EE6222-Machine Vision Assignment 1

Msc CCA Yongqian Huang
Matric No. G1801127L

1. Effects of direct links from the input layer to the output layer (i.e. with and without) and the bias in the output neuron (i.e. with and without)

Considering the limitations of time and computer hardware, I select 10 datasets from the 121 datasets of the UCI repository to implement the evaluation experiments, i.e., Abalone, Car, Ecoli, Magic, Nursery, Page-blocks, Semeion, Wine, Yeast and Zoo. Sinece the training-testing partition of all the selected datasets are avaliable, the RVFL classifer will be trained on the preset training set and tested on the preset test set. Moreover, all input features are normalized by subtracting the mean value and dividing by the square of the variance.

In oder to obtain parameters with the best performance on the test set, parameter tuing of two parameters is performed. One is the number of hidden neurons *N*, which is tuned over 3 to 203 with a step of 20. The other one is λ in the ridge regression, which is equal to 2^C, where C changs from -5 to 14 with a step of 1.

To discuss the effects of direct links from the input layer to the output layer and the bias in the output neuron, I choose radbas as the Activation Function and use the ridge regression for the solution to the ouput. The experiment results are shown in the following Table 1.

Table 1 : Means and variances of the RVFL classification accuracies with different conditions of the direct links from the input layer to the output layer and the bias in the output neuron

Accurarcy	Without Bias Without Direct Link	With Bias Without Direct Link	Without Bias With Direct Link	With Bias With Direct Link
Abalone	0.5771 ± 0.0101	0.5789 ± 0.0098	0.6270 ± 0.0011	0.6271 ± 0.0011
Car	0.7660 ± 0.0051	0.7664 ± 0.050	0.7875 ± 0.0039	0.7875 ± 0.0039
Ecoli	0.7018 ± 0.0333	0.7045 ± 0.0329	0.7901 ± 0.0165	0.7897 ± 0.0166
Magic	0.7409 ± 0.0046	0.7432 ± 0.0044	0.7722 ± 0.0025	0.7730 ± 0.0025
Nursery	0.7290 ± 0.0446	0.7376 ± 0.0415	0.8601 ± 0.0060	0.8610 ± 0.0056
Page-blocks	0.9200 ± 0.0004	0.9204 ± 0.0004	0.9264 ± 0.0003	0.9265 ± 0.0003
Semeion	0.5679 ± 0.0609	0.5691 ± 0.0602	0.8242 ± 0.0030	0.8240 ± 0.0030
Wine	0.7873 ± 0.0505	0.7924 ± 0.0468	0.9348 ± 0.0086	0.9337 ± 0.0089
Yeast	0.4892 ± 0.0132	0.4892 ± 0.0124	0.5461 ± 0.0058	0.5458 ± 0.0059
Zoo	0.7504 ± 0.0481	0.7514 ± 0.0472	0.8874 ± 0.0148	0.8846 ± 0.0161

In Table 1, each column presents means and variances of the RVFL classification accuracies with all the possible *N* and λ based on the condition of the direct links from the input layer to the output layer and the bias in the output neuron. From the second and the third column, we can observe that the bias in the output neuron almost has no effect on the the RVFL classification accuracies. However, the direct links from the input layer to the output layer play an more significant role in the RVFL structure. Apparently, the mean accuracies in the 4th and 5th columns are larger than those in the second and third columns, while the variances of the accuracies in the 4th and 5th columns are much smaller than those in the second and third columns. Thus, the direct links from the input layer to the output layer work as a regularization operator for the randomly initialized parameters and push the RVFL to obtain a whole better performance than the RVFL without the direct links.

2. Performances of 2 activation functions: one from "sigmoid, radbas, sine, sign" and one from "hardlim, tribas"

According to the conclusion above, in order to compare the five activation functions, I implement the direct links from the input layer to the output layer as well as the bias in the output neuron in the network structure and use the ridge regression for the solution to the ouput. The results are shown in the following Table 2.

Table 2 : Means and variances of the RVFL classification accuracies with different activation functions

Accurarcy	sigmoid	radbas	sine	sign	hardlim	tribas
Abalone	0.6213 $\pm$ 0.0011	0.6271 $\pm$ 0.0011	0.6326 $\pm$ 0.0010	0.6180 $\pm$ 0.0013	0.6180 $\pm$ 0.0013	0.6296 $\pm$ 0.0010
Car	0.7713 $\pm$ 0.0029	0.7875 $\pm$ 0.0039	0.8063 $\pm$ 0.0058	0.7703 $\pm$ 0.0013	0.7700 $\pm$ 0.0013	0.7827 $\pm$ 0.0026
Ecoli	0.7746 $\pm$ 0.0189	0.7897 $\pm$ 0.0166	0.7914 $\pm$ 0.0147	0.7833 $\pm$ 0.0048	0.7838 $\pm$ 0.0050	0.7935 $\pm$ 0.0150
Magic	0.7630 $\pm$ 0.0024	0.7730 $\pm$ 0.0025	0.7850 $\pm$ 0.0028	0.7572 $\pm$ 0.0015	0.7572 $\pm$ 0.0015	0.7708 $\pm$ 0.0017
Nursery	0.8587 $\pm$ 0.0054	0.8610 $\pm$ 0.0056	0.8782 $\pm$ 0.0042	0.7976 $\pm$ 0.0186	0.7977 $\pm$ 0.0185	0.8548 $\pm$ 0.0051
Page-blocks	0.9211 $\pm$ 0.0002	0.9265 $\pm$ 0.0003	0.9296 $\pm$ 0.0004	0.9276 $\pm$ 0.0002	0.9271 $\pm$ 0.0002	0.9286 $\pm$ 0.0003
Semeion	0.8325 $\pm$ 0.0017	0.8240 $\pm$ 0.0030	0.8090 $\pm$ 0.0084	0.7830 $\pm$ 0.0050	0.7887 $\pm$ 0.0041	0.8171 $\pm$ 0.0032
Wine	0.9485 $\pm$ 0.0081	0.9337 $\pm$ 0.0089	0.9077 $\pm$ 0.0225	0.9075 $\pm$ 0.0049	0.9147 $\pm$ 0.0045	0.9306 $\pm$ 0.0075
Yeast	0.5366 $\pm$ 0.0068	0.5458 $\pm$ 0.0059	0.5550 $\pm$ 0.0048	0.5262 $\pm$ 0.0035	0.5261 $\pm$ 0.0036	0.5443 $\pm$ 0.0050
Zoo	0.8897 $\pm$ 0.0173	0.8846 $\pm$ 0.0161	0.8563 $\pm$ 0.0207	0.8853 $\pm$ 0.0050	0.8922 $\pm$ 0.0049	0.8876 $\pm$ 0.0147
Mean	0.79173	0.79529	0.79511	0.7756	0.77755	0.79396

Since all the RVFL classifiers have the same direct links from the input layer to the output layer as regulartions in the structures, all the variances of the RVFL classification accuracies are similar. Thus, I focus on comparing the mean values of of the RVFL classification accuracies on the 10 datasets. From the bottom row of Table 2, we can easily see that *radbas* > *sine* > *tribas* > *sigmoid* > *hardlim* > *sign*, where ">" means that the RVFL classifiers with the activation function on the left performs better than the RVFL classifiers with the activation function on the right.

3. **Performances of Moore-Penrose pseudoinverse and ridge regression (or regularized least square solutions) for the computation of the output weights.** In order to investigate the performances of Moore-Penrose pseudoinverse and ridge regression for the computation of the output weights, I choose the RVFL with direct links and *radbas* activation function based on the conclusion above. The results are shown in the following Table 3.

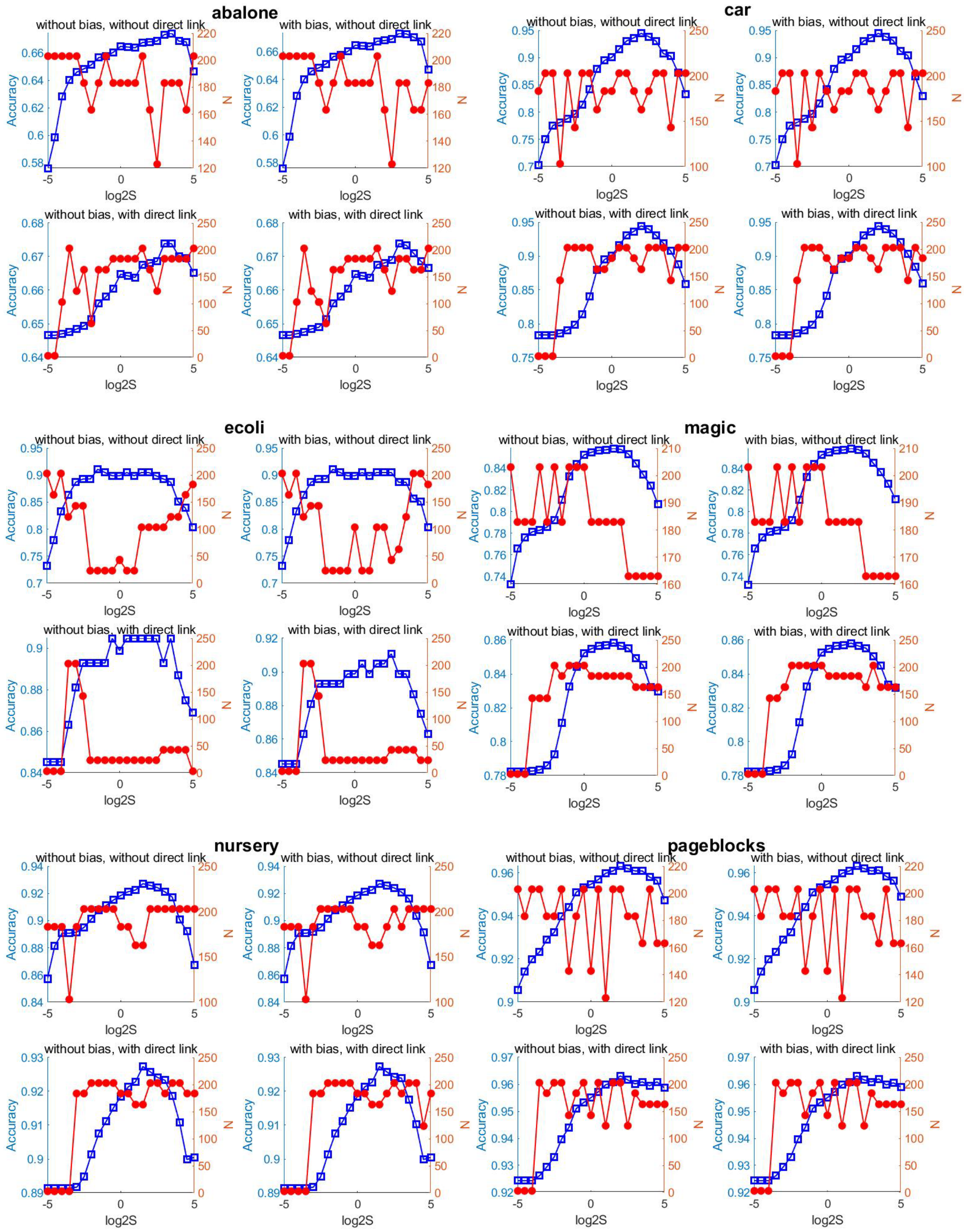
Table 3 : Means and variances of the RVFL classification accuracies based on the condition of the closed-form solution

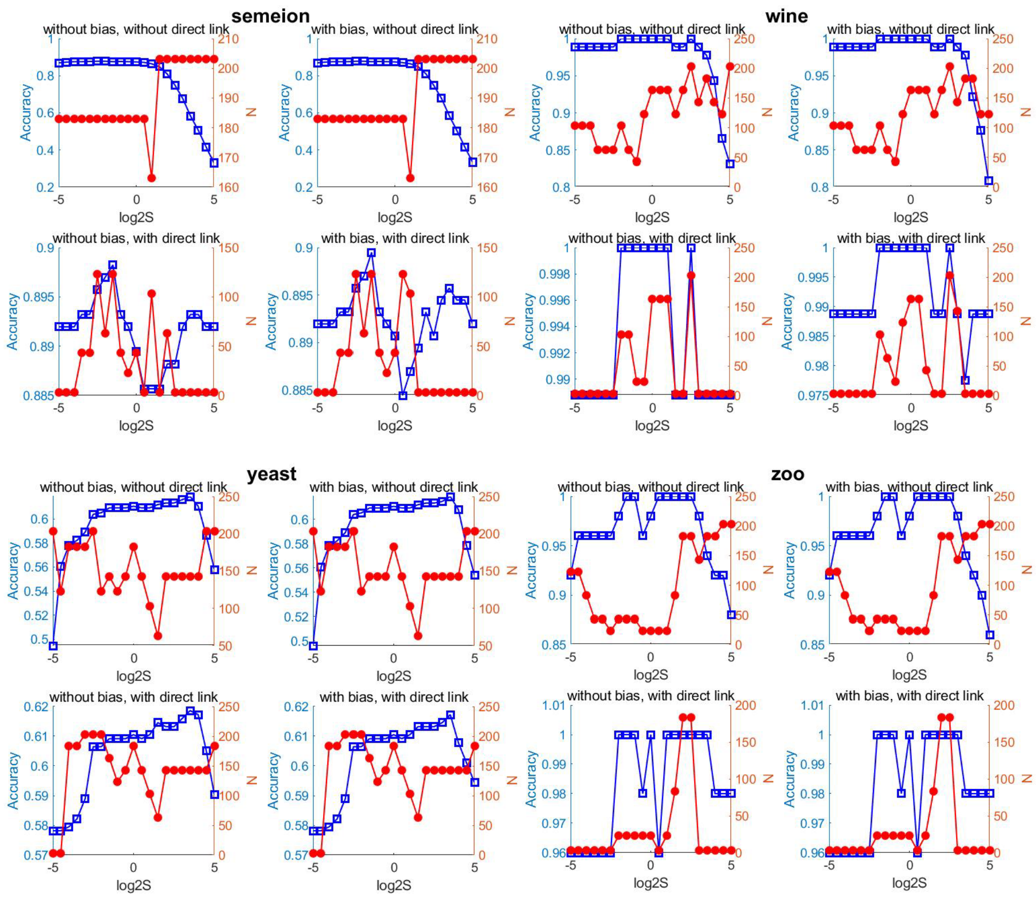
Accurarcy	Ridge regression	Moore-Penrose pseudoinverse
Abalone	0.6271 ± 0.0011	$\mathbf{0.6593 \pm 0.0005}$
Car	0.7875 ± 0.0039	$\mathbf{0.9101 \pm 0.0028}$
Ecoli	$\mathbf{0.7897 \pm 0.0166}$	0.6768 ± 0.0274
Magic	0.7730 ± 0.0025	$\mathbf{0.8431 \pm 0.0006}$
Nursery	0.8610 ± 0.0056	$\mathbf{0.9198 \pm 0.0005}$
Page-blocks	0.9265 ± 0.0003	$\mathbf{0.9540 \pm 0.0001}$
Semeion	$\mathbf{0.8240 \pm 0.0030}$	0.7914 ± 0.0013
Wine	$\mathbf{0.9337 \pm 0.0089}$	0.8155 ± 0.0195
Yeast	0.5458 ± 0.0059	$\mathbf{0.5718 \pm 0.0003}$
Zoo	0.8846 ± 0.0161	$\mathbf{0.9177 \pm 0.0033}$

From Table 3, it is easy to observe that the Moore-Penrose pseudoinverse outperform the ridge regression in 7 datasets (i.e., Abalone, Car, Magic, Nursery, Page-blocks, Yeast, Zoo). Thus, it can be concluded that the Moore-Penrose pseudoinverse can give rise to a better performance based on my ten-dataset experiments. However, due to the small number of the selected datasets, this conclusion is not persuasive and experiemnts with more datasets covered should be done.

4. **Effect of scaling the random features before feeding them into the activation function.** In order to investigate the effect of scaling the random features before feeding them into the activation function, a positive scaling factor $S = 2^t$ where t varies from -5 to 5 with a step of 0.5 is set, and the ridge regression based RVFL with the *radbas* activation function is chosen as the basic RVFL. The results are shown in the following figures.

In the figures, the maximun accuracies with thier suitable numbers of hidden neurons N are shown coreresponding to the scaling factor S with different conditions of bias and direct links. Apparently, the accuracies will drop sharply when the scaling fator S is too large or too small. In the middle range of the scaling factor S, the discrimination ability of the RVFL is postively corroalted with the scaling factor S. Moreover, the over scaling down or up the randomization range of input weights and biases can be compensated by adding the number of the hidden nuerons or implementing the direct links from the input to the output. Due to the consideration of the complexity, the implementation of the direct links from the input to the output is a well recommended operation compared with adding the number of the hidden nuetons.




semeion
wine
yeast
zoo
without bias, without direct link
with bias, without direct link
without bias, with direct link
with bias, with direct link
Accuracy
N
log2S