Gabor Filters as Initializers for Convolutional Neural Networks: A Study on Inductive Bias and Performance on Image Classification

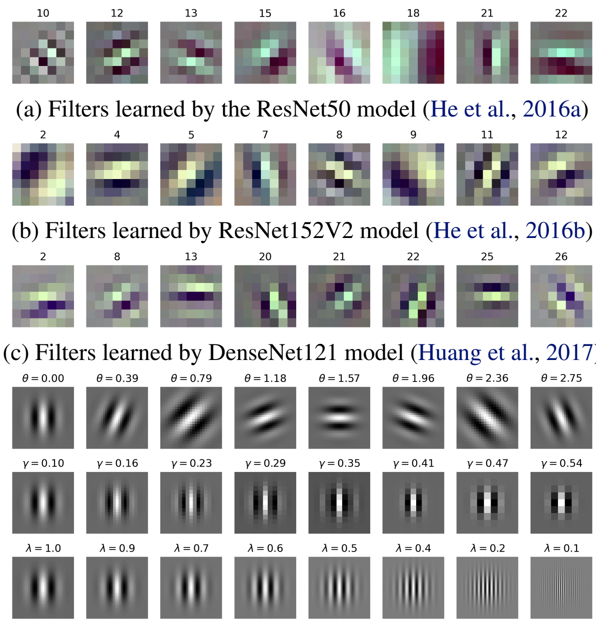
Pablo Rivas [1 2] **Mehang Rai** [3]

Abstract

This study explores the impact of Gabor filters on the performance of Convolutional Neural Networks (CNNs) in image classification tasks. Prior research has indicated that the receptive filters of CNNs often resemble Gabor filters, suggesting their potential as initial receptive filters. We conducted an extensive analysis on a variety of general object datasets, demonstrating that the integration of Gabor filters in the receptive layer enhances CNN performance, as evidenced by improved accuracy, higher Area Under the Curve (AUC), and reduced loss. Furthermore, our findings suggest that CNNs equipped with Gabor filters in the receptive layer can achieve superior performance in a shorter training period compared to traditional random initialization techniques.

1. Introduction

Image understanding has seen significant advances with the advent of neural networks, particularly CNNs and vision transformers (ViTs) (Behnke & Rojas, 1998; Gaussier & Cocquerez, 1992; Sigillito et al., 1991; Reis et al., 2019; Wu et al., 2020; Li et al., 2021; Han et al., 2022). Despite numerous studies on enhancing CNN performance, few have focused on improving their initialization, which is crucial due to issues like vanishing gradients and local optima (Ide & Kurita, 2017; Li et al., 2016). Before CNNs, Gabor filters were used for image processing, extracting texture information and aiding in segmentation (Bai et al., 2019). The receptive field convolutional layer in a CNN, which closely resembles Gabor filters, is critical for the network's performance (Xi et al., 2018; Krizhevsky et al., 2012). However, these layers are often randomly initialized, necessitating significant adjustment through gradient descent.



(a) Filters learned by the ResNet50 model (He et al., 2016a)

(b) Filters learned by ResNet152V2 model (He et al., 2016b)

(c) Filters learned by DenseNet121 model (Huang et al., 2017)

(d) Gabor filters w/ varying λ, θ, and γ (Rai & Rivas, 2020).

Figure 1: Comparison of convolutional filters learned in the receptive field by general-purpose object recognition networks (a)-(c), and Gabor filters with different parameters (d).

In this study, our focus is on initializing the receptive convolutional layer with the Gabor filter, as previous research has demonstrated that only the receptive filters resemble the properties of Gabor filters, see Figure 1. Our study explores the use of Gabor filters for initializing the receptive field of CNNs, potentially improving performance and efficiency (Alekseev & Bobe, 2019). Unlike previous work, we impose no constraints on the Gabor filter structure, allowing the CNN to modify it as needed for more complex feature extraction. This approach can enhance CNN performance, as evidenced by our contributions:

- Improved CNN performance using Gabor filters for object classification in terms of accuracy, AUC, and

[1]Department of Computer Science, Baylor University, Texas, USA [2]Center for Standards and Ethics in Artificial Intelligence, Texas, USA [3]Brazos Innovation Partners, LLC, Texas, USA. Correspondence to: Pablo Rivas <Pablo_Rivas@Baylor.edu>.

Proceedings of the 40^{th} International Conference on Machine Learning, Honolulu, Hawaii, USA. PMLR 202, 2023.

Table 1: Summary of significant research applications of CNNs and Gabor filters within the last seven years

Applications	**References**
Object recognition	(Yao et al., 2016; Alekseev & Bobe, 2019; Luan et al., 2018; Liu et al., 2018)
Age and Gender classification	(Hosseini et al., 2018)
Facial recognition	(Taghi Zadeh et al., 2019)

lower loss.

- Random configuration of Gabor filters in the receptive layer performs better for complex datasets.
- Gabor filters in receptive layers lead to higher CNN performance in less time.
- Restricting CNN's ability to alter Gabor filters during training decreases performance.

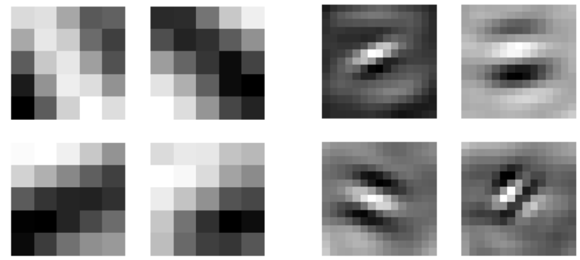
2. Background

Gabor filters, linear filters used for texture analysis and feature extraction, are crucial in image processing applications. They extract patterns at specific frequencies and orientations, with many Gabor filters needed for meaningful features (Jain et al., 1997). These features have been used in various applications, such as retinal, facial, and road feature extraction (El-Sayed et al., 2016; Hemalatha & Sumathi, 2016; Pumlumchiak & Vittayakorn, 2017; Li et al., 2016), and have been integrated into models like Pulse Coupled Neural Networks (PCNNs) and CNNs to enhance performance (Chacon M et al., 2007).

Convolution Neural Networks (CNNs), statistical learning models based on convolution operations, have become popular for image recognition due to their impressive results in various applications (Le Cun et al., 1990; Lawrence et al., 1997; Krizhevsky et al., 2012; Kawano & Yanai, 2014; Schwarz et al., 2015; Fang et al., 2018). Studies have shown that Gabor filters can enhance CNN performance (Yao et al., 2016; Hosseini et al., 2018; Taghi Zadeh et al., 2019).

Previous work has initialized CNN layers with Gabor filters, leading to improved performance and faster convergence (Alekseev & Bobe, 2019; Molaei et al., 2017). Some studies have even initialized multiple CNN layers with different Gabor filters, enhancing model robustness and performance (Luan et al., 2018; Liu et al., 2018).

However, these approaches have limitations. Restricting CNNs to Gabor filters may limit their potential to modify the filter structure for optimal performance. Also, the relationship between Gabor filters and CNN convergence is not well-understood. Lastly, while Gabor filters have been successful in specific tasks, their advantage in general object recognition tasks is unclear. This paper explores the impact of initializing the receptive layer of CNNs with Gabor filters on general object recognition performance.



(a) 5×5 (b) 15×15

Figure 2: Gabor filters with different sizes and resolutions.

3. Methodology

This section outlines our experimental methodology. We first describe the construction of a Gabor filter bank, followed by the datasets used for experimentation. We then detail the CNN architecture, loss function, and training methodologies. Finally, we discuss the success metrics and structure of each experiment.

3.1. Initialization Strategy

A Gabor filter bank is necessary to extract features from an image, as a single Gabor filter can only extract specific texture features. We used the approach proposed in (Meshgini et al., 2012) to design a bank of Gabor filters. We focused on the impact of Gabor filters at the first receptive convolutional layer only.

Our experimental models are categorized into three groups:

1. Random weight initialization, which employs the traditional CNN kernel initialization method using Glorot uniform initialization.
2. Random initialization with a Gabor filter on each channel, where each kernel filter of the receptive layer of CNN is initialized with a random Gabor filter from the filter bank and the CNN is allowed to modify these Gabor filters during training.
3. Repeated Gabor filter on all channels, where a single Gabor filter is assigned to all filters in a particular kernel set with the CNN permitted to alter these Gabor filters during training.

As an example, Figure 2 depicts Gabor filters of different sizes with random orientations.

3.2. Datasets

We considered diverse multi-class datasets for our experiments. The datasets were pre-processed, rescaled, and one-hot encoded before being passed to the respective CNN architecture for training and validation.

3.3. Baseline Architecture

Different CNN architectures were employed depending on the nature of the dataset. The CNN model consisted of convolutional layers, followed by batch normalization, activation, max-pooling, and dropout. The network was expanded with a densely connected neural network followed by batch normalization, activation, dropout, and then the final densely connected neural network.

3.4. Loss Functions

The models were trained to minimize the loss on training data, with emphasis on results in terms of the validation loss. The categorical cross-entropy function was chosen to calculate the validation loss. Adam optimization was chosen as the optimizer for the model based on the validation loss (Kingma & Ba, 2014).

3.5. Success Metrics

The experiments were evaluated based on the metrics from the validation. The main success metrics, using cross-validation, are accuracy, AUC, and loss.

3.6. Experiments

Various CNN models were used to address the nature, distribution, and size of datasets. In order to have a holistic view of how Gabor filters affected CNN, 30 different experiments were performed on the same dataset - 10 experiments for each type of initialization method described in Section 3.1. Gabor filter size was fixed at 15×15 because initial experiments showed this size to be better. The following section will showcase and analyze the outcome of all experiments.

4. Results

A comprehensive set of 10 distinct experiments were conducted across various datasets, each with their unique CNN architecture and receptive convolutional layer kernel configuration. The configurations included random initialization, Gabor filter randomly assigned to each channel, and repeated Gabor filter across all channels. The consistency of the training and validation datasets was maintained for each kernel configuration across all experiments.

The inclusion of Gabor filters potentially enhanced the learning capacity of the Gabor-configured models. The training epochs were not limited, except for the implementation of early stopping if the model showed no significant improvement over a certain number of epochs. The performance of the traditional CNN (randomly initialized) was benchmarked against the Gabor filter configured models in terms of maximum accuracy, AUC at maximum accuracy, and minimum loss, as demonstrated in Tables 2, 3, and 4.

Table 2: Improvement in maximum accuracy of Gabor-configured CNN with respect to traditional CNN

Dataset	Bs Mx Acc		Rand Gabor		Rep Gabor	
	Mean	Std	Mean	Std	Mean	Std
Cats v dogs	0.884	0.00	**+0.023**	**0.01**	**+0.026**	**0.01**
CIFAR-10	0.802	0.00	**+0.021**	**0.00**	**+0.021**	**0.01**
CIFAR-100	0.713	0.00	+0.007	0.01	+0.007	0.01
Caltech 256	0.508	0.01	+0.015	0.01	+0.019	0.01
Stanf. cars	0.233	0.07	**+0.129**	**0.07**	**+0.163**	**0.07**
Tny Imgnet	0.518	0.00	+0.013	0.00	+0.000	0.01
Average	0.610	0.02	**+0.035**	**0.02**	**+0.039**	**0.02**

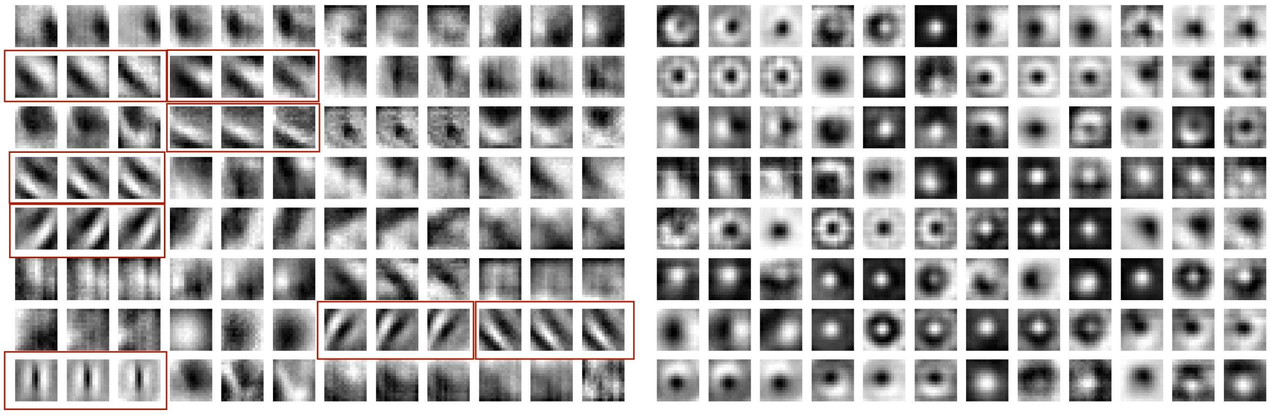
Table 3: Improvement in AUC at maximum accuracy of Gabor-configured CNN with respect to traditional CNN

Dataset	Base AUC		Rand Gabor		Rep Gabor	
	Mean	Std	Mean	Std	Mean	Std
Cats v dogs	0.952	0.00	**+0.014**	**0.00**	**+0.017**	**0.00**
CIFAR-10	0.972	0.00	+0.003	0.00	+0.003	0.00
CIFAR-100	0.962	0.00	+0.001	0.00	+0.002	0.00
Caltech 256	0.889	0.00	+0.008	0.01	+0.004	0.01
Stanf. cars	0.808	0.03	**+0.051**	**0.02**	**+0.063**	**0.03**
Tny Imgnet	0.937	0.00	+0.002	0.00	-0.001	0.01
Average	0.920	0.01	**+0.013**	**0.01**	**+0.014**	**0.01**

The data in Table 2 indicates that the Gabor-configured CNN generally outperformed the traditional CNN in terms of accuracy, especially noticeable in the Cats vs dogs, CIFAR-10, and Stanford cars datasets. The low standard deviation in the Cats vs dogs and CIFAR-10 datasets suggests that Gabor-configured models tend to deliver superior and more consistent performance in terms of accuracy when dealing with less complex datasets. Furthermore, the repeated Gabor configuration generally outperformed the random configuration, although this trend was not observed with increasing dataset complexity.

The kernel filters in the receptive layer of a fully trained traditional CNN, when applied to simpler datasets like Cats vs Dogs, appeared to mimic Gabor filters as shown in Figure 3 (a). This was not observed with more complex datasets, as shown in Figure 3 (b). This suggests that Gabor-initialized models may yield a higher performance gain on simpler datasets.

Tables 3 and 4 present the analysis of AUC at maximum accuracy and minimum loss, respectively. They both indicate that, on average, Gabor-configured models tend to have a higher AUC and lower minimum loss compared to



(a) Kernel filters in the receptive layer of fully trained traditional CNN on Cats vs Dogs dataset.

(b) Kernel filters in the receptive layer of fully trained traditional CNN on CIFAR-100 dataset.

Figure 3: Kernel filters in the receptive layer of fully trained traditional CNN, where three consecutive filters belong to same kernel set

Table 4: Improvement in terms of the minimum loss of Gabor-configured CNN with respect to traditional CNN

Dataset	Bs Min Loss		Rand Gabor		Rep Gabor	
	Mean	Std	Mean	Std	Mean	Std
Cats v dogs	0.296	0.01	-0.044	0.02	-0.056	0.01
CIFAR-10	0.656	0.01	-0.054	0.02	-0.057	0.02
CIFAR-100	1.182	0.02	-0.023	0.02	-0.030	0.02
Caltech 256	2.643	0.07	**-0.104**	**0.08**	**-0.103**	**0.07**
Stanf. cars	4.186	0.36	**-0.781**	**0.29**	**-1.040**	**0.36**
Tny Imgnet	2.739	0.01	-0.053	0.02	-0.004	0.03

traditional CNN. Moreover, the repeated Gabor filter configuration generally outperforms the random Gabor filter configuration, especially when the dataset is less complex.

An analysis of the number of epochs required to train the models revealed that Gabor-configured CNNs tend to learn at a faster rate. While there were instances where Gabor-configured CNNs required more epochs, this was attributed to the fact that Gabor-configured CNNs strive to improve beyond the performance of traditional CNNs. In our experiments, Gabor-configured CNNs achieved the best performance metrics of traditional CNNs in fewer epochs.

The size of the kernel filter and image also played a significant role in performance. It was observed that smaller image sizes did not work as well with Gabor-configured CNNs or traditional CNNs, as smaller details can be missed on smaller image sizes. While there is no direct correlation between performance and image size, larger images provide better detail for CNNs to learn from. Similarly, larger kernels performed better on these larger images because the structure of the Gabor filter is clearer, leading to better feature extraction. This does not suggest a direct correlation between performance and kernel size, but it does indicate that larger kernels tend to perform better.

5. Regarding Vision Transformers

Our study concentrates on the influence of Gabor filters on CNNs for image classification tasks. In contrast, ViTs employ transformer models for image classification, treating images as patch sequences and using self-attention mechanisms (Dosovitskiy et al., 2020). While ViTs show promise, they differ fundamentally from CNNs. Our work aims to improve CNNs by integrating Gabor filters into the receptive layer. Comparing our methodology with ViTs is outside this study's scope, but we recognize ViTs' potential. Future research could investigate using Gabor filters or similar techniques to boost ViTs' performance, merging the strengths of transformers and Gabor filters.

6. Conclusion

Gabor filters have emerged as a potent feature extractor in image processing (Luan et al., 2018; Alekseev & Bobe, 2019; Molaei et al., 2017). Given that the receptive filters of CNNs often resemble Gabor filters, it is plausible that Gabor filters could serve as an effective receptive filter for CNNs. An exhaustive analysis was conducted on a wide array of general object datasets using unrestricted Gabor filter initialization in the receptive layer. The results, as shown in Tables 2, 3, and 4, demonstrate that the integration of Gabor filters in the receptive layer significantly enhances the performance of CNNs, leading to higher accuracy, higher Area Under the Curve (AUC), and lower loss on various datasets. This indicates that Gabor filters contribute to substantial improvements in general object classification. Furthermore, under a restricted training epoch, it was found that

CNNs trained with Gabor filters in the receptive layer could achieve superior performance in a shorter time compared to traditional randomization techniques.

The generation of Gabor filters with varying hyperparameters corresponds to unique image features, and their configuration in the receptive layer influences the performance. For less complex datasets, repeated Gabor filter configurations yield better results, while for more complex datasets, random Gabor filter configurations perform better.

The dimensions of Gabor filters also significantly impact the performance of CNNs, particularly in the case of smaller images. Determining the optimal size of Gabor filters is critical for effective feature extraction.

This research's potential implications are vast, and the findings thus far are promising, indicating that the use of Gabor filters in CNNs can significantly enhance performance and efficiency in image processing tasks. The results presented in this paper provide a strong foundation for future exploration and development in this area.

References

Alekseev, A. and Bobe, A. Gabornet: Gabor filters with learnable parameters in deep convolutional neural network. In *2019 International Conference on Engineering and Telecommunication (EnT)*, pp. 1–4, 2019.

Alekseev, A. and Bobe, A. Gabornet: Gabor filters with learnable parameters in deep convolutional neural network. In *2019 International Conference on Engineering and Telecommunication (EnT)*, pp. 1–4. IEEE, 2019.

Bai, J., Zeng, Y., Zhao, Y., and Zhao, F. Training a v1 like layer using gabor filters in convolutional neural networks. In *2019 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–8. IEEE, 2019.

Behnke, S. and Rojas, R. Neural abstraction pyramid: A hierarchical image understanding architecture. In *1998 ieee international joint conference on neural networks proceedings. ieee world congress on computational intelligence (cat. no. 98ch36227)*, volume 2, pp. 820–825. IEEE, 1998.

Chacon M, M. I., Zimmerman S, A., and Rivas P, P. Image processing applications with a pcnn. In *Advances in Neural Networks: 4th Intl. Symposium on Neural Networks, ISNN 2007*, pp. 884–893. Springer, 2007.

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., et al. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*, 2020.

El-Sayed, M. A., Hassaballah, M., and Abdel-Latif, M. A. Identity verification of individuals based on retinal features using gabor filters and svm. *Journal of Signal and Information Processing*, 7, 2016.

Fang, W., Ding, L., Zhong, B., Love, P. E., and Luo, H. Automated detection of workers and heavy equipment on construction sites: A convolutional neural network approach. *Advanced Engineering Informatics*, 37:139 – 149, 2018.

Gaussier, P. and Cocquerez, J.-P. Neural networks for complex scene recognition: Simulation of a visual system with several cortical areas. In *[Proceedings 1992] IJCNN International Joint Conference on Neural Networks*, volume 3, pp. 233–259. IEEE, 1992.

Han, K., Wang, Y., Chen, H., Chen, X., Guo, J., Liu, Z., Tang, Y., Xiao, A., Xu, C., Xu, Y., et al. A survey on vision transformer. *IEEE transactions on pattern analysis and machine intelligence*, 45(1):87–110, 2022.

He, K., Zhang, X., Ren, S., and Sun, J. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016a.

He, K., Zhang, X., Ren, S., and Sun, J. Identity mappings in deep residual networks. In *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part IV 14*, pp. 630–645. Springer, 2016b.

Hemalatha, G. and Sumathi, C. P. Preprocessing techniques of facial image with median and gabor filters. In *2016 International Conference on Information Communication and Embedded Systems (ICICES)*, pp. 1–6, 2016.

Hosseini, S., Lee, S. H., Kwon, H. J., Koo, H. I., and Cho, N. I. Age and gender classification using wide convolutional neural network and gabor filter. In *2018 International Workshop on Advanced Image Technology (IWAIT)*, pp. 1–3, 2018.

Huang, G., Liu, Z., Van Der Maaten, L., and Weinberger, K. Q. Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4700–4708, 2017.

Ide, H. and Kurita, T. Improvement of learning for cnn with relu activation by sparse regularization. In *2017 international joint conference on neural networks (IJCNN)*, pp. 2684–2691. IEEE, 2017.

Jain, A. K., Ratha, N. K., and Lakshmanan, S. Object detection using gabor filters. *Pattern Recognition*, 30(2): 295 – 309, 1997. ISSN 0031-3203. doi: https://doi.org/10.1016/S0031-3203(96)00068-4.

Kawano, Y. and Yanai, K. Food image recognition with deep convolutional features. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication*, UbiComp '14 Adjunct, pp. 589–593, New York, NY, USA, 2014. Association for Computing Machinery. doi: 10.1145/2638728.2641339.

Kingma, D. P. and Ba, J. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

Krizhevsky, A., Sutskever, I., and Hinton, G. E. Imagenet classification with deep convolutional neural networks. In Pereira, F., Burges, C. J. C., Bottou, L., and Weinberger, K. Q. (eds.), *Advances in Neural Information Processing Systems 25*, pp. 1097–1105. Curran Assoc., Inc., 2012.

Lawrence, S., Giles, C. L., Ah Chung Tsoi, and Back, A. D. Face recognition: a convolutional neural-network approach. *IEEE Transactions on Neural Networks*, 8(1): 98–113, 1997.

Le Cun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., and Jackel, L. D. Handwritten digit recognition with a back-propagation network. In *Advances in Neural Information Processing Systems*, pp. 396–404. Morgan Kaufmann, 1990.

Li, J., Xu, H., Deng, J., and Sun, X. Hyperbolic linear units for deep convolutional neural networks. In *2016 International Joint Conference on Neural Networks (IJCNN)*, pp. 353–359. IEEE, 2016.

Li, Z., Ma, H., and Liu, Z. Road lane detection with gabor filters. In *2016 Intl. Conference on Information System and Artificial Intelligence (ISAI)*, pp. 436–440, 2016.

Li, Z., Liu, F., Yang, W., Peng, S., and Zhou, J. A survey of convolutional neural networks: analysis, applications, and prospects. *IEEE transactions on neural networks and learning systems*, 2021.

Liu, C., Ding, W., Wang, X., and Zhang, B. Hybrid gabor convolutional networks. *Pattern Recognition Letters*, 116: 164 – 169, 2018.

Luan, S., Chen, C., Zhang, B., Han, J., and Liu, J. Gabor convolutional networks. *IEEE Transactions on Image Processing*, 27(9):4357–4366, 2018.

Meshgini, S., Aghagolzadeh, A., and Seyedarabi, H. Face recognition using gabor filter bank, kernel principle component analysis and support vector machine. *International Journal of Computer Theory and Engineering*, pp. 767–771, 2012.

Molaei, S., Shiri, M., Horan, K., Kahrobaei, D., Nallamothu, B., and Najarian, K. Deep convolutional neural networks for left ventricle segmentation. In *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 668–671, 2017.

Pumlumchiak, T. and Vittayakorn, S. Facial expression recognition using local gabor filters and pca plus lda. In *2017 9th Intl. Conference on Information Technology and Electrical Engineering (ICITEE)*, pp. 1–6, 2017.

Rai, M. and Rivas, P. A review of convolutional neural networks and gabor filters in object recognition. In *2020 International Conference on Computational Science and Computational Intelligence (CSCI)*, pp. 1560–1567, 2020. doi: 10.1109/CSCI51800.2020.00289.

Reis, F. A., Almeida, R., Kijak, E., Malinowski, S., Guimarães, S. J. F., and do Patrocínio, Z. K. Combining convolutional side-outputs for road image segmentation. In *2019 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–8. IEEE, 2019.

Schwarz, M., Schulz, H., and Behnke, S. Rgb-d object recognition and pose estimation based on pre-trained convolutional neural network features. In *2015 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 1329–1335, 2015.

Sigillito, V., Sadowsky, J., Bankman, I., and Willson, P. Application of feedforward neural networks to object recognition for image analysis. In *IJCNN-91-Seattle International Joint Conference on Neural Networks*, volume 2, pp. 933–vol. IEEE, 1991.

Taghi Zadeh, M. M., Imani, M., and Majidi, B. Fast facial emotion recognition using convolutional neural networks and gabor filters. In *2019 5th Conf. on Knowledge Based Engineering and Innovation (KBEI)*, pp. 577–581, 2019.

Wu, S., Zhang, X., Wang, X., Li, C., and Jiao, L. Scene attention mechanism for remote sensing image caption generation. In *2020 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–7. IEEE, 2020.

Xi, R., Hou, M., Fu, M., Qu, H., and Liu, D. Deep dilated convolution on multimodality time series for human activity recognition. In *2018 international joint conference on neural networks (IJCNN)*, pp. 1–8. IEEE, 2018.

Yao, H., Chuyi, L., Dan, H., and Weiyu, Y. Gabor feature based convolutional neural network for object recognition in natural scene. In *2016 3rd Intl. Conf. on Information Science and Control Engineering*, pp. 386–390, 2016.