UPCON2021

UPCON 2021

8th IEEE Uttar Pradesh Section International Conference on Electrical, Electronics, and Computer Engineering

Paper ID: 180

ResCNN:An alternative implementation of Convolutional Neural Networks

Avirup Dey*, Sarosij Bose**

*Dept. of Electronics and Telecommunication. Engineering, Jadavpur University.
**Dept. of Computer Science and Engineering, University of Calcutta.

UPCON 2021

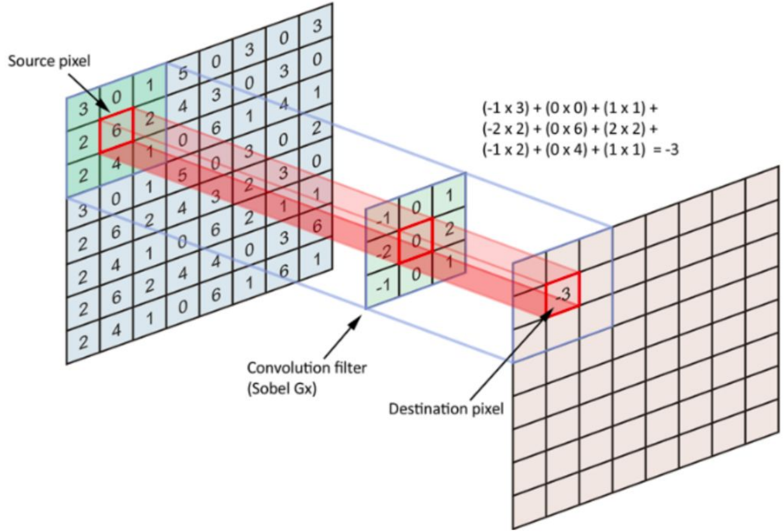
Introduction

Convolutions: An Overview

Convolutions in deep learning have been a goto choice for people to learn latent features of images.

Initial layers learn low level features from images like edges and corners and the deeper we go in the CNN the richer the feature maps become.

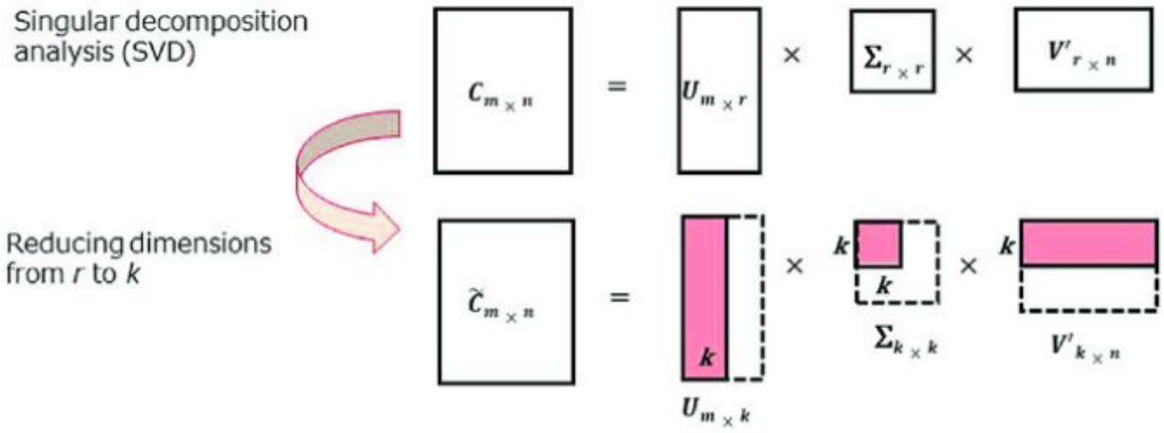
Convolution can be viewed as a kernel sliding over a matrix.



[2]

Introduction

Singular Value Decomposition



[3]

We usually keep the highest singular values because they are enough to approximate the original matrix with reasonable precision.

Introduction

Unrolling

Image Patch 1
```
[ 1.,  2.,  3.,  4.]
[ 5.,  6.,  7.,  8.]
[ 9., 10., 11., 12.]
[13., 14., 15., 16.]
```

Image Patch 2
```
[ 1.,  2.,  3.,  4.]
[ 5.,  6.,  7.,  8.]
[ 9., 10., 11., 12.]
[13., 14., 15., 16.]
```

Image Patch 3
```
[ 1.,  2.,  3.,  4.]
[ 5.,  6.,  7.,  8.]
[ 9., 10., 11., 12.]
[13., 14., 15., 16.]
```

Image Patch 4
```
[ 1.,  2.,  3.,  4.]
[ 5.,  6.,  7.,  8.]
[ 9., 10., 11., 12.]
[13., 14., 15., 16.]
```

Image Patch 5
```
[ 1.,  2.,  3.,  4.]
[ 5.,  6.,  7.,  8.]
[ 9., 10., 11., 12.]
[13., 14., 15., 16.]
```

Image Patch 6
```
[ 1.,  2.,  3.,  4.]
[ 5.,  6.,  7.,  8.]
[ 9., 10., 11., 12.]
[13., 14., 15., 16.]
```

Image Patch 7
```
[ 1.,  2.,  3.,  4.]
[ 5.,  6.,  7.,  8.]
[ 9., 10., 11., 12.]
[13., 14., 15., 16.]
```

Image Patch 8
```
[ 1.,  2.,  3.,  4.]
[ 5.,  6.,  7.,  8.]
[ 9., 10., 11., 12.]
[13., 14., 15., 16.]
```

Image Patch 9
```
[ 1.,  2.,  3.,  4.]
[ 5.,  6.,  7.,  8.]
[ 9., 10., 11., 12.]
[13., 14., 15., 16.]
```

```
[ 1.,  2.,  3.,  5.,  6.,  7.,  9., 10., 11.]
[ 2.,  3.,  4.,  6.,  7.,  8., 10., 11., 12.]
[ 5.,  6.,  7.,  9., 10., 11., 13., 14., 15.]
[ 6.,  7.,  8., 10., 11., 12., 14., 15., 16.]
```

Flattening weight vector

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

```
[1., 2., 3., 4.]
```

Unrolled Input * Flattened Vector

```
[ 44.,  54.,  64.,  84.,  94., 104., 124., 134., 144.]
```

Rolling back the output:

```
[ 44.,  54.,  64.]
[ 84.,  94., 104.]
[124., 134., 144.]
```

UPCON 2021

Problem Statement

1. **Improve performance of convolution operation**

   The computational complexity of the convolution layers stems from the convolution operation with small kernels and cache unfriendly memory access[1]. On the other hand a matrix product makes better use of memory. We intend to bridge the gap between these operations.

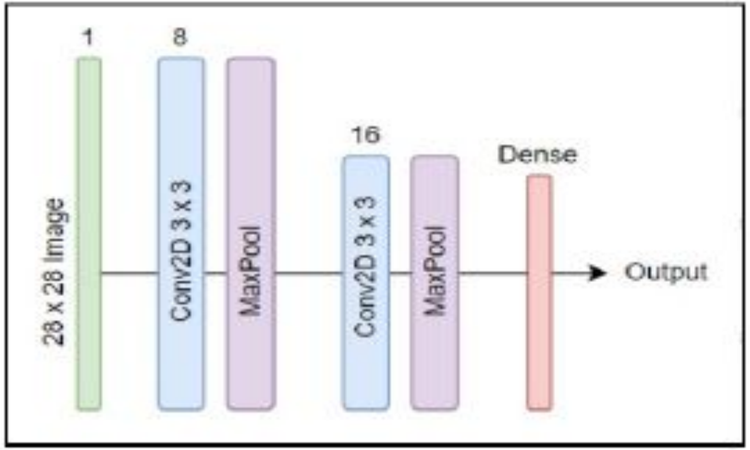
1. **Explore SVD as a preprocessing step to optimise memory usage**

   We know that any image can be be represented using lesser singular values thus cutting down memory overheads. We intend to study how the performance of a neural network varies with such decompositions.

Proposed Method

Approach to study the effect of unrolling

To make a fair comparison between the conventional convolution and our matrix product based approach we will keep the architecture of the CNN same but merely replace the convolution layers in ResCNN with our own.

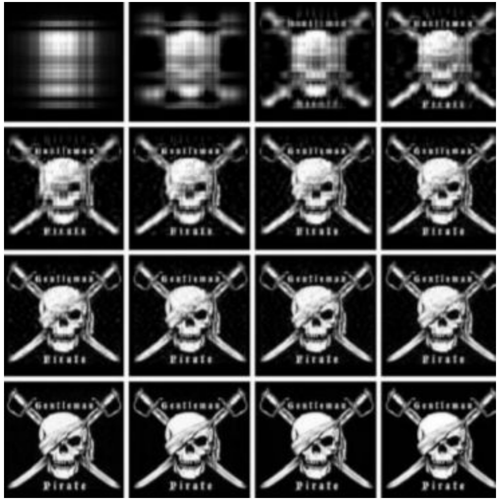


UPCON 2021

Proposed Method

Approach to study the effect of SVD

Varying the number of singular values changes the quality as shown in the image.
Our motive was to find the optimum rank that yields maximum accuracy at a high learning rate.



[4]

UPCON2021

Novelty

1. Our specific implementation (ResCNN) can learn at much higher rates (7x) without compromising on accuracy.
2. We have used SVD, a novel step in pre-processing, which helps in shrinking the size of the dataset, enabling the model to accommodate higher batch sizes.
3. We have shown the variation of model accuracy with the degree of decomposition of the training dataset. This helps us to find the optimum rank for maximising accuracy of the model.

Experimental Setup

1. We compare the performances of the conventional CNN and ResCNN on the metric of training accuracy at different learning rates.

2. The batch-size was chosen as 128 for optimum performance of the models.

3. We noted that a few epochs were enough for the chosen model architecture to achieve over 98% accuracy..

4. To correct variations due to weight initializations in accuracy we averaged the results over 10 runs at each chosen value for the learning rate.

Experiment Results

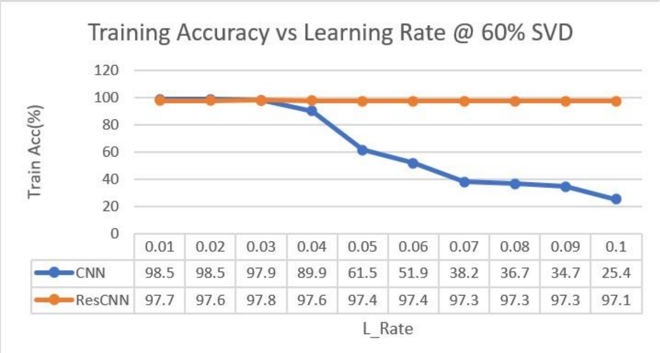
We trained both the models on MNIST [5] handwritten digits dataset and compared their accuracies at different learning rates. Our implementation outperformed the conventional CNN at higher learning rate.

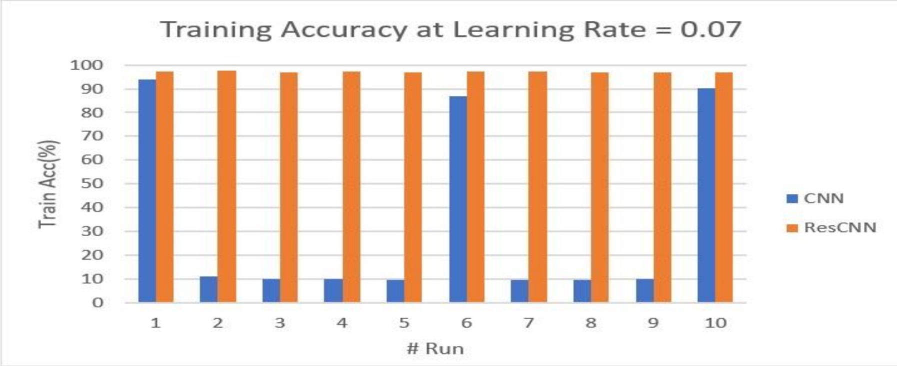
Learning Rate	Acc. of CNN (%)	Acc. of ResCNN (%)
0.01	98.5	97.7
0.02	98.5	97.6
0.03	97.9	97.8
0.04	89.9	97.6
0.05	61.5	97.4
0.06	51.9	97.4
0.07	38.2	97.3
0.08	36.7	97.3
0.09	34.7	97.3
0.10	25.4	97.1

UPCON 2021

Results Analysis and Outcome

CNN vs ResCNN: Comparison





UPCON 2021

Results Analysis and Outcome

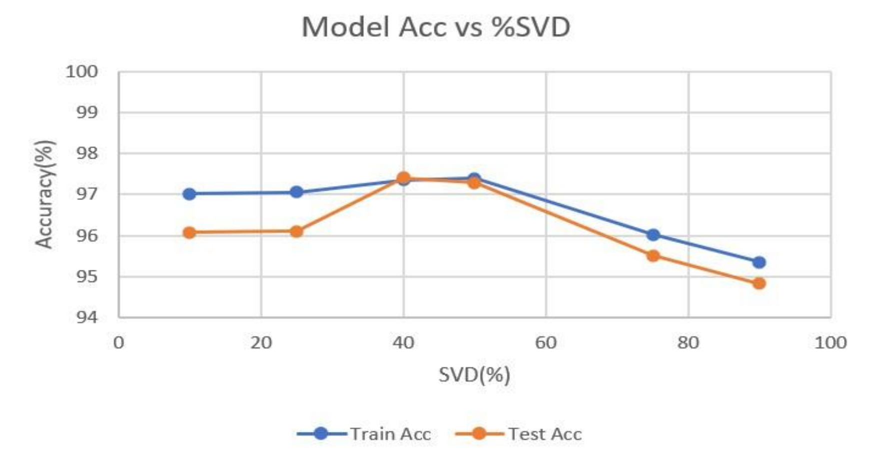
Effect of SVD on ResCNN performance

The accuracy of the model peaks at 50% SVD.

As we increase the SVD, the rank of input shrinks and the model can learn all parameters well. Hence, we see an improvement in accuracy with SVD.

On the other end we see that at high SVD the model accuracy drops. This is because the image loses its salient features at such high levels of decomposition.

While applying the decomposition this needs to be kept in mind.



UPCON 2021

Response to Reviewer- 1 Comments

The introduction section has been updated and we have focussed more on our own contributions and removed the portions which seemed less relevant to us.

We have revised the literature overview and introductions as said.

We ran our experiments on MNIST which is a publicly available dataset so that anyone could replicate our findings. Hence we did not encounter any major problem while working with the dataset.

The implementation of convolution that we have presented in the paper hasn't been sufficiently worked upon in the past. However, we have compared our model with the generic version of neural network that we use. Our comparative study is supported with a sufficient number of plots.

Modified the introduction of the paper and our results accordingly. As we have explicitly noted pointwise in the introduction, our intended goal was to design a more efficient and robust CNN.

The outcomes of our work demand parallel comparison with the existing CNNs so that readers can understand the advantage of our implementation clearly (as in Table 1 and in Figs 7 and 8). We have also presented our results keeping the reader's best interest in mind.

UPCON 2021

Response to Reviewer- 2 Comments

Thank you for taking the time to review our paper. We have updated the citations accordingly.

We designed the experiment with utmost care and sufficient checks were enforced to prevent inconsistencies. Results were averaged across multiple runs and necessary plots have been provided for the reader's references. The results are consistent with the goals mentioned in the introduction.

We have incorporated more diagrams to better elucidate our approach, modified the text to a higher quality to reflect our intended methodology better.

UPCON2021

Response to Reviewer- 3 Comments

We have updated the Introduction and conclusion accordingly to reflect the reviewer's concerns and represent our work in a more decisive manner.

In line with the reviewer's note, we have identified some limitations of our proposed approach in the conclusion section.

We also discuss the probable memory constraints which may arise due to the size of the unrolled image matrices.

As the reviewer may note, we have now explicitly mentioned our novel contributions in this paper in the Introduction itself.

Discussions in the results section and conclusion have been revised accordingly and there are no inconsistencies/ambiguities in the same. Necessary changes have been made to these portions to make it more clear and concise.

We have written about some real life industrial applications a robust model may be needed.

Table 1 compares the performance of our robust model (ResCNN) versus a traditional CNN when subjected to high learning rates. The learning rate is gradually increased as one goes down the Table. This table clearly brings out the manner in which the former just breaks down in terms of accuracy whereas our ResCNN still performs strongly in spite of being subjected to 7x higher learning rates implying the robustness in our proposed implementation.

UPCON 2021

Conclusion

- Our model operates at 7x higher learning rate compared to conventional CNN.
- Using SVD as a preprocessing step helps the neural network to learn the image features under high learning rates. This implies that the network would need lesser epochs to train.
- The entire experiment was done on MNIST datasets but this approach needs further testing on bigger datasets.

UPCON2021

References

1. Kumar Chellapilla, Sidd Puri, and Patrice Simard. High performance convolutional neural networks for document processing. In Tenth international workshop on frontiers in handwriting recognition. Suvisoft, 2006.

2. https://medium.com/@_init_/an-illustrated-explanation-of-performing-2d-convolutions-using-matrix-multiplications-1e8de8cd2544

3. Virginia Klema and Alan Laub. The singular value decomposition: Its computation and some applications. IEEE Transactions on automatic control, 25(2):164–176, 1980.

4. Brady Mathews. Image compression using singular value decomposition (svd). Univ. Utah, 2014.

5. Yann LeCun. The mnist database of handwritten digits. http://yann. lecun. com/exdb/mnist/, 1998.

THANK YOU!