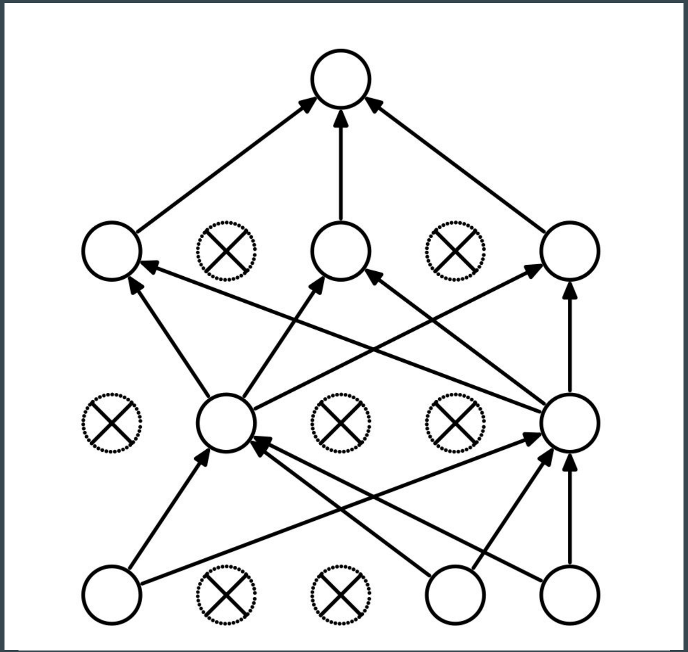
Exploring Dropout: A Powerful Strategy for Overfitting Prevention in Neural Networks

Ece Akdeniz 2004229
Ata Deniz Arslaner 2004040

Table of Content

Introduction

- **Overfitting** in neural networks occurs when the model becomes **too complex** and **tightly fits** the training data, resulting in **poor generalization** to unseen data.
- The paper titled "Dropout: A Simple Way to Prevent Neural Networks from Overfitting" proposes a technique called **dropout** that addresses the problem of overfitting by randomly deactivating neurons during training.
- This project aims to:
  - **reproduce** the results of the paper on the CIFAR-10 dataset by implementing dropout
  - **compare** it with other benchmark regularization methods.



Before and After Dropout

Related Works

- **Batch normalization** normalizes the activations of intermediate layers within a mini-batch during training, reducing internal covariate shift and stabilizing the network.
- **Data augmentation** artificially increases the size of the training set by applying various transformations to introduce diversity and reduce overfitting.

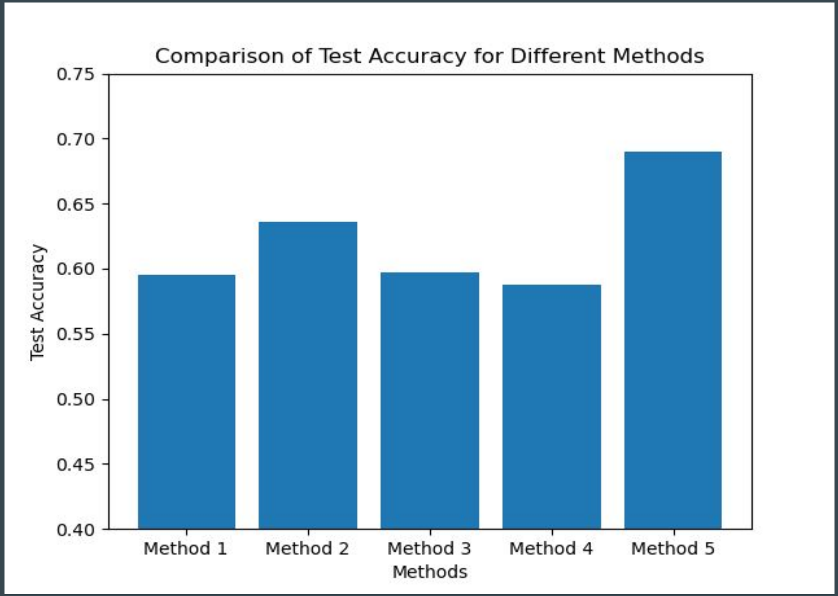
- However, these methods have limitations, such as increased computational overhead and task-dependence.

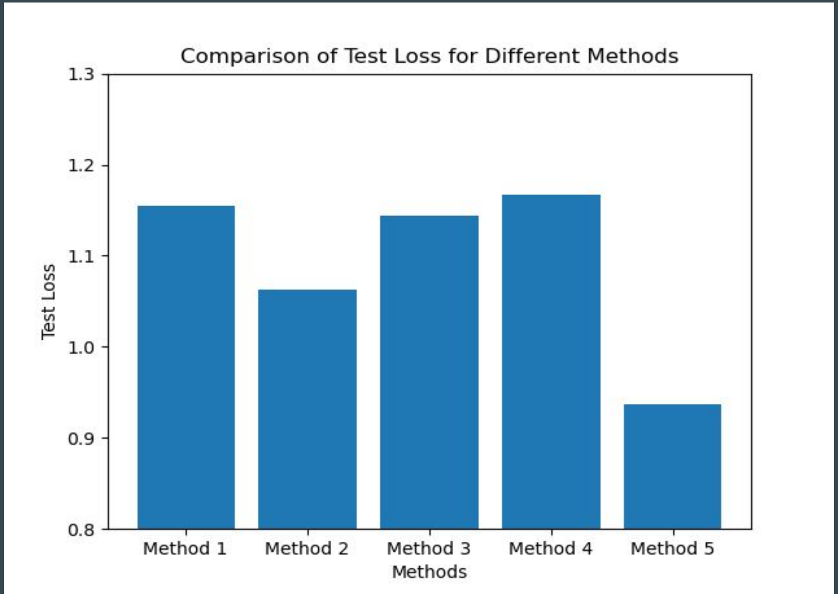
Material and Methods

- The model architecture used in the project follow the description in the paper.
- The CIFAR-10 dataset, which comprises 50,000 32x32 color images categorized into 10 classes, is used for training and testing.
- Stochastic gradient descent with backpropagation is used for training, and the evaluation metric used is likely accuracy.

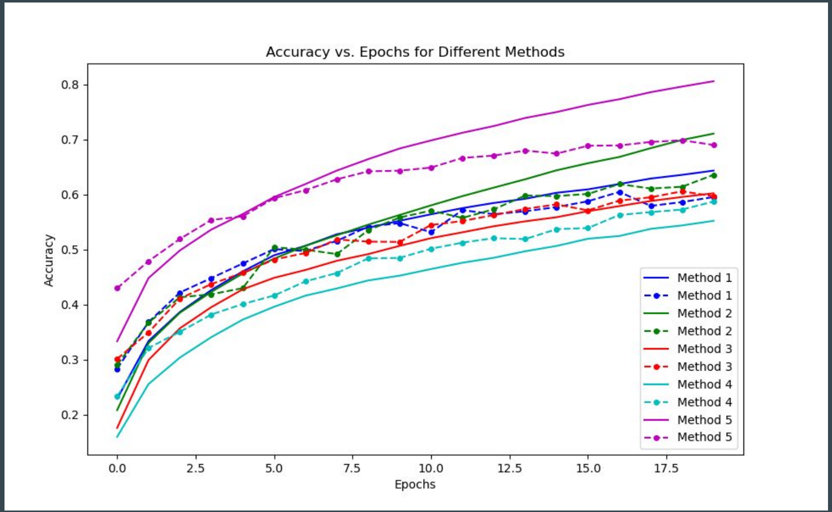
Method 1	CNN + Max Pooling
Method 2	CNN + Max Pooling
Method 3	CNN + Max Pooling + Dropout in Dense Layers
Method 4	CNN + Max Pooling + Dropout in All Layers
Method 5	CNN + Max Pooling + Maxout

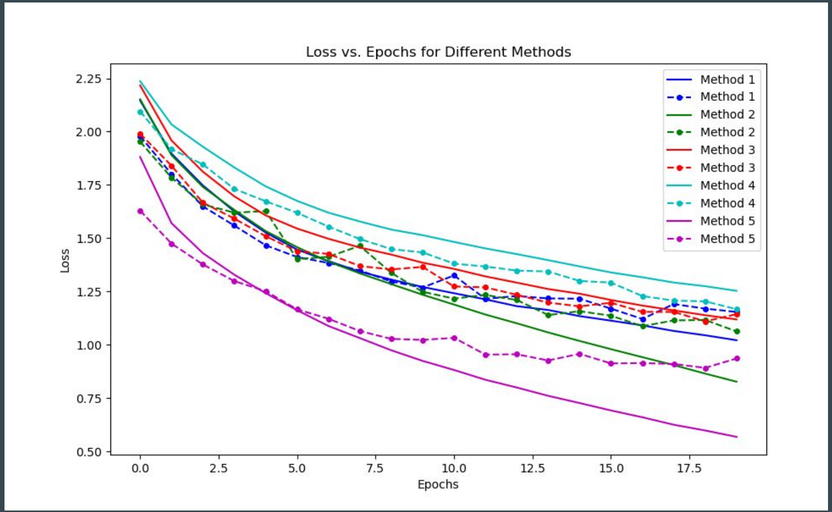
Results - Comparison of 5 Methods





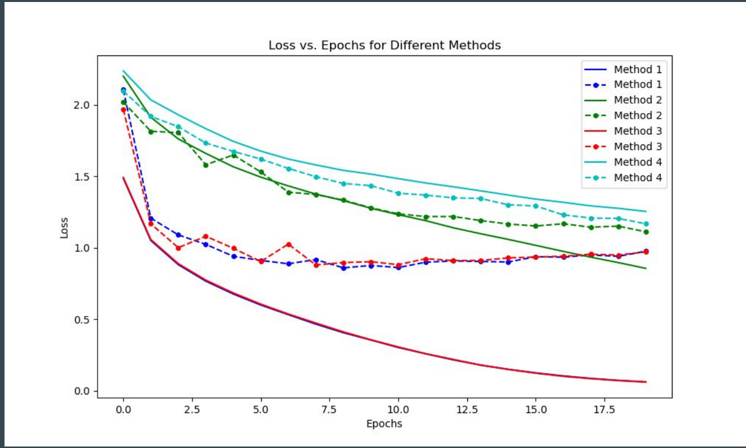
Results - Train vs. Validation Scores of Each Model

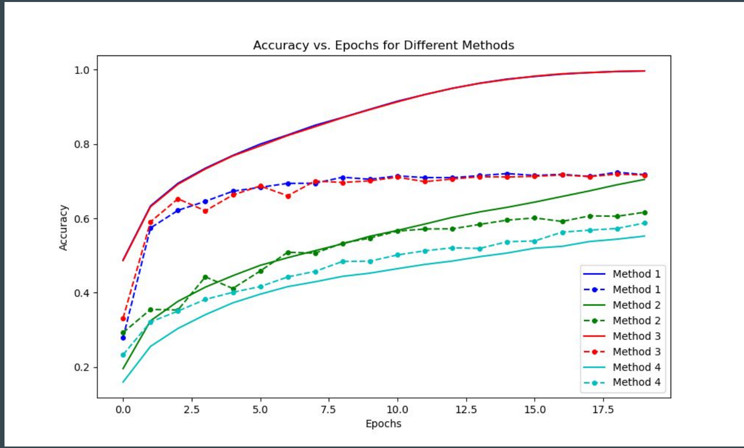




Results - Comparison of Other Regularization Methods

Method 1	Batch Normalization
Method 2	Data Augmentation
Method 3	Batch Normalization + Data Augmentation
Method 4	Dropout





Discussions

- The impact of dropout on neural network performance is analyzed.
- Dropout provides regularization without significantly increasing training time, striking a balance between regularization and efficiency.
- Dropout fosters diverse feature learning, improving the network's ability to handle unseen data.
- The limitations of dropout, such as increased randomness and reduced training capacity, are discussed.

Conclusion

- The project demonstrates the effectiveness of dropout in preventing overfitting in neural networks.
- Dropout consistently improves performance, reduces overfitting, and enhances generalization.
- Dropout strikes a balance between regularization and efficiency, providing practical and efficient regularization without significant training time overhead.
- Future research can explore extending dropout to other models or improving training time efficiency to further enhance neural network performance.

References

M. D. Zeiler and R. Fergus. Stochastic pooling for regularization of deep convolutional neural networks. CoRR, abs/1301.3557, 2013. J. Snoek, H. Larochelle, and R. Adams. Practical Bayesian optimization of machine learning algorithms. In Advances in Neural Information Processing Systems 25, pages 2960–2968, 2012.

I. J. Goodfellow, D. Warde-Farley, M. Mirza, A. Courville, and Y. Bengio. Maxout networks. In Proceedings of the 30th International Conference on Machine Learning, pages 1319– 1327. ACM, 2013.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R. (2014). Dropout: A Simple Way to Prevent Neural Networks from Overfitting. The journal of machine learning research, 15(1), 1929-1958.

Thank You!