IMAGE CLASSIFICATION WITH CIFAR 10 DATASET

Y. Lavanya[1], U. Lahari[2], E.Riya[3], S.J.K.Manideep[4], K.Vamsi[5]

[1]Associate Professor, Department of Electronics and Communication Engineering, RCEE, Eluru.

[2,3,4,5] Students, Department of Electronics and Communication Engineering, Ramachandra College ofEngineering, Eluru.

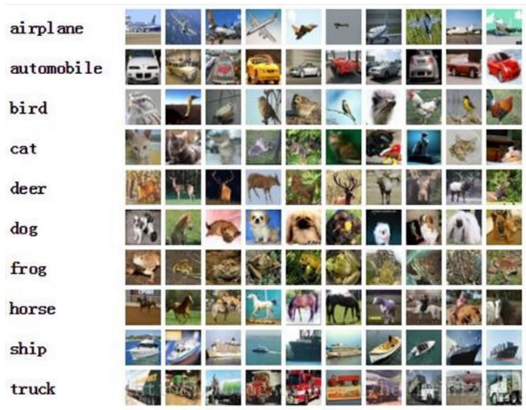
ABSTRACT

In this project, we work on image classification of the CIFAR-10 data set using supervised machine learning techniques. The dataset consists of 60,000 32x32RGB images containing one of 10 object classes, with 6000 images per class. We experiment with various learning algorithms including nearest neighbour classifier, one-vs-all classification, Soft max classifier, two-layer fully connected artificial neural network (ANN), deep convolutional neural network (CNN), and deep residual networks (Res Net). We use cross-validation by splitting the 50,000training data into49,000 training samples and 1,000 validation samples to select the optimized hyperparameters for each parametric classifier. Among all methods, the 56-layer deep residual network yields the best performance with a training accuracy above 99% and a validation accuracy of 93.6%.
Key words: image classification; CIFAR-10; supervised machine learning algorithm; deep convolutional neural network (CNN); deep residual network (Res Net).

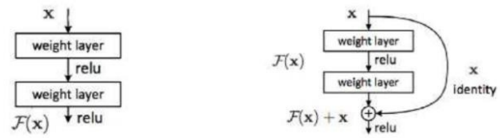
1. INTRODUCTION

In this project, the CIFAR-10 dataset was divided into 50,000 labeled training images and 10,000 unlabeled testing images. We further divided the training set into 49,000 training samples and 1,000 validation samples to select the best model and hyper parameters. The classifier trained on the49,000 samples with the optimized parameters was then used to predict on the testing set and evaluate the prediction accuracy. During this project, we experimented with both non-parametric and parametric methods for image classification. We started with then on-parametric nearest neighbour classifier (NN), which only gave a validation accuracy around23.9%.

Then we experimented with various parametric methods, which included both linear and non- linear classifiers. The linear methods such as One-vs-all classifier and Softmax classifier had validation accuracy around 40%. The non- linear methods, on the other hand, gave a validation accuracy from 50% (two-layer artificial neural net-work (ANN)) to 80% (plain convolutional neural network (CNN)). Finally, we experimented with the deep residual networks (ResNet)6 . After carefully designing the network structures and tuning the hyper parameters, we got the best performance with a 56-layer ResNet. The network showed a training accuracy above 99%, a validation accuracy of 93.6% and a testing accuracy of92.58%. The ResNet is therefore chosen as the final candidate for this project.



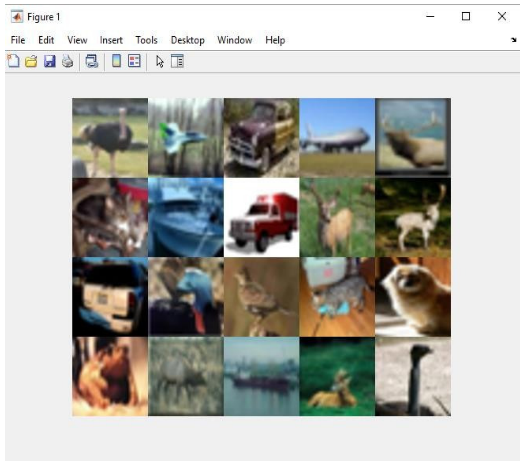
2. RELATED WORK- The overall approach for a medical imaging research area and related data analysis strategies is laid out. This approach aided in formatting the overall research process. Augmentation of data to supplement machine learning methods has been done since the field's conception. More specific studies on exactly which techniques perform given different circumstances will be discussed. Sampling is the only augmentation strategy considered that does not manipulate the underlying data; rather it is strategically removing or repeating pre-existing samples in the dataset. This is usually a first step in dealing with imbalanced or limited data and can take on many different forms. The different methods of sampling and their respective classification performance on unbalanced datasets are explored in. Another popular sampling method for imbalanced datasets is adaptive synthetic (ADASYN) which is presented in. When it comes to medical imaging, augmentation is almost always used in classification to help make networks more robust to the various noise introduced in different collection environments. A comparative analysis of augmentation methods and their impact on classification performance for mammographyimages is carried out in. In this study, classification networks of different configurations are trained with medical images in an attempt to predict accurate diagnoses. The base network used in these experiments is a pre-trained ResNet-50, on ImageNet, and the datasets utilized are publicly available with descriptions of each in Chapter II. The first course of action was establishing a baseline performance to build the comparative analysis from and ensure the methods were applicable. CIFAR-10 [3], the widely known and studied image classification dataset was used for the baseline dataset. After establishing the control result, an array of different augmentations is appliedto the data. The different sampling and augmentation methods are analyzed across artificially imposed imbalance ratios of the CIFAR-10 dataset in or der to rank performance on a well-studied dataset. This study also aided with empirically choosing a control strategy for augmentation/sampling. After initial augmentation studies, well-established deep learning classification methods are applied to the CIFAR-10 data with artificial imbalance and two medical imaging datasets. Once the baseline is calculated for each dataset, the ensemble deep learning (DL) methods are trainedand tested on the datasets. Four different classification systems with various augmentation and sampling configurationsare comparatively analyzed based on the same set of metrics in defined in Table 5.1-2. The focus of this approach is to uncover any combination of increasing complexity of classification architecture and augmentation approaches that perform better and in which scenarios. Some visualization are also calculated from the feature extraction layers of the networks in order to contextualize and better understand the classification results.

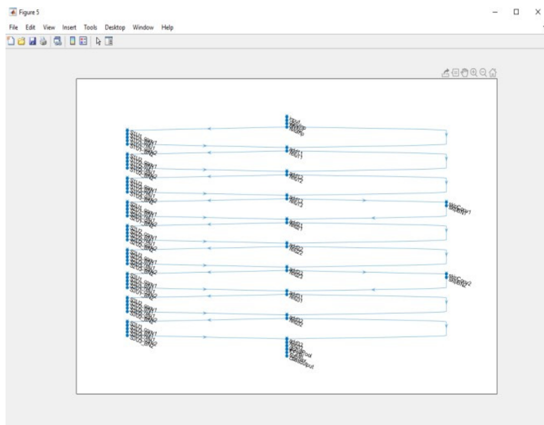


Deep Residual Network (ResNet)- Thegeneral convolutional neural network (CNN) consists of multiple layers that transform the input image volume into an output volume holding the class scores.The several distinct types of layers are convolutional layer, RELU layer, POOL layer and fully-connected layer. For a convolutional neural network, the most important layer is the convolution layer, where each entry in the output volume canbe interpreted as an output of a neuron thatlooks at only a small region in the input and shares parameters with neurons in the same activation map7. A lot of breakthroughs for image classification have been made with deep CNN by stacking more and more convolutional layers. However, when deeper networksare able to start converging, the degradation problem occurs6. The accuracy gets saturated and then degrades rapidly, and adding more layers to a suitable deep model will lead to highertraining error. The degradation problem was addressed by KaimingHe, et al by introducing a deep residual learning framework6. In this approach, instead of directly fitting the desired mapping, the layers are made to explicitly fit a residual mapping. Formally, suppose the desired mapping is H(x), the plain CNN directly fits the mapping H(x), while the ResNet fits another mapping of F(x) = H(x) − x. Therefore, the original mapping is recast into F(x) + x, and it could be realized by feed forward neural networks with “short connections”. The short connections simply perform identity mapping and their outputs areadded to the outputs of the stacked layers6

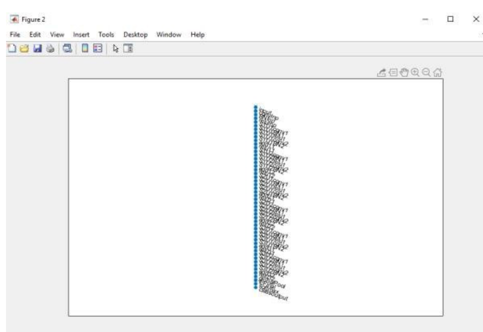
.Figure 3 shows the comparison between the building block of plain CNN and Res Net. By stacking the building blocks together, very deep residual networks (30-100 layers)could be developed without the degradation problem.

4. RESULT- For many applications, usinga network that consists of a simple sequence of layers is sufficient. However, some applications require networks with a more complex graph structure in which layers can have inputs from multiple layers and outputs to multiple layers. These types of networks are often called directed acyclic graph (DAG) networks.

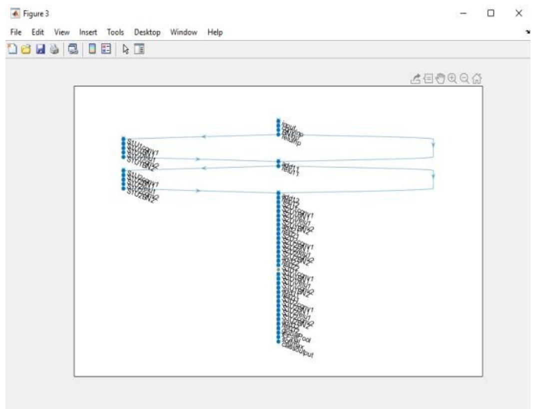




A residual network is a type of DAG network that has residual (or shortcut) connections that bypass the main network layers. Residual connections enable the parameter gradients to propagate more easily from the output layer to the earlier layers of the network, which makes it possible to train deeper networks. This increased network depth can result in higher accuracies on more difficult tasks.

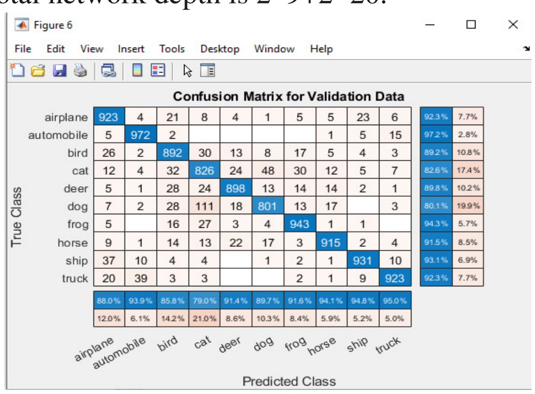


Add residual connections around the convolutional units. Most residual connections perform no operations and simply add element-wise to the outputs of the convolutional units.



When the layer activations in the convolutional units change size (that is, when they are down sampled spatially and up sampled in the channel dimension), the activations in the residual connections must also change size. Change the activation sizes in the residual connections by using a 1-by-1 convolutional layer together with its batch normalization layer.

Create a residual network with nine standard convolutional units (three units per stage) anda width of 16. The total network depth is 2*9+2=20.



5.CONCLUSION

In this project, weachieved the best classification accuracy of theCIFAR-10 datasetwith 56-layer deep residual network. Through this project, by experimenting with multiple linear/non-linear classifiers and tuning the hyper-parameters, wecould concludethat linear classifiers (OVA, Softmax, etc.) normally have a bottleneck accuracy around40%. Non-linear neural networks have a much better performance, with fully-connectedANN having accuracyhigher than 50%, and plain CNN around 80- 90%. ModifiedCNN, such as fractional maxpooling and ResNet will further increase the accuracy toabove 90%. Besides, we learned that the control of overfitting isimportant for classifiers withlarge set of parameters. By adding a penalty term for linear classifiers, or usingdrop-out technique, we could effectively control overfitting and improve the predictionaccuracy on the testing dataset.

6. REFERENCE

[1]Krizhevsky, Alex,Vinod Nair, and Geoffrey Hinton. "The CIFAR-10 dataset."online: http://www.cs.toronto.edu/kriz/cifar. html(2014).

[2]Lee, Chen-Yu, Patrick W. Gallagher, and Zhuowen Tu. "Generalizing pooling functions in convolutional neural networks: Mixed, gated, and tree." Artificial Intelligence and Statistics. 2016.

[3] Springenberg, Jost Tobias, et al. "Striving for simplicity: The all convolutional net." arXiv preprint arXiv:1412.6806 (2014).

[4] Krizhevsky, Alex, and G. Hinton. "Convolutional deep belief networks on cifar- 10." Unpublished manuscript 40 (2010): 7.

[5] Graham, Benjamin. "Fractional max- pooling." arXiv preprint arXiv:1412.6071 (2014).

[6] He, Kaiming, et al. "Deep residual learning for image recognition." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.

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

[8] jseppanen/cifarlasagne: Cifar-10 imageclassification with lasagne, https://github.com/jseppanen/cifar_lasagne. (Accessed on 04/24/2016).

[9] Recipes/deep residual learning cifar-10.pyat master lasagne/recipes, https://github.com/Lasagne/Recipes/blob/mast er/papers/deep_residual_learning/Deep_Resid ual_Learning_CIFAR -10.py.(Accessed on04/24/2016)

[10] R. Almahdi, R. Hardie, B. N. Narayanan, and M. S. de Silva, "Deep learning ensemble methods for skin lesion segmentation to aid in melanoma detection," 07 2019.

[11] M. A. Hearst, "Support vector machines," IEEE Intelligent Systems, vol. 13, no. 4, p. 18– 28, Jul. 1998. [Online]. Available:https://doi.org/10.1109/5254.708428

[12] Fu Jie Huang and Y. LeCun, "Large-scale learning with svm and convolutional for generic object categorization," in 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06), vol. 1, 2006, pp. 284–291.

[13] B. N. Narayanan, V. S. P. Davuluru, and R. Hardie, "Two-stage deep learning architecture for pneumonia detection and its diagnosis in chest radiographs," 02 2020.

[14] A. Jim´enez-S´anchez, S. Albarqouni, and D. Mateus, "Capsule networks against medicalimaging data challenges," in Intravascular Imaging and Computer Assisted Stenting and Large-Scale Annotation of Biomedical Data and Expert Label Synthesis, D. Stoyanov, Z. Taylor, S. Balocco, R. Sznitman, A. Martel, L. Maier-Hein, L. Duong, G. Zahnd, S. Demirci, S. Albarqouni, S.-L. Lee, S. Moriconi, V. Cheplygina, D. Mateus, E. Trucco, E. Granger, and P. Jannin, Eds. Cham: Springer International Publishing, 2018, pp. 150–160.

[15] J. Wanga, Z. Fangb, N. Langc, H. Yuanc, M.-Y. Sud, and P. Baldi, "A multi-resolution approach for spinal metastasis detection using deep siamese neural networks," Computers in Biology and Medicine, vol. 84, pp. 137–146, 2017.

[16] N. M. Khan, N. Abraham, M. Hon, and L.Guan, "Machine learning on biomedical images: Interactive learning, transfer learning, class imbalance, and beyond," in 2019 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR), 2019, pp. 85–90.

[17] D. S. Fong, L. Aiello, T. W. Gardner, G. L. King, G. Blankenship, J. D. Cavallerano, F.L. Ferris, and R. Klein, "Retinopathy in diabetes," Diabetes Care, vol. 27 Suppl 1, pp. S84–87, Jan 2004.

[18] Z. Zhong, L. Zheng, G. Kang, S. Li, and Y. Yang, "Random erasing data augmentation," in Proceedings of the AAAI Conference on Artificial Intelligence (AAAI), 2020.

[19] M. Aubry, S. Paris, S. W. Hasinoff, J. Kautz, and F. Durand, "Fast local laplacian filters," ACM Transactions on Graphics, vol. 33, no. 5, p. 1–14, 2014.

[20] S. Gao, M. Cheng, K. Zhao, X. Zhang, M.Yang, and P. H. S. Torr, "Res2net: A new multi-scale backbone architecture," IEEETransactions on Pattern Analysis and Machine Intelligence, pp. 1–1, 2019.

[21] J. Schmidhuber, "Deep learning in neuralnetworks: An overview," Neural Networks,vol. 61, pp. 85 – 117, 2015. [Online].Available: http://www.sciencedirect.com/science/article/pii/S0893608014002135

[22] F. F. Rosenblatt, "The perceptron: a probabilistic model for information storage and organization in the brain." Psychological review, vol. 65, pp. 386–408, 1958.

[23] J. Mairal, P. Koniusz, Z. Harchaoui, and C. Schmid, "Convolutional kernel networks," Advances in

neural information processing systems, 06 2014.

[24] M. Buscema, "Back propagation neural networks," Substance use & misuse, vol. 33, pp. 233–70, 02 1998.

[25] D. Kingma and J. Ba, "Adam: A method for stochastic optimization," International Conference on Learning Representations, 122014.