



Gated Channel Attention Network for Cataract Classification on AS-OCT Image

Zunjie Xiao¹, Xiaoqing Zhang¹, Risa Higashita^{1,2(✉)}, Yan Hu¹, Jin Yuan³,
Wan Chen³, and Jiang Liu^{1,4,5,6(✉)}

¹ Department of Computer Science and Engineering, Southern University of Science
and Technology, Shenzhen, China

² Tomey Corporation, Nagoya, Japan

³ Zhongshan Ophthalmic Center, Sun Yat-sen University, Guangzhou, China

⁴ Cixi Institute of Biomedical Engineering, Ningbo Institute of Materials Technology
and Engineering, Chinese Academy of Sciences, Ningbo, China

⁵ Guangdong Provincial Key Laboratory of Brain-inspired Intelligent Computation,
Department of Computer Science and Engineering, Southern University of Science
and Technology, Shenzhen, China

⁶ Research Institute of Trustworthy Autonomous Systems, Southern University
of Science and Technology, Shenzhen, China

Abstract. Nuclear cataract (NC) is the leading cause of blindness and vision impairment globally. Accurate NC classification is significant for clinical NC diagnosis. Anterior segment optical coherence tomography (AS-OCT) is a non-contact, high-resolution, objective imaging technique, which is widely used in diagnosing ophthalmic diseases. Clinical studies have shown that there is a significant correlation between the pixel density of the lens region on AS-OCT images and NC severity levels; however, automatic NC classification on AS-OCT images has not been seriously studied. Motivated by clinical research, this paper proposes a gated channel attention network (GCA-Net) to classify NC severity levels automatically. In the GCA-Net, we design a gated channel attention block by fusing the clinical priority knowledge, in which a gated layer is designed to filter out abundant features and a Softmax layer is used to build the weakly interacting for channels. We use a clinical AS-OCT image dataset to demonstrate the effectiveness of our GCA-Net. The results showed that the proposed GCA-Net achieves 94.3% in accuracy and outperformed strong baselines and state-of-the-art attention-based networks.

Keywords: AS-OCT · Nuclear cataract · Gated channel attention · Deep learning

1 Introduction

Cataract is the leading cause of reversible blindness and vision impairment worldwide [5]. Early treatment can address vision impairment and restore vision to

Z. Xiao and X. Zhang—Equal contribution.

© Springer Nature Switzerland AG 2021

T. Mantoro et al. (Eds.): ICONIP 2021, LNCS 13110, pp. 357–368, 2021.

https://doi.org/10.1007/978-3-030-92238-2_30

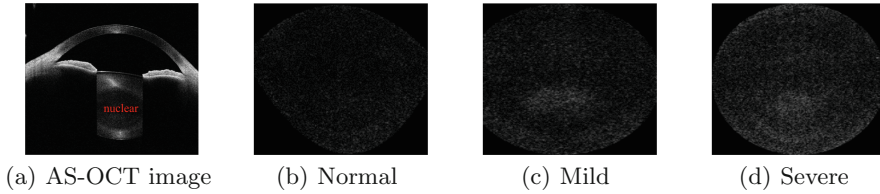


Fig. 1. The whole AS-OCT image shown in (a), and the center area is the nucleus. (b) Normal nucleus image; (c) Mild NC image where the nuclear opacity is asymptomatic; (d) Severe NC image where the nuclear opacity is symptomatic.

improve the cataract patient's quality of life. According to the location of the opacities, cataracts can be generally classified into three types: nuclear cataract (NC), cortical cataract (CC), and posterior subcapsular cataract (PSC). NC is the most common type of cataract, characterized by the increase of light scattering in the nucleus region of the crystalline lens area. In clinical practice, slit-lamp image is routinely used to diagnose NC based on standard cataract classification systems. Lens opacity classification system III (LOCS III) [4] is a well-accepted slit lamp image-based cataract classification system. With the development of nuclear opacity pathology, nuclear cataract can be divided into three stages [15]. (1) Normal: healthy or without nuclear opacity in the slit-lamp image; (2) mild (grade=1 or 2 in LOCS III): the nuclear opacity is asymptomatic; (3) severe (grade \leq 3 in LOCS III): the nuclear opacity is symptomatic. Mild NC can be relieved by clinical intervention, while severe NC needs to prepare for surgery as soon. Figure 1 shows the representative figures of AS-OCT nuclear areas at the three stages.

Anterior segment optical coherence tomography (AS-OCT) is a non-contact, high-resolution tomography technique, which can objectively and quickly obtain overall information of the entire lens. AS-OCT images have gradually been used in the diagnosis of various anterior segment ocular diseases such as glaucoma, cataracts, and keratitis [5]. For NC diagnosis, AS-OCT image can capture the nucleus region clearly while other ophthalmology images like fundus images cannot. The clinical study has shown that the average lens density (ALD) has a strong linear relationship with the nucleus region of AS-OCT images based on the LOCS III [21], which provided clinical support for automatic cataract classification on AS-OCT images. Following [21], clinical research [3, 14, 16, 20] further got the similar statistics results. Motivated by the preliminary works, [27] studied NC classification based on AS-OCT image, which uses the convolutional neural network (CNN), but they achieved poor performance.

Average nucleus density (AND) is a clinical indicator on AS-OCT image for nuclear cataract diagnosis, which is defined as the average pixel density in the nucleus region [21]. Figure 2 shows the distribution of AND in different stages of nuclear cataract. It can be seen that there are significant differences in the AND distribution among different NC stages, while many images are difficult to

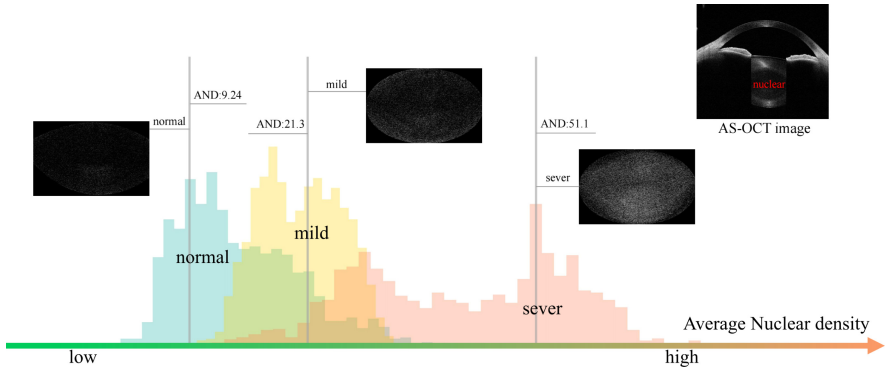


Fig. 2. The different nuclear opacity stages are reflected in the OCT image; the histogram reflects the sample distribution of three nuclear cataract severity levels (different colors mean different NC stages).

classify the severity of cataracts simply by AND (the overlap area as shown in Fig. 2).

In recent years, channel attention mechanism has become one of the most popular attention mechanisms due to its simplicity and effectiveness, which directly learns importance weights of each channel. In channel attention block, global average pooling (GAP) is used for integrating channel-wise information, which calculates the mean value of each channel. GAP collects the global mean value, which enhances the representation ability for global information, especially *AND*. Inspired by this relationship, we propose a simple yet effective gated channel attention network (GCA-Net) for NC classification automatically. In the GCA-Net, this paper designs a novel gated channel attention block, where a gating operator is used to mask and applies a weakly-interacting operator to model the global channel information.

The main contributions of this paper are as follows: (1) We develop a novel convolutional neural network (CNN) model named GCA-Net to discriminate opacity information for classifying NC levels into three severity levels. (2) This paper designs a simple yet effective channel attention (GCA) block comprised of three stages: gating, squeezing, and interacting, to capture the global information. (3) The results on a clinical AS-OCT image show that our GCA-Net surpasses state-of-the-art attention-based networks.

2 Related Work

2.1 Cataract Classification

In recent years, research scholars have proposed many advanced machine learning and deep learning methods for automatic cataract classification on different

ophthalmology image modalities [26]. [12] proposes an automatic NC classification system that contains three stages (region detection, pixel feature extraction, and level prediction) based on the ACHIKO-NC slit-lamp dataset, and achieves an average error of 0.36. Xu et al. also performed NC classification on the ACHIKO-NC dataset, using the group sparse regression (GSR) method and achieved 83.4% accuracy [25]; [24] proposed the semantic similarity method for slit lamp image-based NC classification and obtained better performance than GSR. [1] achieves an accuracy of 95% using support vector machines (SVM) to classify NC on ultrasound images, but the ultrasound image data sets used for their work are from animals. Li et al. achieved accurate cataract screening by improving the Haar wavelet transform algorithm on fundus images [2].

Compared with machine learning methods, deep learning methods are skilled at capturing useful feature representations. Gao et al. proposed a hybrid model of convolutional neural network (CNN) and recurrent neural network (RNN) based on slit-lamp images and achieved 82.5% accuracy for NC classification [6]. A team of Sun Yat-sen University proposed a congenital cataract screening platform based on deep learning [13]. Xu et al. proposed a global-local hybrid CNN network by fusing different parts of pathological information that achieves better performance than previous methods on fundus images [23, 26].

There are relatively few NC classification studies on AS-OCT images. Some clinical studies have verified its reliability on NC classification based on LOCS III [3, 16, 21]. [27] tried preliminary NC classification using deep learning methods on AS-OCT images. We combine clinical and methodological research to propose our own method.

2.2 Attention Mechanism

Attention mechanisms have empowered CNN models and achieved state-of-the-art results on various learning tasks [19]. In general, attention mechanisms can be mainly summarized into two groups, channel attention mechanism and spatial attention mechanism. SENet [10] firstly proposed the channel attention mechanism. It performs the GAP for channel squeeze, then reconstructs inter-dependencies of the channels through fully-connected (*fc*) layers, finally a *Sigmoid* layer is applied to generate channel weights for each channel. GENet [9] introduces a learnable layer for better exploiting the context feature, and FcaNet [19] increases the diversification of extracted features by extracting multi-band information. Bottleneck Attention Module (BAM) [17] and Convolutional Block Attention Module (CBAM) [22] combine the two attention mechanisms for getting the fused attention weights. To improve efficiency, ECANet [18] uses one-dimensional convolution layers to replace the original fully-connected layers in SENet.

3 Method

In this section, we first revisit the classical channel attention mechanism. Then we elaborate our GCA block in detail.

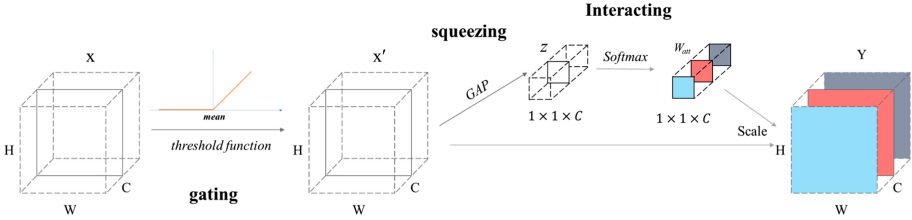


Fig. 3. A gated channel attention block.

3.1 Revisiting of Channel Attention

Channel attention is one of the most widely used attention module in CNNs. It uses a learnable block to adjust the importance of each channel and enhance the feature representation ability of the model. Given $X \in \mathbb{R}^{C \times W \times H}$ is the input feature tensor, where C denotes the number of channels, H and W denote the height and width of the feature map, respectively. The output $Y \in \mathbb{R}^{C \times W \times H}$ has the same shape of X with re-weighting of each channel. SENet [10] is the most classic channel attention mechanism consist of squeeze and excitation operation. The formula can be written as:

$$Y = \mathbf{F}_{scale}(W_{att}, X), \tag{1}$$

$$W_{att} = \mathbf{F}_{ex}(\mathbf{F}_{sq}(X)), \tag{2}$$

where $W_{att} \in \mathbb{R}^C$ is the channel attention weight, \mathbf{F}_{scale} refers to channel-wise multiplication, \mathbf{F}_{sq} represents the squeeze function GAP, and \mathbf{F}_{ex} is the excite function to transform the squeeze info to attention weights. Generally, the squeeze step compresses channel information, and excitation step calculates the channel weights W_{att} . For the first step, it usually use parameter-free function like global average pooling (GAP) [10] or global max pooling (GMP) [22] to compute channel-statistics information. For the second step, it adopts fc layers for inter-channel dependency reconstruction.

In this paper, we found that the dependency among channel-statistics information is weak, and fc layers do not work well for AS-OCT image-based NC classification. This is because *AND* is an important indicator for NC diagnosis on AS-OCT images. Hence, we design a simple yet effective channel attention block named gated channel attention (GCA) block and will be introduced in the next section.

3.2 Gated Channel Attention Block

Figure 3 shows the diagram of the structure of a gated channel attention (GCA) block, which comprises three stages: gating, squeezing, and interacting.

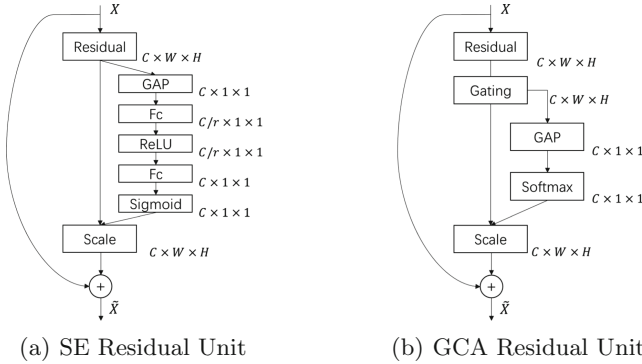


Fig. 4. The schema of the SE residual unit (left) and the GCA residual unit (right).

Gating: To suppress the redundant features in a feature map, we devise a gated unit to mask the irrelevant features. According to the clinical studies in Sect. 2.1, the higher density region has higher relevance with cataract. To this end, we proposed a high-value gate for masking the low-value influence. It is an adaptive *threshold* function in which we use the global average value from each feature map as the threshold value. This is because [11] demonstrated that pooling value below average suppressed neuron activations in a CNN model. Formally, the gated tensor $X' \in \mathbb{R}^{C \times W \times H}$ is generated by masking the low-value of input tensor $X \in \mathbb{R}^{C \times W \times H}$, such that the c -th channel is formulated by:

$$(X'_c)_{ij} = \mathbf{F}_{gating}(X_c)_{ij} = \text{Max}(\text{Mean}(X_c), (X_c)_{ij}), \tag{3}$$

where *Mean* function calculates the mean value of the feature map, *Max* function returns the largest item of input.

Squeezing: We use a squeezing operator to follow the gating operator, which is used to compute the channel-statistics feature information from each channel. This paper uses global average pooling (GAP) as squeezing operator, equivalent to the *AND* indicator for NC diagnosis. It can be written as follows:

$$z_c = \mathbf{F}_{GAP}(X'_c) = \frac{1}{W \times H} \sum_{i=1}^W \sum_{j=1}^H (X'_c)_{ij}, \tag{4}$$

where z_c denotes the output of GAP in c -th channel.

In the experiments, we test the effects of different pooling operators.

Interacting: In the third stage, we propose a weakly interacting operator to construct weak dependencies of inter-channel and set the relative weights for channels. The fully-connection operator is the first proposed method for channel interacting in channel attention block. However, it brings higher model complexity, and [18] simplifies the interacting stage using local-connection. We further reduce the interacting complexity, and achieve channel interacting base on a

Softmax function. This paper uses the following formulation to get attention weights:

$$(W_{att})_c = \text{Softmax}(z)_c = \frac{e^{z_c}}{\sum_{i=1}^C e^{z_i}}, \quad (5)$$

where W_{att} is the channel attention weight same as formula 2.

As shown in the formula 5, the attention weight $(W_{att})_c$ of each channel can be obtained through the dependencies between a single channel (z_c) and all channels (z). Thus, *Softmax* function can be regarded as a weakly-connection among channels. On the contrary, *Sigmoid* obtain the channel weights independently with a lack of interaction. In the experiments, we will make a comparison between these two interaction methods.

The final output of the GCA block is obtained by rescaling X' with the channel weights W_{att} :

$$Y_c = F_{scale}(X'_c, (W_{att})_c) = (W_{att})_c X'_c, \quad (6)$$

where Y_c is the c -th channel of final output, $\mathbf{F}_{scale}(X'_c, (W_{att})_c)$ is a channel-wise multiplication between the weight W_{attc} and the feature map X'_c .

Discussion: To demonstrate the effectiveness of our GCA block, we use ResNet18 and ResNet34 as the backbone networks. We use them based on two reasons: 1) ResNet is a universal backbone, and ResNet18 and ResNet34 have low computational cost. 2) Most attention mechanism blocks have been verified to be effective on the ResNet backbone. The final GCA-Net is stacked by repeated GCA units shown in Fig. 4(b).

4 Experiments

4.1 Dataset and Evaluation Measures

We use a clinical AS-OCT images dataset, which is collected through the CASIA2 ophthalmology device (Tomey Corporation, Japan). The original AS-OCT image is shown as Fig. 1(a). However, only the nucleus area is associated with NC classification [21], and we extracted the nucleus part of the whole AS-OCT image manually as shown in Fig. 1(b)(c)(d).

The AS-OCT image dataset contains 17200 AS-OCT images from 543 participants with the average age of 61.3 ± 18.7 (range: 14~95) years old, and there are 135 males and 335 females among the participants with gender information. The participants were asked to collect images of one eye or both eyes, and the total number of collected eyes is 860 (440 left eyes and 420 right eyes). Each eye has 20 AS-OCT images, and We discarded 999 images without complete nucleus region due to the occlusion of the eyelids during collection. Finally, we use 16201 AS-OCT images for NC classification.

We divide the dataset based on participants into three disjoint subsets: training dataset, validation dataset, and testing dataset. Table 2 summarizes the distribution of three NC stages on the three datasets.

Table 1. The AS-OCT image distribution of NC stages on different datasets.

Dataset	Normal	Mild	Severe
Training	896	3219	5504
Validation	317	793	2331
Testing	390	830	1921
Total	1603	4842	9756

We resize the nucleus images to 224*224 and perform the random rotation and random horizontal flipping for data augmentation. All models are implemented on the Pytorch platform and trained on a TITAN-V GPU with 12GB memory. We use the stochastic gradient descent (SGD) optimizer with the batch size of 64. The initial learning rate is set to 0.0015 and decreased by a factor of 10 every 10 epochs after 100 epochs.

We use three commonly-used evaluation metrics: *Acc*, *F1* and *Kappa* value to evaluate the performance of the model [7]. The calculation formulas are as follows:

$$Acc = \frac{TP + TN}{TP + FP + TN + FN}, \quad (7)$$

$$Recall = \frac{TP}{TP + FN}, \quad (8)$$

$$Precision = \frac{TP}{TP + FP}, \quad (9)$$

$$F1 = 2 \times \frac{Recall \times Precision}{Recall + Precision}, \quad (10)$$

where TP, FP, TN, and FN denote the numbers of true positives, false positives, true negatives, and false negatives, respectively.

$$Kappa = \frac{p_0 - p_e}{1 - p_0}, \quad (11)$$

where p_0 is the relative observed agreement among raters, and p_e is the hypothetical probability of chance agreement. Furthermore, we use $\#P$ to denote the number of parameters and *GFLOPs* [10] to measure the computation.

4.2 Comparison with State-of-art Attention Attention Blocks

Table 2 compares the proposed GCA block with state-of-art attention blocks on ResNet18 and ResNet34. Our GCA-Net achieves the best NC classification results among all methods. It obtains the accuracies of **94.24%** and **94.31%**, respectively, and outperforms state-of-art attention blocks by more than **3%** accuracy. Furthermore, It also consistently improves performance over other methods on *F1* and *Kappa* value, demonstrating the effectiveness of the proposed GCA-Net. Moreover, compared with ResNets and comparative attention-based CNN models, the GCA-Net parameters are equal to ResNets and are

Table 2. Comparison with state-of-the-art attention blocks.

Method	Backbone	Acc	F1	Kappa	#P	GFLOPs
ResNet [8]	ResNet18	91.02	90.98	83.43	11.18M	1.82
SENet [10]		90.61	90.88	82.44	11.27M	1.82
CBAM [22]		89.18	89.03	79.21	11.27M	1.82
GCA-Net		94.24	94.76	89.48	11.18M	1.82
ResNet [8]	ResNet34	88.57	88.57	78.27	21.29M	3.67
SENet [10]		88.35	88.39	77.75	21.44M	3.67
CBAM [22]		91.21	90.93	83.54	21.44M	3.67
GCA-Net		94.31	94.55	89.45	21.29M	3.67

smaller than SENet and CBAM. Furthermore, our GCA-Net does not add additional GFlops through comparisons to other state-of-the-art attention methods. In general, Our GCA-Net works better between accuracy and complexity.

4.3 Ablation Study

Table 3. Effects of pooling operators in GCA based on ResNet18 (✓ denotes using gating operator before squeezing and ✗ denotes not).

Squeeze	Gating	Acc	F1	Kappa
Global max pooling	✓	93.51	93.80	85.95
Global std pooling	✓	93.25	93.59	87.84
Global average pooling	✗	93.64	93.67	88.31
Global average pooling	✓	94.24	94.76	89.48

Effects of Different Pooling Operators. Table 3 shows the classification results of three different pooling operators in the GCA block based on ResNet18. Compared with global max pooling and global std pooling, the GAP achieves the best results on three evaluation measures. This is because GAP can be taken as another representation of average nucleus density (AND) from the nucleus region. Furthermore, the results also demonstrate that the gating operator significantly improves the classification results for the GCA block.

Table 4. Classification results of channel interaction operators in GCA block based on ResNet18.

Squeezing operator	Acc	F1	Kappa
Fully-connection [10]	91.27	91.08	83.41
Local-connection [18]	93.31	93.34	87.77
Non-connection	92.04	92.12	85.44
Weakly-connection	94.24	94.76	89.48

Effect of Different Channel Interaction. Table 4 presents the classification results of four interaction operations: fully-connection, local-connection, non-connection (*Sigmoid*) and weakly-connection (*Softmax*). Our weakly-connection interaction operation obtains the best classification results among four interaction operations. Two reasons can explain these: 1) *Softmax* operation not only sets the relative weights for channels, but also suppresses the unimportant channels. 2) Inter-channel dependencies are weak, and it is difficult to build good dependencies among channels in training.

5 Conclusion

This paper proposes a simple yet effective gated channel attention network named GCA-Net to classify severity levels of nuclear cataract automatically on AS-OCT images. In the GCA-Net, we design a gated channel attention (GCA) block to mask redundant features and use the *Softmax* layer to set relative weights for all channels, which is motivated by the clinical study of average nucleus density (AND). The results on a clinical AS-OCT image dataset demonstrate that our GCA-Net achieves the best classification performance and outperforms advanced attention-based CNN models. Moreover, the computation complexity of our GCA-Net is equal to previous methods, which indicates that it has the potential to deploy our method on the real machine.

In the future, we will collect more AS-OCT images to verify the overall performance of the GCA-Net and plug the GCA block in other CNN models to test its effectiveness.

Acknowledgment. This work was supported in part by Guangdong Provincial Department of Education (2020ZDZX3043, SJJG202002), Guangdong Provincial Key Laboratory (2020B121201001), Shenzhen Natural Science Fund (JCYJ20200109140820699 and the Stable Support Plan Program 20200925174052004).

References

1. Caixinha, M., Amaro, J., Santos, M., Perdigão, F., Gomes, M., Santos, J.: In-Vivo automatic nuclear cataract detection and classification in an animal model by ultrasounds. *IEEE Trans. Biomed. Eng.* **63**(11), 2326–2335 (2016)

2. Cao, L., Li, H., Zhang, Y., Zhang, L., Xu, L.: Hierarchical method for cataract grading based on retinal images using improved Haar wavelet. *Inf. Fusion* **53**, 196–208 (2020)
3. Chen, D., Li, Z., Huang, J., Yu, L., Liu, S., Zhao, Y.E.: Lens nuclear opacity quantitation with long-range swept-source optical coherence tomography: correlation to LOCS III and a Scheimpflug imaging-based grading system. *Br. J. Ophthalmol.* **103**(8), 1048–1053 (2019)
4. Chylack, L.T., et al.: The lens opacities classification system iii. *Arch. Ophthalmol.* **111**(6), 831–836 (1993)
5. Gali, H.E., Sella, R., Afshari, N.A.: Cataract grading systems: a review of past and present. *Curr. Opin. Ophthalmol.* **30**(1), 13–18 (2019)
6. Gao, X., Lin, S., Wong, T.Y.: Automatic feature learning to grade nuclear cataracts based on deep learning. *IEEE Trans. Biomed. Eng.* **62**(11), 2693–2701 (2015)
7. Hao, H., et al.: Open-Appositional-Synechial anterior chamber angle classification in AS-OCT sequences. In: Martel, A.L., et al. (eds.) *MICCAI 2020. LNCS*, vol. 12265, pp. 715–724. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-59722-1_69
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. Hu, J., Shen, L., Albanie, S., Sun, G., Vedaldi, A.: Gather-excite: exploiting feature context in convolutional neural networks. *arXiv preprint arXiv:1810.12348* (2018)
10. Hu, J., Shen, L., Sun, G.: Squeeze-and-excitation networks. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 7132–7141 (2018)
11. Kobayashi, T.: Global feature guided local pooling. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 3365–3374 (2019)
12. Li, H., Lim, J.H., Liu, J., Wong, T.Y.: Towards automatic grading of nuclear cataract. In: *2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 4961–4964. IEEE (2007)
13. Long, E., et al.: An artificial intelligence platform for the multihospital collaborative management of congenital cataracts. *Nat. Biomed. Eng.* **1**(2), 1–8 (2017)
14. Makhotkina, N.Y., Berendschot, T.T., van den Biggelaar, F.J., Weik, A.R., Nuijts, R.M.: Comparability of subjective and objective measurements of nuclear density in cataract patients. *Acta Ophthalmol.* **96**(4), 356–363 (2018)
15. Ozgokce, M., et al.: A comparative evaluation of cataract classifications based on shear-wave elastography and B-mode ultrasound findings. *J. Ultrasound* **22**(4), 447–452 (2019)
16. Panthier, C., Burgos, J., Rouger, H., Saad, A., Gatinel, D.: New objective lens density quantification method using swept-source optical coherence tomography technology: Comparison with existing methods. *J. Cataract Refract. Surg.* **43**(12), 1575–1581 (2017)
17. Park, J., Woo, S., Lee, J.Y., Kweon, I.S.: Bam: bottleneck attention module. *arXiv preprint arXiv:1807.06514* (2018)
18. Qilong, W., Banggu, W., Pengfei, Z., Peihua, L., Wangmeng, Z., Qinghua, H.: ECA-Net: Efficient channel attention for deep convolutional neural networks (2020)
19. Qin, Z., Zhang, P., Wu, F., Li, X.: Fcanet: frequency channel attention networks. *arXiv preprint arXiv:2012.11879* (2020)
20. Wang, W., et al.: Objective quantification of lens nuclear opacities using swept-source anterior segment optical coherence tomography. *Br. J. Ophthalmol.* (2021)

21. Wong, A.L., et al.: Quantitative assessment of lens opacities with anterior segment optical coherence tomography. *Br. J. Ophthalmol.* **93**(1), 61–65 (2009)
22. Woo, S., Park, J., Lee, J.Y., Kweon, I.S.: CBAM: convolutional block attention module. In: Proceedings of the European Conference on Computer Vision (ECCV), pp. 3–19 (2018)
23. Xu, X., Zhang, L., Li, J., Guan, Y., Zhang, L.: A hybrid global-local representation CNN model for automatic cataract grading. *IEEE J. Biomed. Health Inform.* **24**(2), 556–567 (2019)
24. Xu, Y., Duan, L., Wong, D.W.K., Wong, T.Y., Liu, J.: Semantic reconstruction-based nuclear cataract grading from slit-lamp lens images. In: Ourselin, S., Joskowicz, L., Sabuncu, M.R., Unal, G., Wells, W. (eds.) MICCAI 2016. LNCS, vol. 9902, pp. 458–466. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46726-9_53
25. Xu, Y., et al.: Automatic grading of nuclear cataracts from slit-lamp lens images using group sparsity regression. In: Mori, K., Sakuma, I., Sato, Y., Barillot, C., Navab, N. (eds.) MICCAI 2013. LNCS, vol. 8150, pp. 468–475. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-40763-5_58
26. Zhang, X., Fang, J., Hu, Y., Xu, Y., Higashita, R., Liu, J.: Machine learning for cataract classification and grading on ophthalmic imaging modalities: a survey. arXiv preprint [arXiv:2012.04830](https://arxiv.org/abs/2012.04830) (2020)
27. Zhang, X., et al.: A novel deep learning method for nuclear cataract classification based on anterior segment optical coherence tomography images. In: 2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC), pp. 662–668. IEEE (2020)