International Journal of Applied Engineering Research, ISSN 0973-4562 Volume 13, Number 13 (2018) pp. 11012–11017
 https://dx.doi.org/10.37622/IJAER/13.13.2018.11012-11017

Unadorned Gabor based Convolutional Neural Network Overrides Transfer Learning Concept

Jyothi R. L.
University of Kerala, College of Engineering Chengannur, Kerala, India.

Abdul Rahiman M.
Kerala Technological University, Kerala, India.

Anilkumar A.
College of Engineering Chengannur, Kerala, India.

Abstract: The efficiency of Convolutional Neural Networks (CNN) is highly influenced by the size of dataset. To train CNN systems from the scratch, dataset of very large size is essential. Transfer learning concept was introduced to overcome this deficiency of CNN. Even though in recent years transfer learning networks has attained high popularity, the adaption of transfer learning networks to entirely different dataset would produce very low recognition results. In this work a Gabor based CNN network is introduced which works highly efficient compared to transfer learning networks. Another deficiency of CNN is that it is not robust to rotation. Even though the notion of Gabor filter being induced in CNN has been suggested earlier, this work introduces an entirely different and very simple Gabor based CNN which produces high recognition efficiency in dataset of very small size and works invariant to rotation. This work would shed new light in deep Learning research where researchers are forced to focus on and build highly complex CNN network structures. Small subset of MNIST dataset with rotated images is used to learn the proposed architecture.

Keywords: Unadorned Gabor based Convolutional Neural Network (UGCNN); Convolutional Neural Network(CNN), transfer learning small dataset; rotation invariant.

INTRODUCTION

Deep Learning has dramatically advanced the state of art in vision, speech and other related areas. The basis for the concept of current convolutional Neural Network (CNN) has been introduced in 1989 [1]. Record beating performance reported by CNN in Imagenet dataset 2012 has shed a new light for researchers in the field of machine learning [2]. Many research papers has already been published that demonstrates outstanding performance of Convolutional Neural Network in various areas. The performance of CNN network is highly influenced by (i) Availability of large dataset (ii) Powerful GPU implementations [3].

CNN would work efficiently only when the dataset for learning is very large. For dealing with small dataset the concept of transfer features were introduced. Estimation of apt kernel values for convolution filter design will enhance the efficiency of CNN network. Convolution kernel values are initialized with random Gaussian distribution which is relearned in further iterations through back propagation and stochastic gradient descent method [4]. But when the kernels of convolution matrix are initialized with random distribution it will take a longer time for network to reach a conclusion stage. Moreover the CNN initialized with random values for convolution kernels will produce a very low recognition result when the dataset is very small. CNN needs a large dataset to learn to be an efficient recognition system. In transfer network the convolution layers in the CNN networks are replaced with pre-learned convolution layer values and fine-tuned based on target dataset. Transfer network with no fine tuning would produce very low recognition efficiency [5]. Performance of the transfer network reduces when the difference between the base dataset (dataset based on which training is carried out) and target dataset increases[5]. Therefore transfer network is not good choice in building an efficient CNN based recognition system. Moreover CNN produces low recognition result when tested with rotated version of the same images used for learning [6] [7].

In this work a new modified version of CNN is developed by incorporating Gabor features in designing convolution kernels of the convolution layer for dealing with small dataset. Here the layers of the first convolution layer of a simple CNN is initialized with Gabor filters of 16 different orientations. The new version of CNN known as Unadorned Gabor based CNN (UGCNN) learn small subsets of MNIST datasets with efficiency greater than pre-trained CNN (CNN learned based on transfer features). More UGCNN produces high recognition using rotated version of the images. Section II explains the related works. Structure and working of UGCNN is explained in Section III. Section IV demonstrates experimental analysis that illustrates the working efficiency of UGCNN architecture.

RELATED WORK

The learning process in CNN network deals with finding optimum weights for the convolution kernels. Time taken for the network to learn depends upon time taken by it to learn and update the apt weight values for its convolution kernels that produces high recognition efficiency. When weights of convolution filters are initialized with random Gaussian distribution it would take a long time to reach a concluded weight state and with very small dataset the working of the recognition system highly depreciates. To reduce the deficiency of existing CNN architecture Gabor based convolution layers are introduced. The idea of incorporating Gabor kernels is not a novel one. In [8] character images from MNIST dataset are first subjected to extraction of Gabor features. The features extracted are then fed to basic CNN. With MNIST dataset of 60,000 image instances, the system produced a recognition efficiency of 99.32. In [9] initially Gabor features of 4 orientations are extracted thereby creating 4 images of different orientations. All the four images are then separately passes to CNN for learning. Experiments with AT& T database show an efficiency of 89.50 percentages. In [10] the derived Gabor features are analysed with CNN and a recognition accuracy of 93 percentages is claimed.The proposed work is different from the above three cited existing works .In the proposed work the kernels of first convolution layer in basic CNN architecture is replaced with Gabor features of different orientations. Moreover the proposed architecture is learned and tested with small sized data subset. In [11] the convolution layer in CNN (with two set of convolution layers) is replaced with Gabor kernels to increase the recognition accuracy, computational time and energy requirements. Three variations of analysing the convolution kernels with Gabor features were analysed. Here all the variations analysed Gabor features with 12 orientations. First variation replaced only the first convolution layer with fixed Gabor kernel values but the kernel values are never relearned. This variation obtained a recognition accuracy of 99.38 percentages with 60,000 instances of MNIST dataset. The second variation replaced the convolution kernels in first and second layer with fixed Gaussian kernels without relearning. The accuracy of the system reduced to 94.15 percentages. The third variation replaced some of the kernels in each convolution layer with Gabor filter values and others with random initialization of values which is further relearned. The efficiency of this variation is 99 percentages. The proposed UGCNN is different from this work as it initializes only the first layer convolution kernels with Gabor features and the kernels are subjected to relearning. In [6] receptive fields of all convolution layers of CNN were designed with Gabor filters of 4 orientation and 2 scales. When trained with 5000 images instances of 98.15 percentages is reported. And when trained with 60,000 images of MNIST data the recognition efficiency is 99.43 percentages and when learned with MNIST rotated images the recognition efficiency is 94.92 percentages. But in the proposed work convolution kernels of only the first layer is initialized with Gabor features of 16 different orientations using a single Gabor scale and produces high recognition efficiency with small sized MNIST data subset with rotated version of MNIST images included. In [12] convolution kernels in convolution layers are initialized with combination of random initialized kernels and Gabor kernel coefficients. Here Gabor filters are designed by taking into account frequency and time features and no orientation is taken into account. More over the network is highly complex with 59 Gabor based filters and 61 random value initialized filters in each layer. But in the proposed work the network is highly simple using only 16 Gabor filter values for convolution. In [13] pre-trained convolution kernels (transferred) are subjected to convolution by Gabor filters of 4 orientations and kernel size 7x7. The scale value of the convolution filter increases with increasing layer. Here an efficiency of 99.58 percentage is obtained when the network is learned with MNIST dataset of 60000 images and when MNIST rotated images are used the network produced an recognition efficiency of 98.90.In proposed UGCNN the network is learned from scratch instead of using pre-trained network(transfer features).

Unique features of UGCNN:

- Network structure of UGCNN resembles CNN designed with three sets of convolution relu pairs.
- Only the initial convolution layer is initialized with Gabor features, as more general features are learned in this layer [3].
- All the convolution layers of UGCNN are subjected to retraining thereby increasing the efficiency of the system.
- Produces excellent result with small dataset.
- UGCNN works invariant to rotation, translation and scaling.

UNADORNED GABOR BASED NEURAL NETWORK

Unadorned Gabor Based Neural Network is a simple Gabor based neural network with Gabor filters used as kernels in the first convolution layer. Here weight vectors of kernels in the first convolution layer are initialized with Gabor feature values. Convolutional Neural network is not robust towards rotation. Using Gabor filters in the initial convolution layers makes the CNN network invariant to rotation. Some of the existing works has already used Gabor filters for creating convolution layer filters but most of the work has taken into account the scale feature of the Gabor features and not orientation features for designing weight values. Taking into account the scale values is of less importance in building an efficient recognition system as most of the input values can be projected to a particular scale where the feature space can be easily classified. Some of the existing work used fixed Gabor filters with different orientations values for initialling some of the kernels in the convolution layers and other kernels are initialized with random weights. The Gabor filters are frozen and never subjected to learning by back propagation but randomly initialized layers in those works are subjected to relearning. Some of the networks have used combination of orientation values and scale values for initialling weight values of kernels of filters in all the convolutions layers. Such a way of initialization of convolution layers with Gabor have no theoretical base. But in Unadorned Gabor Based Neural Network only the first convolution layer is initialized with Gabor weights and other layers are initialized with random Gaussian distribution. The basic general features like edges, corners, lines are learned by the first convolution layer. Therefore learning of directional features should be incorporated into the first convolution layer instead of other layers which learns more specific features [3]. Moreover as Gabor features are spectral features the network can be easily generalized with small dataset which overrides the concept of transfer learning. Unadorned Gabor Based Neural Network uses a simple CNN structure without making the structure complicated to make the network more complex for generalization purpose.

Gabor Filters

The Concept of Gabor filter was introduced by Dennis Gabor [14] to serve as the basis for Fourier transforms in information theory. It capture localized regions of spectrum and temporal information over a broader time interval and acts like band pass filters which are convolved with the input image to extract the image features. Frequency and orientation representations

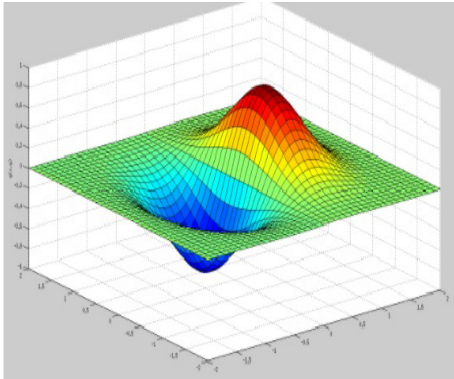

FIGURE 1. Response of a single Gabor filter with σ=1 and θ =0

of Gabor filters are similar to those of the human visual system. A particular advantage of Gabor filters is their degree of invariance to scale, rotation, and translation. The impulse response is a product of a sinusoidal function and a Gaussian function. The Gabor filter coefficients are calculated as

$$G(x, y) = exp\left(\left(-\frac{(M^2+\gamma^2N^2)}{2\sigma^2}\right) cos\left(\frac{2\pi}{\lambda}M\right)\right) \quad (1)$$

Here $M = xcos\theta - ysin\theta$ and $N = xcos\theta + ysin\theta$.

The range of x and y determines the size of Gabor filter. The size of the filter has direct effect in building an efficient recognition system. Even though deep learning is trying to build a recognition system that can be applied wide range of applications, selection of a generalized size filter may reduce the efficiency of recognition system. To large and too small size may have inverse effect in recognition systems. Large filters may sharpen the noise in recognition system and small sized filter may useful information. Fig. 1 shows the response of Gabor filter with σ=1 and θ =0.

γ is the aspect ratio, and λ determines wavelength. These parameters together determine the frequency change (rate of change of adjacent values) in kernel data distribution in both directions of a 2 dimensional kernel. A smaller wavelength means a denser sinus wave. A larger wavelength means larger waves.

σ determines the scale (effective width). Data values in the kernel is randomly distributed in such a way that variance of data should be of particular scale σ. Larger values of σ captures a broader range of frequencies resulting in tighter band pass and a poorer spatial localization. By selecting different range of σ values the Gabor filter tends to work in different scales.

θ determines the orientation of Gabor kernels. Orientation aspect of Gabor kernels makes object recognition using Gabor filter invariant to rotation. As Gabor filters are combination of sinusoidal wave and Gaussian function σ and θ determines

International Journal of Applied Engineering Research, ISSN 0973-4562 Volume 13, Number 13 (2018) pp. 11012–11017
©Research India Publications, https://dx.doi.org/10.37622/IJAER/13.13.2018.11012-11017

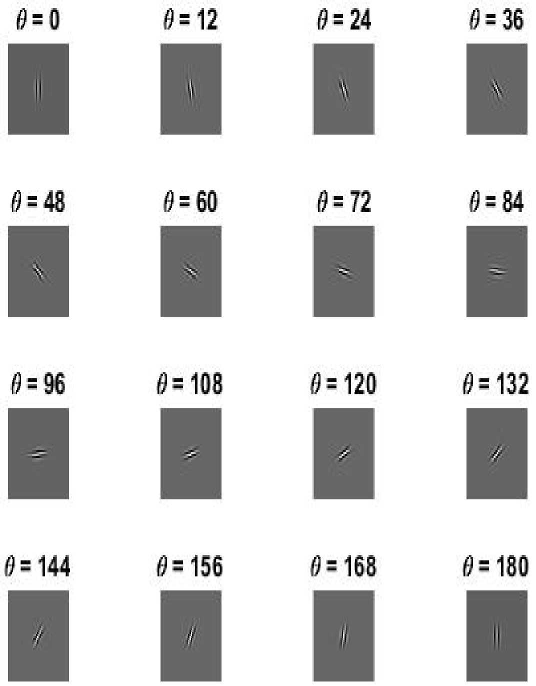

FIGURE 2. Gabor convolution filters at 16 different orientations

Gaussian component and γ and λ determines sinusoidal component. Fig 2 shows Gabor convolution filters at 16 different orientations.

Architecture

The CNN network is divided into 4 groups. First two groups consist of Convolution Layer, RELU Layer and Max pooling Layer. Third group consists of only Convolution Layer and RELU Layer. The fourth group consists of fully connected layer. First group consists of Convolution layer initialized with Gabor Kernels, second and third group consists of Convolution Layer initialized with random Gaussian distribution.

The first convolution layer is designed based on 16 Gabor filters with uniformly distributed θ(orientation) values (0 12 24 36 48 60 72 84 96 108 120 132 144 156 168 180). A Gabor kernel of size 3×3 is used as convolution filter. The size of the Gabor filter has been selected based on experimental analysis which is depicted in Section IV. Scale (σ) is chosen as 4.1, γ as 5.2 and λ as 0.3.Selection of all the values are based on experimental analysis which is shown in Section IV. A single column and row of padding with zero values is done in the top, bottom, left and right of the input image before performing convolution. The convolution layer of the first group convert a single input image into 16 output images which is then passed to the ReLU Layer.

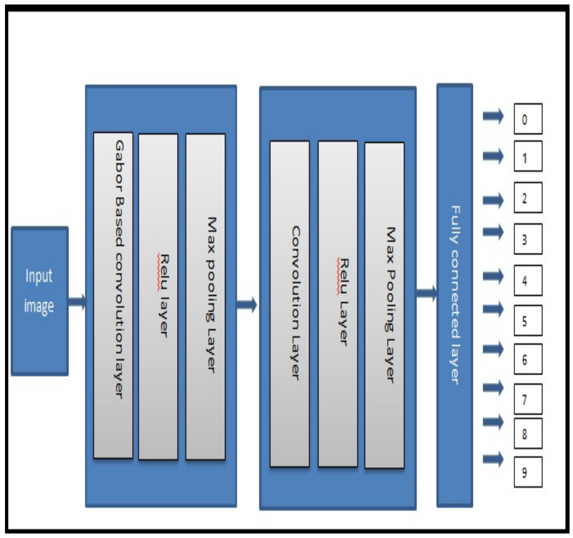

FIGURE 3. Basic structure of Unadorned Gabor Based Convolution Neural Network

Convolution Layer in second and third group works on kernels initialized with random Gaussian distribution values with standard deviation 0.0001.Thirty two filters are used for convolution in the first group and 64 filters are used for convolution in the in the second group. Biases are initialized with Gaussian distribution of mean value 1 and standard deviation 0.00001. The learning rate factor of the bias is set to 2.

Rectified Linear Unit (ReLU) layer performs a threshold operation to each element, where any input value less than zero is set to zero. ReLU Layer is used in first three groups.

Maxpooling layer perform maxpooling with kernel size of 2×2 that works with a stride of '2' is used in first and second group.

Fully connected layer with output size 10 is used in this network. The weight values in the fully connected layer is initialized with Gaussian distribution with mean zero and standard deviation 0.0001. The bias values are initialized with Gaussian distribution with a mean of 1 and deviation of 0.0001. The learning rate factor of the bias is set to 2. Softmax is used as activation function in the output layer of the fully connected layer. Fig. 3 shows Basic structure of Unadorned Gabor Based Convolution Neural Network.

EXPERIMENTAL ANALYSIS

Selection of Gabor filter parameters

The size of the Gabor filter is designed based on experimental analysis. Efficiency of the system is tested with Gabor filters of different size. The system is tested with Gabor filters of size 3×3, 5×5 and 7×7. Highest efficiency has been produced

International Journal of Applied Engineering Research, ISSN 0973-4562 Volume 13, Number 13 (2018) pp. 11012–11017
©Research India Publications, https://dx.doi.org/10.37622/IJAER/13.13.2018.11012-11017

when the kernel size is 3×3. When size of the filter is 3×3 the efficiency of the system is 98.68 and when the kernel size is 5×5 the efficiency of the system is 97.69 and when the size of kernel is 7×7 the efficiency of the system is 97.46.Therefore kernel size of 3×3 is selected. The selection of σ, γ and λ is also carried out based on experimental analysis. It can be found that when σ is 4.1, γ is 5.2 and λ is 0.3 the system produces the highest recognition efficiency. Combinations of different values of σ and γ ranging between 1 and 10 is tested for the system and it is found that σ values between 4 and 5 and γ values between 5 to 6 produces highest recognition rate. Again testing of the recognition system is repeated with different combinations of σ between 4.1 to 4.9 and γ between 5.1 to 5.9 and it is found highest efficiency of UGCNN is reported when σ is 4.1 and γ is 5.2. Fixing σ as 4.1 and γ as 5.2 different λ values ranging from 0 to 0.5 is tested for UGCNN and it is found that fixing λ values to 0.3 will increase the efficiency of the system. For selection of parameters of Gabor filters testing is carried out in a subset of MNIST dataset.

Training UGCNN with MNIST dataset

A small subset of MNIST dataset consisting of about 1000 image instances of digits from 0 to 9 with rotated images of digits are used for learning UGCNN. The experiments were carried out with Nvidia GTX 1070. CNN when learned with MNIST dataset of 60,000 digit images an efficiency of 99.68 percentages has been reported [15]. Small 5 dataset subsets consisting of 1000 images have been created from MNIST dataset. A simple CNN network consisting of 4 groups are created with first three groups consists of convolution layer followed by ReLU layer. Convolution ReLU layers in first and second group is followed by maxpooling layer. Convolution kernels of the first convolution layer is made of 16 kernels initialized with random Gaussian distribution. Convolution layer in the second group in the second layer is made of 32 filters and in third group is made of 64 filters. All these convolution kernels are initialized with Gaussian distribution. Last Group is made of a single fully connected layer. This simple CNN is trained with 1000 images of each dataset out of which only 750 used for training and others are used for validation and testing. The top-1 recognition efficiency of simple CNN working with different MNIST dataset subsets is shown in the TABLE 1.

Same subsets are used to relearn a transfer feature CNN network. The transfer CNN network is created by copying first three layers of the pre-trained CNN to target CNN. The newly created transferred CNN network consists of 3 transferred layers and one fully connected layers. The layers in the network are fine-tuned based on each set of datasets used for learning. TABLE II shows the recognition efficiency of CNN designed based on transfer features.

Table 1. Top-1 Efficiency of Simple CNN dataset for different data subsets created from MNIST dataset

DATASET	Top-1 Efficiency
Datasubset1	35.6
Datasubset2	32.8
Datasubset 3	40.5
Datasubset 4	39.6
Datasubset 5	37.6

Table 2. Top-1 Efficiency of Transfer CNN for different data subsets created from MNIST dataset

DATASET	Top-1 Efficiency
Datasubset 1	93.6
Datasubset 2	92.8
Datasubset 3	90.5
Datasubset 4	89.6
Datasubset 5	89.8

Table 3. Top-1 Efficiency of UGCNN for different data subsets created from MNIST dataset

DATASET	Top-1 Efficiency
Datasubset 1	98.65
Datasubset 2	98.68
Datasubet 3	98.81
Datasubset 4	98.83
Datasubset 5	98.72

Same data subsets were learned with newly proposed UGCNN architecture. The recognition efficiency is shown in table TABLE III.

From the tabulated result it can be seen that newly proposed UGCNN architecture overrides transfer learning concept in dealing with small datasets. Fig 4 shows the comparative analysis of CNN, transfer learning based CNN and UGCNN in recognition of MNIST dataset of 1000 image instances.Proposed UGCNN architecture produced a highest efficiency of 98.83 while working with rotated images which is very much higher than the efficiency reported in [6] for rotated MNIST images.

CONCLUSION

In this work a simple version of Convolutional Neural Network is developed where the convolution kernels of the first Convolution layer is replaced by Gabor features of various orientations. The proposed work produced high recognition

International Journal of Applied Engineering Research, ISSN 0973-4562 Volume 13, Number 13 (2018) pp. 11012–11017
©Research India Publications, https://dx.doi.org/10.37622/IJAER/13.13.2018.11012-11017

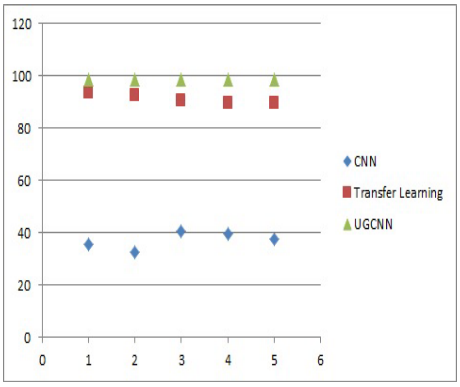

FIGURE 4. Comparative Analysis of CNN, Transfer Learning Based CNN and UGCNN

efficiency of 98.83 when learned with MNIST dataset containing 1000 instances of each digit category(with rotated images included). This work produced high recognition efficiency compared to basic CNN network and Transfer learning based CNN when tested with small sized dataset. This work overcomes the deficiency of CNN in working with small sized dataset and rotated images. Moreover this work can be extended to enhance the efficiency of very deep neural networks like VGGNet, AlexNet, ReSNet etc.

REFERENCES

[1] LeCun, Yann, Bernhard Boser, John S. Denker, Donnie Henderson, Richard E. Howard, Wayne Hubbard, and Lawrence D. Jackel, 1989, "Backpropagation applied to handwritten zip code recognition.", *Neural computation*, vol.1, no. 4 pp. 541–551.

[2] Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton, 2012, "Imagenet classification with deep convolutional neural networks." In Advances in neural information processing systems, pp. 1097–1105.

[3] Zeiler, Matthew D., and Rob Fergus, 2014, "Visualizing and understanding convolutional networks." In European conference on computer vision, pp. 818–833. Springer, Cham.

[4] Kingma, Diederik, and Jimmy Ba, 2014, "Adam: A method for stochastic optimization." arXiv preprint arXiv:1412.6980

[5] Yosinski, Jason, Jeff Clune, Yoshua Bengio, and Hod Lipson, 2014, "How transferable are features in deep neural networks?." In Advances in neural information processing systems, pp. 3320–3328.

[6] Verkes, Govert, 2017, "Receptive Fields Neural Networks using the Gabor Kernel Family."

[7] Nguyen, Anh, Jason Yosinski, and Jeff Clune, 2015, "Deep neural networks are easily fooled: High confidence predictions for unrecognizable images." In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 427–436.

[8] Caldern, Andrs, Sergio Roa, and Jorge Victorino, 2003, "Handwritten digit recognition using convolutional neural networks and gabor filters." Proc. Int. Congr. Comput. Intell.

[9] Kinnikar, Ashwini, Moula Husain, and S. M. Meena, 2016, "Face Recognition Using Gabor Filter and Convolutional Neural Network." In Proceedings of the International Conference on Informatics and Analytics, p. 113. ACM.

[10] Xu, Yajun, Fengmei Liang, Gang Zhang, and Huifang Xu, 2016, "Image Intelligent Detection Based on the Gabor Wavelet and the Neural Network." Symmetry 8, no. 11: 130.

[11] Sarwar, Syed Shakib, Priyadarshini Panda, and Kaushik Roy, 2017, "Gabor Filter Assisted Energy Efficient Fast Learning Convolutional Neural Networks" arXiv preprint arXiv:1705.04748.

[12] Chang, Shuo-Yiin, and Nelson Morgan, 2014, "Robust CNN-based speech recognition with Gabor filter kernels." In Fifteenth Annual Conference of the International Speech Communication Association.

[13] Luan, Shangzhen, Baochang Zhang, Chen Chen, Xianbin Cao, Qixiang Ye, Jungong Han, and Jianzhuang Liu, 2017, "Gabor Convolutional Networks." arXiv preprint arXiv:1705.01450.

[14] Jain, Anil K., Nalini K. Ratha, and Sridhar Lakshmanan, 1997, "Object detection using Gabor filters." Pattern recognition 30, no. 2: 295–309.

[15] Simard, Patrice Y., David Steinkraus, and John C. Platt, 2003, "Best practices for convolutional neural networks applied to visual document analysis." In ICDAR, vol. 3, pp. 958–962.