Comparative study of AlexNet & ResNet

Tarasov Andrei, Potekhin Yuriy, Kalmykov Maxim
Mentor: Potapov Yurii

Motivation

Modern neural networks easily solve image classification tasks, but numerous architectures exist. Our goal is to select the optimal architecture for Tiny ImageNet by comparing AlexNet and ResNet, as well as to gain a practical understanding of neural and convolutional network architectures and their implementation in PyTorch. This research is based on the papers presented in **[1]** and **[2]**.

AlexNet

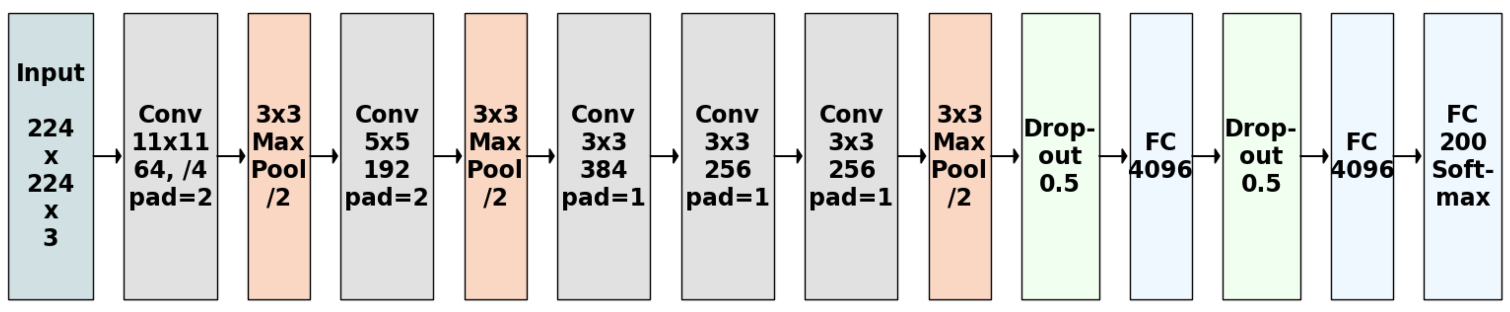

Figure 1: AlexNet architecture

AlexNet is a convolutional neural network consisting of 5 convolutional layers followed by 3 fully connected layers. It uses ReLU activation functions. To reduce overfitting, AlexNet employs Data Augmentation and Dropout in the fully connected layers. The network also uses overlapping Max Pooling to improve feature extraction. The classic implementation takes a 224×224 image as input and returns a probability vector of length 1000. In our experiment, the classifier is replaced with a vector of length 200.

The model we trained contains **57,823,240** parameters.

ResNet

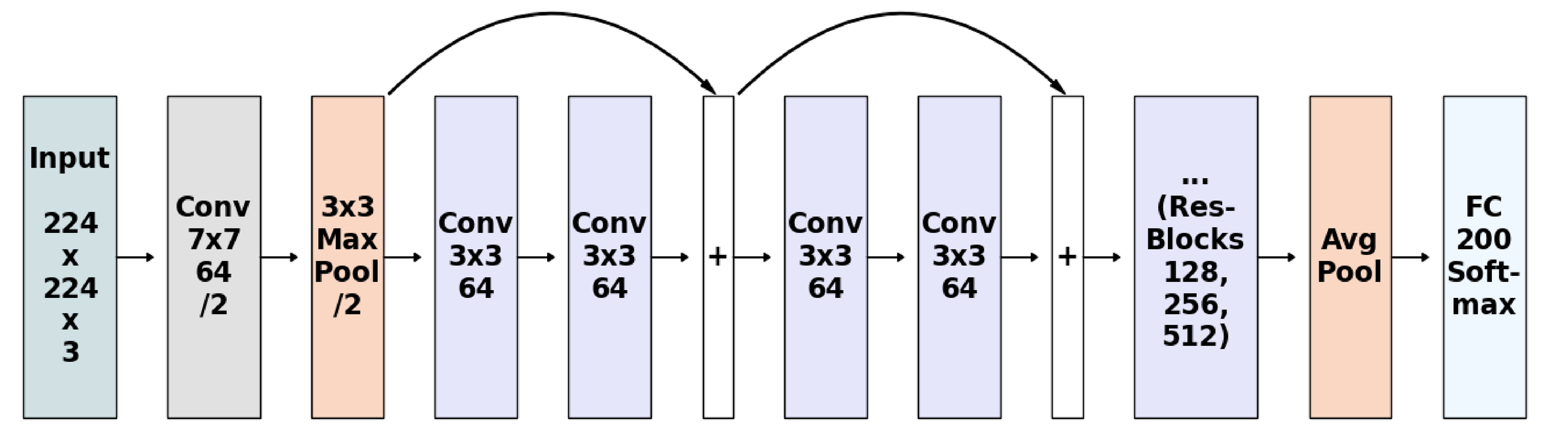

Figure 2: ResNet18 architecture

ResNet-18 is a deep convolutional neural network consisting of 18 layers organized into a series of residual blocks. Each residual block contains two convolutional layers with batch normalization and ReLU activations, as well as a skip connection that adds the block's input to its output. These skip connections help mitigate the vanishing gradient problem and allow for much deeper networks to be trained.

The model we train contains **11,279,112** parameters.

Dataset

Tiny ImageNet contains 100,000 images of 200 classes (500 for each class) downsized to 64×64 colored images. Each class has 500 training images, 50 validation images, and 50 test images.

We applied standard **augmentations**, including Random Horizontal Flip (p = 0.5), Color Jitter with brightness (low = 0.9, high = 1.08) and contrast (low = 0.9, high = 1.08), Random Resized Crop (size = 64, scale range = $(0.8, 0.95)$).

Figure 3: Data augmentations

Related works

Work **[3]** introduced the **Tiny ImageNet** dataset and reported the performance of various models on the classification task. Using a convolutional model with a comparable number of layers and parameters to AlexNet, the authors achieved a top-1 accuracy of **0.42**. In a 2025 paper **[4]**, it was demonstrated that a baseline ResNet-18 model without pretrain attains a top-1 accuracy of **0.55** on the **Tiny ImageNet**.

Results

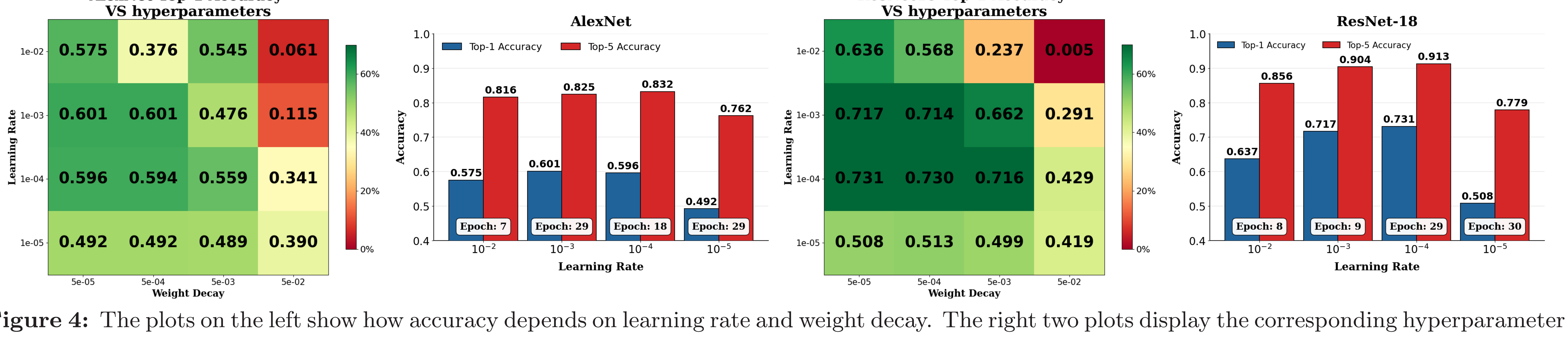

Figure 4: The plots on the left show how accuracy depends on learning rate and weight decay. The right two plots display the corresponding hyperparameter sensitivity analysis for ResNet.

We implemented the Alexnet and Resnet architectures described in **[1]** and **[2]** on the plain pytorch framework. We trained these neural networks on the Tiny ImageNet dataset.

During our research, we studied how various data augmentations and hyperparameter values affect the training process. Without applying data augmentation, AlexNet scored **0.3747** and ResNet scored **0.4077**.

Additionally, we explored the possibility of using pre-trained weights to solve our problem. The best result on non pre-trained weights was significantly lower: **0.416** top-1 accuracy for AlexNet and **0.519** top-1 accuracy for ResNet.

Conclusions

Our experiments showed that **ResNet outperforms AlexNet**, which is probably due to the use of skip connections, despite ResNet having fewer parameters.

In addition, we observed that pre-trained weights significantly improve performance by accelerating convergence and improving final accuracy. Data augmentation also had a positive effect, although its impact was relatively significant.

Finally, the results obtained in our implementation largely coincide with those presented in previously published literature, indicating the correctness and reproducibility of our configuration and training methodology.

References

[1] **Krizhevsky, Alex and Sutskever, Ilya and Hinton, Geoffrey E. Imagenet classification with deep convolutional neural networks**. *Advances in neural information processing systems*, 25, 2012.

[2] **He, Kaiming and Zhang, Xiangyu and Ren, Shaoqing and Sun, Jian. Deep residual learning for image recognition**. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[3] Yann Le and Xuan Yang. Tiny imagenet visual recognition challenge. *CS 231N*, 7(7):3, 2015.

[4] Aidar Amangeldi, Angsar Taigonyrov, Muhammad Huzaid Jawad, and Chinedu Emmanuel Mbonu. Cnn and vit efficiency study on tiny imagenet and dermamnist datasets. *arXiv preprint arXiv:2505.08259*, 2025.