suite of abdominal organs. By achieving these goals, we strive to enhance the overall robustness and generalizability of our proposed approach.

Acknowledgements This project is supported by NIH funding: R01-CA246704, R01-CA240639, U01-DK127384-02S1, and U01-CA268808.

Disclosure of Interests The authors declare no competing interests.

References

1. Arnold, M., Abnet, C.C., Neale, R.E., Vignat, J., Giovannucci, E.L., McGlynn, K.A., Bray, F.: Global burden of 5 major types of gastrointestinal cancer. Gastroenterology **159**(1), 335–349 (2020)
2. Bilic, P., Christ, P., Li, H.B., Vorontsov, E., Ben-Cohen, A., Kaissis, G., Szeskin, A., Jacobs, C., Mamani, G.E.H., Chartrand, G., et al.: The liver tumor segmentation benchmark (lits). Medical Image Analysis **84**, 102680 (2023)
3. Chollet, F., et al.: Keras. https://keras.io (2015)
4. Glorot, X., Bordes, A., Bengio, Y.: Deep sparse rectifier neural networks. In: Proceedings of the fourteenth international conference on artificial intelligence and statistics. pp. 315–323. JMLR Workshop and Conference Proceedings (2011)
5. Goodfellow, I.J., Warde-Farley, D., Mirza, M., Courville, A., Bengio, Y.: Maxout networks (2013)
6. He, K., Zhang, X., Ren, S., Sun, J.: Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In: Proceedings of the IEEE international conference on computer vision. pp. 1026–1034 (2015)
7. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 770–778 (2016)
8. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 770–778 (2016)

9. Jha, D., Ali, S., Tomar, N.K., Johansen, H.D., Johansen, D., Rittscher, J., Riegler, M.A., Halvorsen, P.: Real-Time Polyp Detection, Localization and Segmentation in Colonoscopy using Deep Learning. IEEE Access **9**, 40496–40510 (2021)
10. Jha, D., Riegler, M.A., Johansen, D., Halvorsen, P., Johansen, H.D.: DoubleU-Net: A Deep Convolutional Neural Network for Medical Image Segmentation. In: Proceedings of the IEEE 33rd International Symposium on computer-based medical systems (CBMS). pp. 558–564 (2020)
11. Jha, D., Smedsrud, P.H., Riegler, M.A., Johansen, D., De Lange, T., Halvorsen, P., Johansen, H.D.: ResUNet++: An Advanced Architecture for Medical Image Segmentation. In: Proc. of int. symposium on multimedia. pp. 225–2255 (2019)
12. Jha, D., Tomar, N.K., Ali, S., Riegler, M.A., Johansen, H.D., Johansen, D., de Lange, T., Halvorsen, P.: Nanonet: Real-time polyp segmentation in video capsule endoscopy and colonoscopy. In: Proceedings of the International Symposium on Computer-Based Medical Systems (CBMS). pp. 37–43 (2021)
13. Jha, D., Tomar, N.K., Sharma, V., Bagci, U.: TransNetR: Transformer-based Residual Network for Polyp Segmentation with Multi-Center Out-of-Distribution Testing. In: Proceedings of the Medical Imaging with Deep Learning (MIDL). pp. 1372–1384 (2023)
14. Kidger, P., Lyons, T.: Universal approximation with deep narrow networks (2020)
15. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization (2017)
16. Maas, A.L., Hannun, A.Y., Ng, A.Y., et al.: Rectifier nonlinearities improve neural network acoustic models. In: Proceedings of the International Conference on Machine Learning. vol. 30, p. 3 (2013)
17. Mei, X., Liu, Z., Robson, P.M., Marinelli, B., Huang, M., Doshi, A., Jacobi, A., Cao, C., Link, K.E., Yang, T., et al.: RadImageNet: RadImageNet: An Open Radiologic Deep Learning Research Dataset for Effective Transfer Learning. Radiology: Artificial Intelligence **4**(5), e210315 (2022)
18. Misra, D.: Mish: A self regularized non-monotonic activation function (2020)
19. Nair, V., Hinton, G.E.: Rectified linear units improve restricted boltzmann machines. In: Proceedings of the 27th International Conference on Machine Learning (ICML). pp. 807–814 (2010)
20. Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., et al.: Pytorch: An imperative style, high-performance deep learning library. Advances in neural information processing systems **32** (2019)
21. Reyes, M., Ballester, M.A.G., Li, Z., Kozic, N., Summers, R.M., Linguraru, M.G.: Interpretability of anatomical variability analysis of abdominal organs via clusterization of decomposition modes. In: Proceedings of the International Conference of the IEEE Engineering in Medicine and Biology Society. pp. 355–358 (2008)
22. Ronneberger, O., Fischer, P., Brox, T.: U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Proceedings of the 18th International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 234–241 (2015)
23. Tomar, N.K., Shergill, A., Rieders, B., Bagci, U., Jha, D.: TransResU-Net: Transformer based ResU-Net for real-time colonoscopy polyp segmentation. arXiv preprint arXiv:2206.08985 (2022)
24. Valanarasu, J.M.J., Patel, V.M.: Unext: Mlp-based rapid medical image segmentation network (2022)
25. Xu, B., Wang, N., Chen, T., Li, M.: Empirical evaluation of rectified activations in convolutional network. arXiv preprint:1505.00853 (2015)