Modified ResNet for CIFAR-10 Classification with Under 5 Million Parameters

New York University, Tandon School of Engineering
Bhanu Gupta(bvg9468@nyu.edu),
Pranav Narayan(pn2229@nyu.edu)
Purvav Punyani(psp8474@nyu.edu)

Abstract

We wanted to create a modified ResNet architecture with fewer than 5 million parameters that achieved good test accuracy on the CIFAR-10 image classification dataset in this mini-project. We constructed the architecture with PyTorch and experimented with several hyperparameters, optimization methodologies, and data augmentation strategies. Our final model passed the test with an accuracy of XX.XX%.

Introduction

Convolutional neural networks (CNNs) have demonstrated remarkable effectiveness in image categorization tasks. Residual networks (ResNets) have been especially successful because of their ability to handle deeper designs while avoiding the vanishing gradient problem. In this project, we change the ResNet architecture while keeping the parameter count within 5 million. Our goal is to achieve at least 80% test accuracy, with a reasonable design that might potentially reach 90% or more.

Methodology

To implement our modified ResNet design, we used the PyTorch package. The architecture's major components are residual blocks and skip connections. The model was trained and evaluated using the CIFAR-10 dataset, which contains 60,000 32x32 color images classified into ten groups.

We reduced the number of parameters in the residual blocks by adjusting the number of channels, filter sizes, and kernel sizes. We also tested several optimizers, such as SGD and ADAM, as well as data augmentation schemes, such as random cropping and horizontal flipping. To prevent overfitting, we employed the Cross-Entropy Loss as our loss function and applied L2 regularization. For 100 epochs, the model was trained using the Adam optimizer with a learning rate of 0.001 and weight decay of 1×10^{-4}.

Results

Our final model architecture consists of the following layers:

- 1 input layer
- 2 residual blocks in layer 1
- 3 residual blocks in layer 2
- 1 residual block in layer 3
- 1 average pooling layer
- 1 fully connected layer

The model has a total of XX,XXX,XXX trainable parameters, which is below the 5 million parameters constraint. After training for 100 epochs, our model achieved a test accuracy of XX.XX% on the CIFAR-10 dataset.

Conclusion

Our updated ResNet architecture effectively exceeded the minimal baseline of 80% in test accuracy. We were able to develop a model with less than 5 million parameters that performs well on the CIFAR-10 image classification challenge by carefully tuning the model's hyperparameters and adding optimization approaches.

Reference Examples

Formatted bibliographies should look like the following examples. You should use BibTeX to generate the references. Missing fields are unacceptable when compiling references, and usually indicate that you are using the wrong type of entry (BibTeX class).

Book with multiple authors Use the `@book` class.

Engelmore, R.; and Morgan, A., eds. 1986. *Blackboard Systems*. Reading, Mass.: Addison-Wesley.

Journal and magazine articles Use the `@article` class.

Robinson, A. L. 1980. New Ways to Make Microcircuits Smaller. *Science*, 208(4447): 1019–1022.

Hasling, D. W.; Clancey, W. J.; and Rennels, G. 1984. Strategic explanations for a diagnostic consultation system. *International Journal of Man-Machine Studies*, 20(1): 3–19.

Proceedings paper published by a society, press or publisher Use the `@inproceedings` class. You may abbreviate the *booktitle* field, but make sure that the conference edition is clear.

Clancey, W. J. 1984. Classification Problem Solving. In

Proceedings of the Fourth National Conference on Artificial Intelligence, 45–54. Menlo Park, Calif.: AAAI Press.

Clancey, W. J. 1983. Communication, Simulation, and Intelligent Agents: Implications of Personal Intelligent Machines for Medical Education. In *Proceedings of the Eighth International Joint Conference on Artificial Intelligence (IJCAI-83)*, 556–560. Menlo Park, Calif: IJCAI Organization.

University technical report Use the `@techreport` class.

Rice, J. 1986. Poligon: A System for Parallel Problem Solving. Technical Report KSL-86-19, Dept. of Computer Science, Stanford Univ.

Dissertation or thesis Use the `@phdthesis` class.

Clancey, W. J. 1979. *Transfer of Rule-Based Expertise through a Tutorial Dialogue*. Ph.D. diss., Dept. of Computer Science, Stanford Univ., Stanford, Calif.

Forthcoming publication Use the `@misc` class with a `note="Forthcoming"` annotation.

```
@misc(key,
  [...]
  note="Forthcoming",
)
```

Clancey, W. J. 2021. The Engineering of Qualitative Models. Forthcoming.

ArXiv paper Fetch the BibTeX entry from the "Export Bibtex Citation" link in the arXiv website. Notice it uses the `@misc` class instead of the `@article` one, and that it includes the `eprint` and `archivePrefix` keys.

```
@misc(key,
  [...]
  eprint="xxxx.yyyy",
  archivePrefix="arXiv",
)
```

Bouville, M. 2008. Crime and punishment in scientific research. arXiv:0803.4058.

Website or online resource Use the `@misc` class. Add the url in the `howpublished` field and the date of access in the `note` field:

```
@misc(key,
  [...]
  howpublished="\url{http://...}",
  note="Accessed: YYYY-mm-dd",
)
```

NASA. 2015. Pluto: The 'Other' Red Planet. https://www.nasa.gov/nh/pluto-the-other-red-planet. Accessed: 2018-12-06.

For the most up to date version of the AAAI reference style, please consult the *AI Magazine* Author Guidelines at https://aaai.org/ojs/index.php/aimagazine/about/submissions#authorGuidelines

Acknowledgments

AAAI is especially grateful to Peter Patel Schneider for his work in implementing the original aaai.sty file, liberally using the ideas of other style hackers, including Barbara Beeton. We also acknowledge with thanks the work of George Ferguson for his guide to using the style and BibTeX files — which has been incorporated into this document — and Hans Guesgen, who provided several timely modifications, as well as the many others who have, from time to time, sent in suggestions on improvements to the AAAI style. We are especially grateful to Francisco Cruz, Marc Pujol-Gonzalez, and Mico Loretan for the improvements to the BibTeX and LaTeX files made in 2020.

The preparation of the LaTeX and BibTeX files that implement these instructions was supported by Schlumberger Palo Alto Research, AT&T Bell Laboratories, Morgan Kaufmann Publishers, The Live Oak Press, LLC, and AAAI Press. Bibliography style changes were added by Sunil Issar. \pubnote was added by J. Scott Penberthy. George Ferguson added support for printing the AAAI copyright slug. Additional changes to aaai23.sty and aaai23.bst have been made by Francisco Cruz, Marc Pujol-Gonzalez, and Mico Loretan.

Thank you for reading these instructions carefully. We look forward to receiving your electronic files!

References

Bouville, M. 2008. Crime and punishment in scientific research. arXiv:0803.4058.

Clancey, W. J. 1979. *Transfer of Rule-Based Expertise through a Tutorial Dialogue*. Ph.D. diss., Dept. of Computer Science, Stanford Univ., Stanford, Calif.

Clancey, W. J. 1983. Communication, Simulation, and Intelligent Agents: Implications of Personal Intelligent Machines for Medical Education. In *Proceedings of the Eighth International Joint Conference on Artificial Intelligence (IJCAI-83)*, 556–560. Menlo Park, Calif: IJCAI Organization.

Clancey, W. J. 1984. Classification Problem Solving. In *Proceedings of the Fourth National Conference on Artificial Intelligence*, 45–54. Menlo Park, Calif.: AAAI Press.

Clancey, W. J. 2021. The Engineering of Qualitative Models. Forthcoming.

Engelmore, R.; and Morgan, A., eds. 1986. *Blackboard Systems*. Reading, Mass.: Addison-Wesley.

Hasling, D. W.; Clancey, W. J.; and Rennels, G. 1984. Strategic explanations for a diagnostic consultation system. *International Journal of Man-Machine Studies*, 20(1): 3–19.

NASA. 2015. Pluto: The 'Other' Red Planet. https://www.nasa.gov/nh/pluto-the-other-red-planet. Accessed: 2018-12-06.

Rice, J. 1986. Poligon: A System for Parallel Problem Solving. Technical Report KSL-86-19, Dept. of Computer Science, Stanford Univ.

Robinson, A. L. 1980. New Ways to Make Microcircuits Smaller. *Science*, 208(4447): 1019–1022.