Journey into Filter Size and Its Effects, Limitations and Challenges Exploratory Study

Mohamed Aboukhair[a], Abdelrahim Koura[a], and Mohammed Kayed[a]

[a] Computer Science Department, Faculty of Computers and Artificial Intelligence, Beni-Suef University, Beni Suef 62511, Egypt

Abstract

This research delves into the investigation of the impact of filter size, deviating from the conventional notion of preferring smaller filter sizes (3x3), with the aim of shifting the focus towards the potential of larger filters and encouraging researchers to invest more effort in exploring their capabilities. The effectiveness of larger filters will increase as computational power increase, however, currently there exists no CNNs model that achieves optimal performance with bigger filters. Therefore, it is imperative to consider the necessity of integrating larger filter sizes in our models, as computing power continues to advance in the future. As advancements in computer power continue to occur and larger image sizes become more prevalent, the obstacles that have traditionally hindered researchers from utilizing larger filter sizes will diminish. This will lead to increased exploration and study of larger filter sizes in the future. Our role is to provide guidance by elucidating the effects, limitations, and challenges that must be addressed in order to capitalize on the unique opportunities and broader potential offered by larger filters. To the best of our knowledge, we have identified three such opportunities and conducted an analysis of filter size impact using all previous research that were related to filter size effects comparison.

Keywords: CNN, Classification, Filter Size Effects, CNN structures

1 Introduction

* what is the significant about computer vision?

The advent of ample processing resources and vast data sets has cleared the path for the success of deep learning, elevating the importance of computer vision. This cutting-edge technology possesses immense value across various fields, including automated driving, object recognition, and image analysis. Through the study of computer vision, we can optimize manufacturing processes and spearhead innovative technological advancements[1].

* how much the researchers and industry foucs on computer vision?

There has been a noteworthy escalation in the number of papers presented at machine learning conferences wherein at least one author had affiliations with the industry. Specifically, the percentage of industry-affiliated papers at the International Conference on Machine Learning (ICML) increased from 20-25% in 2017 to 45% in 2018. Moreover, the number of publications by researchers from Google's DeepMind alone grew twofold from 6% in ICML'17 to 13% in ICML'18. Although the precise ratio of industry-sponsored research in computer vision remains unspecified, the immense success of deep learning has facilitated the growth of computer vision applications, resulting in significant investments by numerous companies in research and development in this domain. Industry-sponsored researchers have emerged as a prominent component of the computer vision community and have made significant contributions to all major conferences and journals in this field[2].

* how much importance does deep learning offer in computer vision?

The proliferation of deep learning in computer vision has supplanted conventional machine learning algorithms and engenders pronounced benefits in feature extraction, rendering it ubiquitous in computer vision, as well as in other domains such as physics, biology, and manufacturing. Notably, the paper identifies several prevalent deep learning algorithms in computer vision, including dropout, convolutional neural networks (CNN), full convolutional networks (FCN), and generative adversarial networks (GAN). Consequently, deep learning has assumed a critical role in computer vision and is a highly sought-after research area in both academic and industrial settings[3].

The application of deep learning has yielded remarkable successes in various domains, including computer vision, natural language processing, and video/speech recognition. In the field of computer vision, deep learning has facilitated notable advancements in tasks such as image classification, object detection, and segmentation. In the realm of natural language processing, deep learning has been utilized for tasks such as sentiment analysis, machine translation, and text generation. In the domain of video/speech recognition, deep learning has enabled significant progress in tasks such as speech recognition, speaker identification, and emotion recognition. It is noteworthy that the significance of deep learning in these areas is contingent upon the specific context and field in which it is employed[4].

* how much importance does filter has in CNN Architectures ?

Kernels or filters are essential components of deep learning architectures, especially Convolutional Neural Networks (CNNs), which use a variety of kernels (filters) in convolutional layers to convolve the entire picture and intermediate feature maps, yielding different feature maps. The filters are used to extract features from input pictures, and their size might affect the network's computational efficiency. Szegedy et al. (2016), for example, advocated replacing filters with a size of 5x5 (7x7) with two stacked 3x3 filters to minimize the number of parameters and computational costs while keeping the receptive field size the same. As a result, filters play a critical role in the performance and efficiency of CNNs[4].

This presents a thought-provoking yet insufficiently examined matter: the facility to utilize larger filters within our deep learning models. To the best of our knowledge, it appears that the use of 3x3 filters in convolutional neural network (CNN) designs is preferred. This preference extends to original models such as Inception, which incorporate larger filters within their modules. Subsequent studies have demonstrated that the substitution of bigger filters with two stacked 3x3 filters is more effective, further solidifying the dominance of 3x3 filters in all CNN models. The emphasis on 3x3 filters in convolutional neural network (CNN) designs is not without tradeoffs. While the use of larger filter sizes may lead to better learners, the associated increase in the number of parameters presents challenges for the optimizer. Overfitting is also a concern with larger filters, which are known to be powerful learners. Moreover, the exponential growth in parameters necessitates increased processing capacity, further complicating matters.

The present research endeavors to elucidate the challenges and drawbacks entailed in the application of larger filter sizes in convolutional neural network (CNN) architectures by conducting a comprehensive review of relevant literature. In addition, this study seeks to assess the benefits and limitations of larger filters and identify exceptional prospects by addressing the impediments and constraints associated with their use.

The contributions of this paper are summarized as follows:

- An exploration analysis of filter size effects on CNN
- Bigger filters challenges and opportunities

The remaining sections of this paper are organized as follows: Section 2 presents a comprehensive review of the previous literature relevant to our work. In Section 3, we discuss the output from the research that we targeted and show the rare opportunities that a bigger filter can offer. Finally, in Section 4, we offer concluding remarks summarizing the key findings and contributions.

2 Previous Literature

To the best of our knowledge, there is no research that shows all previous work of filter size effects on CNN models as exploratory research to show the challenges of using bigger filters as well as possible opportunities While it is not the richest topic of research yet still one of the most important topics in the deep learning field. we selected all research that meets two Rules: first, the research must have a comparison of filter sizes, and second, any research with structure and modules that contain bigger filters.

2.1 Preliminaries of CNN filter

[A Survey of the Recent Architectures of Deep Convolutional Neural Networks]Convolutional Neural Network (CNN) is a type of neural network that is widely used in machine learning, particularly in vision-related applications. It is known for its ability to learn representations from grid-like data, such as images, and has shown significant performance improvement in various machine-learning tasks. CNNs are composed of multiple layers, including convolutional layers, non-linear processing units, and subsampling layers, which help in extracting useful features from data and classifying them.

[Review of deep learning: concepts, CNN architectures, challenges, applications, future directions]CNNs are inspired by the structure of neurons in human and animal brains and are designed to automatically identify relevant features in input data without human supervision. They consist of multiple layers, including convolutional layers that use convolutional filters to extract features from input images.

[A Survey of the Recent Architectures of Deep Convolutional Neural Networks]CNNs have been applied to visual tasks since the late 1980s. In 1989, LeCuN et al. proposed the first multilayered CNN named ConvNet, which was based on Fukushima's Neocognitron. This marked the foundation for modern 2D CNNs. LeCuN's work introduced supervised training using the backpropagation algorithm, and ConvNet showed successful results in optical character and fingerprint recognition tasks. It was further improved with the introduction of LeNet-5, which played a significant role in classifying characters in document recognition applications.

[A Survey of the Recent Architectures of Deep Convolutional Neural Networks]Deep CNNs faced challenges in the early 2000s due to complex architecture and limited hardware resources. However, the use of activation functions like RelU and advancements in hardware, such as GPUs, revived the research in CNNs.

The availability of large image databases like ImageNet and platforms like Kaggle and Colab further accelerated CNN research.

[A Survey of the Recent Architectures of Deep Convolutional Neural Networks]A CNN consists of several essential components that enable it to learn representations from grid-like data and perform tasks such as feature generation and classification. These components include convolutional layers, non-linear processing units, and subsampling layers.

[A Survey of the Recent Architectures of Deep Convolutional Neural Networks]Convolutional Layers: These layers perform the convolution operation, which helps extract useful features from locally correlated data points. Convolutional kernels(filters) are used to scan the input data and generate feature maps.

[A Survey of the Recent Architectures of Deep Convolutional Neural Networks]Non-linear Processing Units: After the convolution operation, the output is passed through a non-linear activation function. This function introduces non-linearity into the feature space and helps in learning abstractions and semantic differences in images.

[A Survey of the Recent Architectures of Deep Convolutional Neural Networks]Subsampling Layers: Subsampling layers summarize the results obtained from the previous layers and make the input invariant to geometrical distortions. They help reduce the dimensionality of the feature maps and capture the most important information.

[A Survey of the Recent Architectures of Deep Convolutional Neural Networks]These components work together in a hierarchical manner, allowing CNN to learn hierarchical representations of the input data. The hierarchical learning, automatic feature extraction, and weight-sharing capabilities of CNNs contribute to their effectiveness in various machine learning applications, especially in vision-related tasks.

[Review of deep learning: concepts, CNN architectures, challenges, applications, future directions] A filter or kernel in a Convolutional Neural Network (CNN) is a small matrix of weights that is convolved with the input image to extract features. Filters or kernels play a crucial role in Convolutional Neural Networks (CNNs) as they extract features from input data. These filters slide over the input image, performing a dot product operation to generate an output feature map. The weights of the filters are adjusted during the training process, allowing them to learn to extract significant features. The use of filters in CNNs offers several benefits. Firstly, filters enable weight sharing, which reduces the number of trainable network parameters. This weight-sharing feature helps enhance generalization and avoid overfitting. Secondly, filters allow the CNN model to automatically identify relevant features without human supervision. This automatic feature identification is a key advantage of CNNs compared to their predecessors.

[Understanding of a Convolutional Neural Network; Review of deep learning: concepts, CNN architectures, challenges, applications, future directions] The filters in CNNs are designed to detect specific patterns or features in the input data. For example, a filter may be designed to detect edges, corners, or textures. By applying multiple filters to the input data, CNNs can learn to recognize complex patterns and features. The output of the convolutional layer is a set of feature maps, each representing the activation of a specific filter. These feature maps are then passed through non-linear activation functions to introduce non-linearity and enhance the network's ability to learn complex relationships.

[Understanding of a Convolutional Neural Network] filters or kernels values are the CNN Weights and because CNNs supports Weight sharing which is a key feature of CNNs. It refers to the practice of using the same set of weights across different spatial locations in the input data. By sharing weights, CNNs are able to achieve translation invariance, meaning they can detect and recognize features regardless of their positions in the image. This is particularly useful in computer vision tasks, where the location of features may vary.

The choice of filters is an important design consideration in CNN architecture, as it directly impacts the network's ability to extract meaningful information and achieve accurate results.

[A Survey of the Recent Architectures of Deep Convolutional Neural Networks]Different sizes of filters can be used to capture different levels of granularity in the data. Small-size filters are effective in extracting fine-grained details, while large-size filters capture coarse-grained information. By adjusting the filters, CNNs can perform well on both coarse and fine-grained details, improving their overall performance.

[Review of deep learning: concepts, CNN architectures, challenges, applications, future directions] Bigger filter sizes in convolutional neural networks (CNNs) can have several effects. Firstly, they increase the receptive field of the network, allowing it to capture larger spatial patterns in the input data. This can be beneficial for tasks that require an understanding of global context or larger structures. Bigger filters can increase the receptive field of the network. The receptive field refers to the area of the input that influences the activation of a particular neuron. By using bigger filters, the network can take into account a larger context when making predictions, which can improve its overall performance. Bigger filter sizes can lead to increased model complexity and computational requirements. The number of parameters in the network increases with larger filters, which can make training and inference more computationally expensive.

2.2 Key Challenges

The only significant challenges to filter size existed in the usage of larger filters, while smaller filters are advised for a variety of reasons. This section illustrates the primary issues that models encounter when using larger filter sizes, clarifying the restrictions while describing the challenges.

Output Dimension Challenge

One of the problems that may arise in using bigger filters is Dimension reduction leading to limiting the depth model that can reach with usage of bigger filter size That happens because of filter size. filter size factor can be removed by using padding yet this will lead to overpadding the higher we go in-depth which puts Dimension Challenge as one of the limitations that lash back the usage of bigger filters leading to prefer 3x3 filter size that allows for going deeper in depth which benefits the accuracy of the model.[cite paper that supports depth and paper that supports filter size reduction and pooling]

Computational Power Challenge

computational power problem is one deal coming with the usage of bigger filter size due to the increase of time and model complexity. time complexity increases because of the increase of parameters of the model due to the filter being based on numbers which increases when the filter size increases. the problem of time complexity does not only affect the time needed to train the model it also leads to bigger filters to get poor scores[5].

Overfitting Challenge

Overfitting is more likely to occur with complex models that have more flexibility as bigger filters increase model complexity, they can easily help overfitting to occur because of the increase of model complexity[6, 7]. this can be handled by many means like dropout and regularization yet it is unexplored enough as there are few model architectures that use bigger filters as part of their Structure. As far as we know there is no technique against overfitting that targets the use of bigger filters, all techniques are generalized to target overfitting yet one way to lower the chances of overfitting is by lowering the complexity in other words even if bigger filters are implemented, overfitting may highly occur which make this one of hardest challenge.

2.3 Filter size effects

As we said in previous sections the research topic of filter size effects is not well explored which leads only to six research about comparing filter size effects as far as we know and almost all of them assume that the use of small filters is better. Our research will use their analysis plots to present the bigger picture as we have no access to their results data.

Paper 1

the first paper is “Analysis of Filter Size Effect In Deep Learning” by Camgözlü and Kutlu in 2021. the research presents the effects of three different filter sizes (3x3;5x5;9x9) in two different ways on the Mendeley data-set, the first way is every filter size standalone in the model which means each model has only one filter. the second way is way more interesting as the baseline in this research is six convolutional layers, the researcher increases or decreases the filter size every two convolutional layers leading to the use of 3x5x9 filter size in this order or in 9x5x3 order[8].

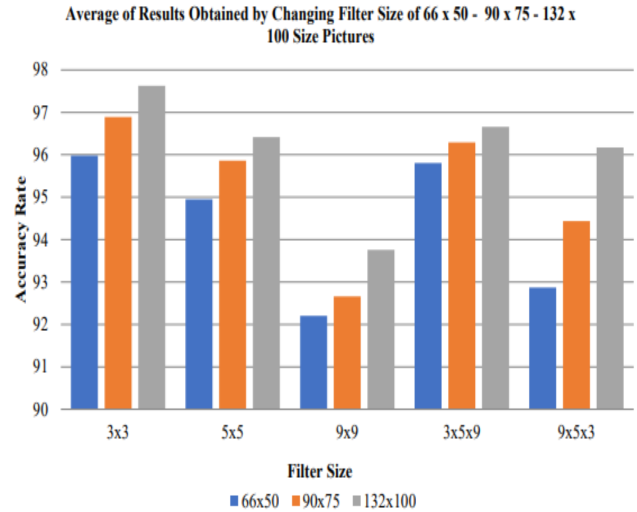

Figure 1: Accuracy Over Filter Size with Images Sizes

From Figure 1, we can notice how filter size affects the accuracy in many different ways, the first thing that will be noticed is how good is 3x3 filter size, second thing to notice is how good a bigger image size affects the model accuracy, third thing to notice that using a combination of incremental filter sizes is the second best results which gives more promising results than bigger filter sizes alone and the interesting part is last thing to notice if not ignored that using a combination of decremental filter sizes has huge difference between every image size more than any other approach. the last notice is interesting because it does not show the effects of decremental filter sizes in the model but it shows how good bigger filter potential is.

From Figure 2, we can notice how the filter size and image size affect the time complexity in different ways, the first thing to notice is how good 3x3 filter size in time complexity as well as smaller image sizes, the last thing to notice is how the order in multi-filter sizes affect time complexity which is logical as running a bigger filter size on the input will take more time leading to second worst results in time complexity to be decremental filter sizes in the model.

From Figure 3, we can notice how the filter size affects the parameter number of the model, the pars represents the parameter number while the line represents the accuracy of each approach. we notice that the researcher unintentionally made a couple of illusions in this figure. The first illusion image size does not affect the parameter number of the model, it only affects the time that takes the number of filters to be applied to the image but the fully connected layer parameters do depend on the dimension of input that reaches it which made the illusion of different parameter number for each im-

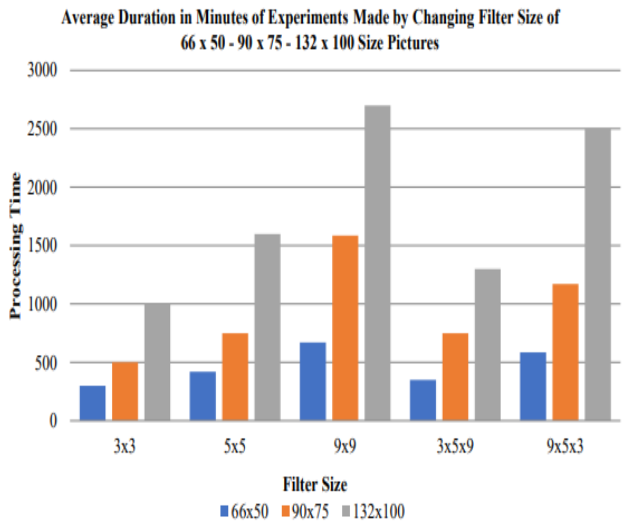

Figure 2: Time over Over Filter Size with Images Sizes

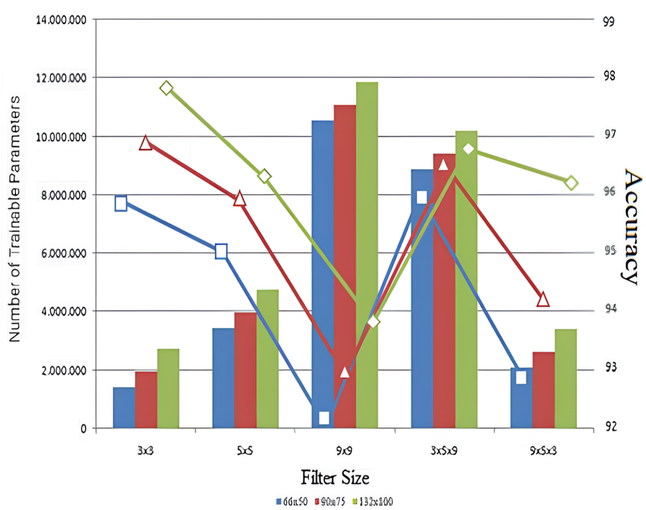

Figure 3: Accuracy Over Parameters Over Filter Size

age size that clearly appears in the difference between each image size in each approach which is the same difference in all approaches. The second illusion is fascinating as 3x5x9 has lower time than 9x5x3 from Figure 2 yet 3x5x9 has way higher parameters than 9x5x3. this can be explained by the illusion of using too many filter numbers the more we go in-depth causing the last layers to be dense in parameters which cause a 9x9 filter in the last layers to have more parameters yet work with smaller dimension in parallel resulting to be faster than 9x9 filter working with bigger input with lower filter number, yet the researcher does not supply the structure that was used but it is the only explanation.

Paper 2

the second paper is "Effect of filter sizes on image classification in CNN: a case study on CFIR10 and Fashion-MNIST datasets" by Khanday, Dadvandipour, and Lone in 2021 [9]. The research presents the effects of different filter sizes (3x3;5x5;7x7) in two datasets(CFIR10 and Fashion-MNIST).

Filter Size	Training Data	Validation Data	Test Data
3×3	0.942625	0.7275	0.7304
5×5	0.923275	0.7261	0.7297
7×7	0.87725	0.7067	0.635

Table 1: dataset-Cifar10-acc

Filter size	Training data	Validation data	Test data
3×3	0.929	0.9235	0.9268
5×5	0.926	0.9196	0.9264
7×7	0.918	0.910	0.911

Table 2: dataset-FashionMNIST-acc

From Table [1;2], we notice only that the higher the filter size the worse results get as well as filter 5x5 is the nearest to 3x3 filter size from the accuracy perspective leading to the common thought that 3x3 filter size is better than 5x5 filter size but in fact, this is not the case as the size of the image in Fashion-MNIST is 28x28 and the size of the image in CFIR10 is 32x32 which means that the effects of bigger filters were tested on the smaller image which gives a huge advantage for small filter size to get higher results easily.

Paper 3

the third paper is "The Impact of Filter Size and Number of Filters on Classification Accuracy in CNN" by Ahmed and Karim in 2020 [10]. The research presents the effects of different filter sizes (3x3;5x5) with different numbers of each filter being applied in the convolutional layer.

From Table 3, we notice the same pattern of favoring smaller filter sizes because of the input image size which is 96x96 yet in this research we can see a different perspective which is the number of filters, the results improve with a higher filter number in the case of smaller filter 3x3 size yet in the 5x5 filter size it has drop-down then improving but under its best score. the effects of a bigger filter size are indeed overfitting and increasing the number of filters will just make the output worse than its lower filter number original state. the number of filters of 128,256 is still better than the number of filters of 64,128 which somehow decreases the effect of overfitting yet it is lower than the number of filters of 32,64.

From Table 4, we notice that the number of filters affects the time complexity exponentially, and the ratio of time increases with higher filter number increase even more with higher filter size.

Dataset	Convolution layer	Number of filters	Filter Size	Train accuracy	Test accuracy	Loss value
KTH	1, 2	32, 64	3×3	0.999%	97.64%	0.002
	1, 2	64, 128		0.999%	97.96%	0.003
	1, 2	128, 256		0.998%	98.09%	0.005
	1, 2	32, 64	5×5	0.997%	97.58%	0.007
	1, 2	64, 128		0.999%	97.33%	0.002
	1, 2	128, 256		0.998%	97.46%	0.005

Table 3: accuracy over filter size

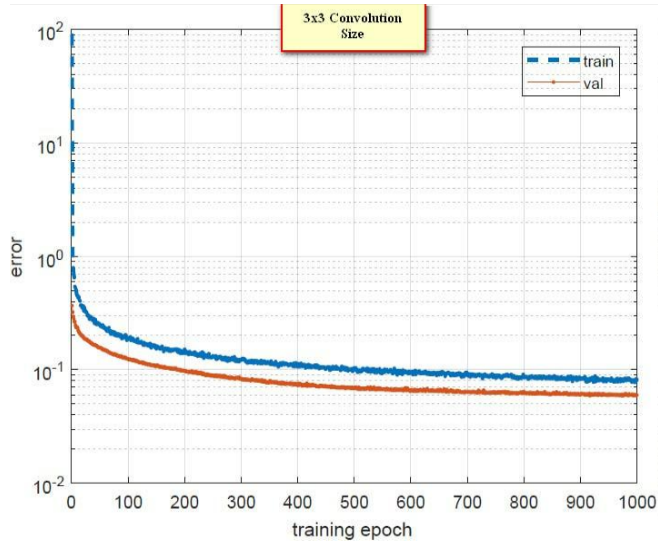

Figure 4: 3x3 train vs val loss

Dataset	Number of filters	Filter Size	Training Time in second per epoch
KTH	32	3×3	38 s
	64		96 s
	128		259 s
	32	5×5	67 s
	64		171 s
	128		516 s
	32	7×7	160 s
	64		428 s
	128		1400 s

Table 4: time over filter size

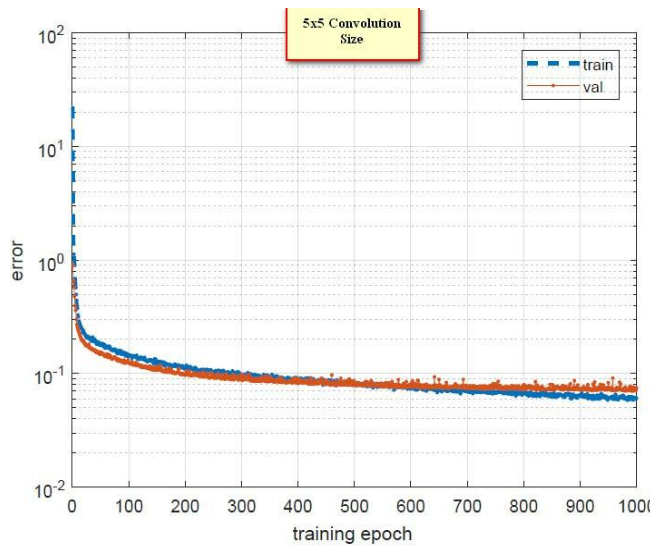

Figure 5: 5x5 train vs val loss

Paper 4

the forth paper is "Convolution Kernel Size Effect on Convolutional Neural Network in Histopathological Image Processing Applications" by Ozturk et al. in 2018 [11]. The research presents the effects of different filter sizes (3x3;5x5;7x7;9x9) in the training process on The CAMELYON17 challenge dataset which has the second biggest image size (128x128) in all previous papers.

From Figure [4;5], we notice that the 3x3 filter has a stable curve where the line of validation loss is closing to train loss which indicates stable results which is 6.1% error in the other side 5x5 had stable results then overfitting take over after 500 epoch leading into a decrease in loss of train lower than validation loss which is effect of overfitting resulting into 7.5% error which is expected from overfitting model.

From Figure [6;7], we notice that the 7x7 filter has a stable curve in factor of total epochs but has dropouts cause a little mess in the line yet the curve of training and validation loss are closing to each other better than the 3x3 filter from Figure 4 resulting into 6% error, a better error than the 3x3 filter which shows our point of view that bigger filter sizes have undiscovered potential as well as the greatest danger which is overfitting like in 9x9 filter which has to overfit after 100 epoch easily leading to 9.1% error. this effect supports the potential of bigger filter sizes leading to the support of bigger filter sizes being applied and tested with bigger image sizes.

Paper 5

the fifth paper is "Convolution Neural Networks and Impact of Filter Sizes on Image Classification" by Khanday and Dad-

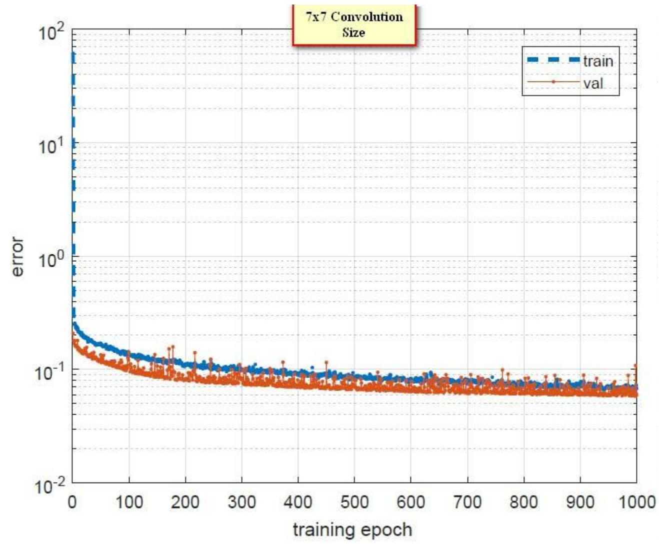

Figure 6: 7x7 train vs val loss

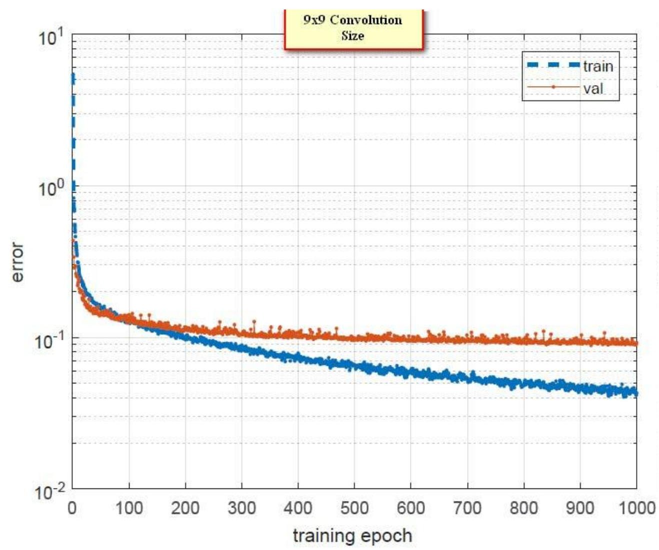

Figure 7: 9x9 train vs val loss

vandipour in 2020 [12]. The research presents the effects of different filter sizes (3x3;5x5;7x7) in the digits MNIST dataset and gets similar results to paper 2 favoring a smaller filter size with 1% improvement while using an image size of 28x28.

Paper 6

the sixth paper is “Filter size optimization on a convolutional neural network using FGSA” by Poma et al. in 2020 [13]. The research presents the FGSA algorithm (The fuzzy gravitational search algorithm) to automatically search for the best filter size that increases the Recognition rate in the test set in the face Recognition task, the only limitation is the need for a small convolution model for the method to not need high computational power. the FGSA is used on the CROPPED YALE database(380 records) which has an image size of 640(w) x 480 (h). The results were very interesting as the best filter size is 9x9. The effect of favoring a bigger filter is not a coincidence as the image size is too big which asked for a bigger filter to extract the information.

3 Discussion

3.1 Filter size effects

In the previous section Filter size effects, we present all the previous work related to comparing filter sizes showing the favoritism of 3x3 filter size while explaining why 3x3 filter size is tested on cases that favor its use in the first place. In this section, we will discuss these effects with the possibility of improving the test cases while showing collective analysis. From Figure 3, the image size is too limited to a max size of 132x100 which gives an advantage to smaller filter sizes and also why 7x7 was not used or shown in the experiment as from figure 6 it shows 7x7 filter size being applied with slightly bigger image size and has more promising results than 5x5 filter size which shows that test missed two important factors to judge bigger filters size to have better or worse results. From Table 4, the effect of increment of the time complexity exponentially can explain also the second illusion of 9x5x3 with high time complexity yet has a lower parameter number within Figure 3 mentioned in the previous section of Paper1[11, 8].

3.2 Bigger filter Architecture

The Inception module, proposed in the seminal work ”Going Deeper with Convolutions,”[14] introduced a novel structure incorporating multiple filters. However, it should be clarified that the Inception module was not designed to function as a standalone layer or replace individual layers in a model. It faced challenges related to computational limitations[15] and managing output size growth associated with 5x5 Convolutional layers, which were addressed through input reduction strategies. Furthermore, the Inception module was recommended for utilization primarily in the upper layers of a model. Subsequent research, exemplified by ”Rethinking the Inception Architecture for Computer Vision,”[16] aimed to enhance the module by substituting 5x5 Convolutional layers with two consecutive 3x3 Convolutional layers. This modification aimed to improve efficiency and effectiveness while maintaining the overall module architecture. Nevertheless, it revealed that the original Inception module was not intended to incorporate filter sizes larger than 3x3. Further exploration and development of alternative approaches, such as the proposed multi-filter layer in this research, are crucial to address the limitations and challenges associated with the use of multiple filter sizes.

3.3 Computational Power Increase

3.4 Opportunities

Capture larger spatial patterns

The utilization of larger filter sizes has the potential to capture larger patterns, which is advantageous for tasks that require the perception of global context or larger objects. Larger filter sizes have the potential to yield significant improvements in results within the appropriate context. In order to fully capitalize on these benefits, it is necessary to solve the challenges and limitations associated with larger filters.

Stronger Learner

As technology advances, the utilization of strong learners without the constraints of overfitting and computational power can greatly assist in enhancing deep learning models. The need for more powerful learners opens up vast potential for deep learning models in various applications. Even a slight increase in accuracy, such as 1%, can be appreciated in fields like medicine, where it can assist in diagnosing patients more accurately and potentially saving more lives.

Handling Bigger Image Better

The issue of image input size limitations is often overlooked or disregarded by machine learning engineers. While past machine learning engineers constraints may have forced this issue, new advancements in technology, such as TPUs, have increased image input size limits by over tenfold. However, the exploration and implementation of larger image sizes is still lacking due to limitations and challenges, including difficulty in handling larger image sizes with smaller filter sizes. Despite the benefits of utilizing larger images, such as increased information extraction and reduced pixelation and data loss from downsampling, deep learning users often resort to downsampling as a recommended solution to circumvent issues. Two significant challenges of utilizing larger images include limited vision with 3x3 filters as image size increases and exponential increases in time complexity for model processing.

4 Conclusion

We are stuck with 3x3 filters for bigger filter limitations and this is also why we are stuck with using smaller image sizes because we can not use bigger filter sizes. We present the challenges and limitations that come with the use of bigger filter sizes in the form of problems that need to be solved to access the use of bigger filter sizes. We analyze the effects of filter size showing bias towards smaller filter sizes in the test case design as lowering image size or skipping certain filter size lead to bias that favor the filter size 3x3 over bigger filter sizes. we show the possible benefits and Opportunities that bigger filter sizes have to offer which is a stronger learner, the ability to capture larger spatial patterns, and handle bigger image sizes better. Opportunities present themselves for a solution that neutralizes the challenges that the use of bigger filters requires to solve.

References

[1] Ahmed Fathi Bekhit. "Introduction to Computer Vision". In: *Computer Vision and Augmented Reality in iOS: OpenCV and ARKit Applications*. Berkeley, CA: Apress, 2022, pp. 1–20. ISBN: 978-1-4842-7462-0. DOI: 10.1007/978-1-4842-7462-0_1. URL: https://doi.org/10.1007/978-1-4842-7462-0_1.

[2] Iuliia Kotseruba, Manos Papagelis, and John K. Tsotsos. "Industry and Academic Research in Computer Vision". In: *CoRR* abs/2107.04902 (2021). arXiv: 2107.04902. URL: https://arxiv.org/abs/2107.04902.

[3] Qing Wu et al. "The application of deep learning in computer vision". In: *2017 Chinese Automation Congress (CAC)*. 2017, pp. 6522–6527. DOI: 10.1109/CAC.2017.8243952.

[4] Junyi Chai et al. "Deep learning in computer vision: A critical review of emerging techniques and application scenarios". In: *Machine Learning with Applications* 6 (2021), p. 100134. ISSN: 2666-8270. DOI: https://doi.org/10.1016/j.mlwa.2021.100134. URL: https://www.sciencedirect.com/science/article/pii/S2666827021000670.

[5] Owais Khanday and Samad Dadvandipour. "Convolution Neural Networks and Impact of Filter Sizes on Image Classification". In: *Multidiszciplináris Tudományok* 10 (Apr. 2020), pp. 55–60. DOI: 10.35925/j.multi.2020.1.7.

[6] Osval Antonio Montesinos López, Abelardo Montesinos López, and Jose Crossa. "Overfitting, model tuning, and evaluation of prediction performance". In: Cham: Springer International Publishing, 2022, pp. 109–139.

[7] Matias Valdenegro-Toro and Matthia Sabatelli. "Machine Learning students overfit to overfitting". In: (2022).

[8] Yunus Camgözlü and Yakup Kutlu. "Analysis of Filter Size Effect In Deep Learning". In: *CoRR* abs/2101.01115 (2021). arXiv: 2101.01115. URL: https://arxiv.org/abs/2101.01115.

[9] Owais Khanday, Samad Dadvandipour, and Mohd Aaqib Lone. "Effect of filter sizes on image classification in CNN: a case study on CFIR10 and Fashion-MNIST datasets". In: *IAES International Journal of Artificial Intelligence (IJ-AI)* 10 (Dec. 2021), p. 872. DOI: 10.11591/ijai.v10.i4.pp872-878.

[10] Wafaa Ahmed and Abdulamir Karim. "The Impact of Filter Size and Number of Filters on Classification Accuracy in CNN". In: (Apr. 2020), pp. 88–93. DOI: 10.1109/CSASE48920.2020.9142089.

[11] Saban Ozturk et al. “Convolution Kernel Size Effect on Convolutional Neural Network in Histopathological Image Processing Applications”. In: *2018 International Symposium on Fundamentals of Electrical Engineering (ISFEE)*. IEEE, Nov. 2018. DOI: 10.1109/isfee.2018.8742484. URL: https://doi.org/10.1109/isfee.2018.8742484.

[12] Owais Khanday and Samad Dadvandipour. “Convolution Neural Networks and Impact of Filter Sizes on Image Classification”. In: *Multidiszciplináris Tudományok* 10 (Apr. 2020), pp. 55–60. DOI: 10.35925/j.multi.2020.1.7.

[13] Yutzil Poma et al. “Filter size optimization on a convolutional neural network using FGSA”. In: *Intuitionistic and Type-2 Fuzzy Logic Enhancements in Neural and Optimization Algorithms: Theory and Applications*. Studies in computational intelligence. Cham: Springer International Publishing, 2020, pp. 391–403.

[14] Christian Szegedy et al. “Going Deeper with Convolutions”. In: *CoRR* abs/1409.4842 (2014). arXiv: 1409.4842. URL: http://arxiv.org/abs/1409.4842.

[15] Sergey Ioffe and Christian Szegedy. “Batch Normalization: Accelerating deep network training by reducing internal covariate shift”. In: (2015).

[16] Christian Szegedy et al. “Rethinking the Inception Architecture for Computer Vision”. In: *CoRR* abs/1512.00567 (2015). arXiv: 1512.00567. URL: http://arxiv.org/abs/1512.00567.