Neural Network Models for Object Recognition

Module Name: Machine Learning August 2023

Tutor Name: Dr Stelios Sotiriadis

Team Presentation: Group 1

Student Names: Ruth Allison, Nicholas Bandy, Tasweem Beelunkhan, Danilo De Sousa, Lojayne Diab

Introduction

Computer vision applications are in demand

Deep learning mimics how the brain processes information and allows for computer vision

The rationale for having a validation set

Training Set:
The sample used to fit the model

Test set:
Sample used to provide an unbiased evaluation of the final model fit (Brownlee, 2020).

Validation set:
The validation set plays a crucial role in preventing bias during model evaluation

Details of the data validation set training set

Data Set 60 000
32x32 color images

Training Set
75%

Validation Set
15%

Test Set
10%

Structure of the ANN

A CNN consists of two basic components:

- A base for feature learning, and
- A head for classification.

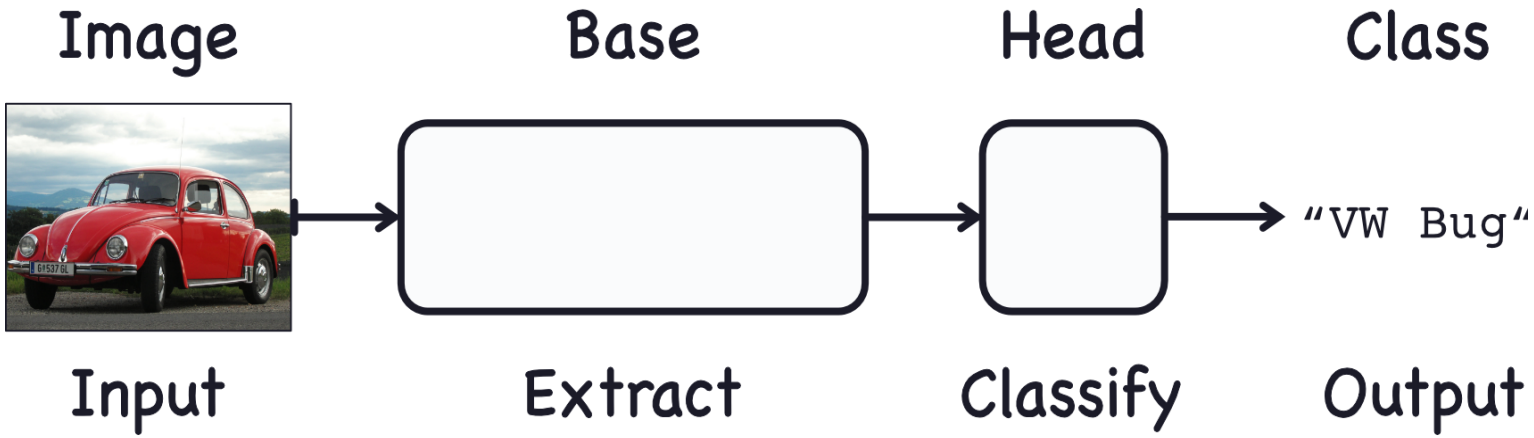

Figure 1: Basic CNN structure (Holbrook, 2019)

CNN Base

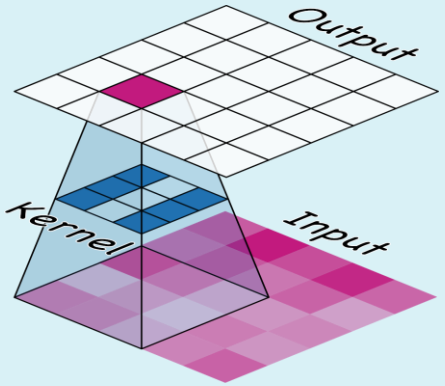

Figure 2: A kernel scanning the input for features (Holbrook, 2019)

- Filters, or kernels, act as lenses that scan the input image and give a weighted sum of the pixel values

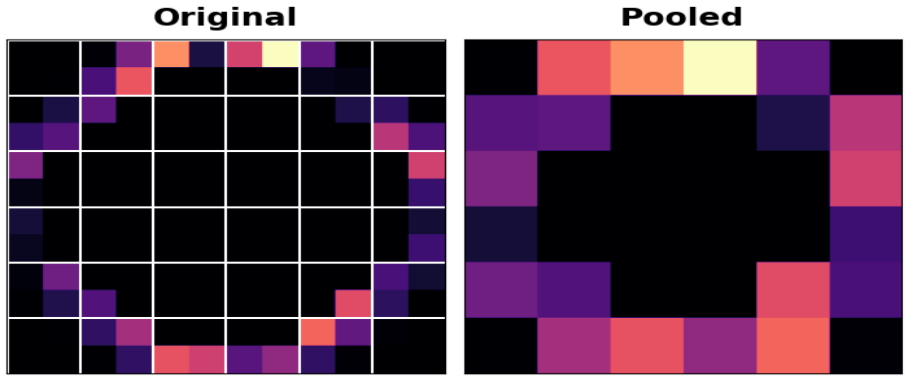

Figure 3: Reduction of dimensions by pooling (Holbrook, 2019)

- Pooling layers reduce the feature maps resolution.
- This reduces the sensitivity of the output to shifts and distortions

CNN Head

- In a fully connected layer, Neurons are connected to all neurons in the previous layer, and are one dimensional

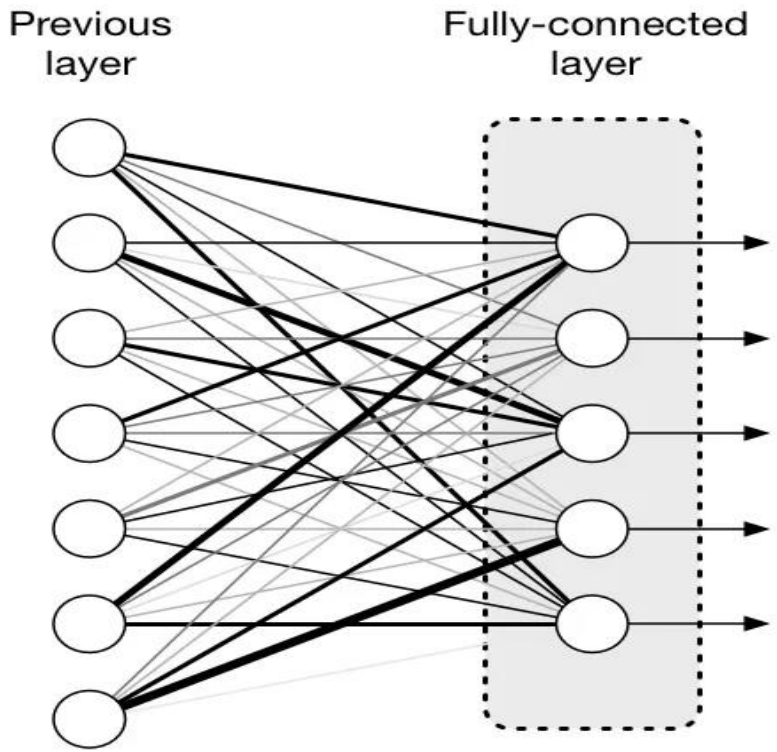

Figure 4: Connections of a fully connected layer (Teco, 2019)

Activation function

Activation functions used in the proposed model:

ReLU:

hidden layers

Softmax:
Output layer

Epochs

- Lower number of epochs vs large number of epochs (Ali at el, 2021).
- Epoch in one of the hyperparameter to be tune (Goel at el, 2021).
- Early stopping rule to avoid overfitting (Nurnoby at el, 2020).
- Proposed model was trained with 150 epochs, model reached a stable performance at the 45 epochs , with 87.38% accuracy on the testing dataset.

```
900/900 [==============================] - 11s 12ms/step - loss: 0.2581 - accuracy: 0.9128 - val_loss: 0.4421 - val_accuracy: 0.8666
Epoch 43/150
900/900 [==============================] - 11s 12ms/step - loss: 0.2503 - accuracy: 0.9157 - val_loss: 0.4368 - val_accuracy: 0.8608
Epoch 44/150
900/900 [==============================] - 11s 12ms/step - loss: 0.2503 - accuracy: 0.9150 - val_loss: 0.4264 - val_accuracy: 0.8738
```

Figure 5: Model training output

Loss function

- The categorical cross entropy loss function and sparse categorical cross entropy loss functions were used

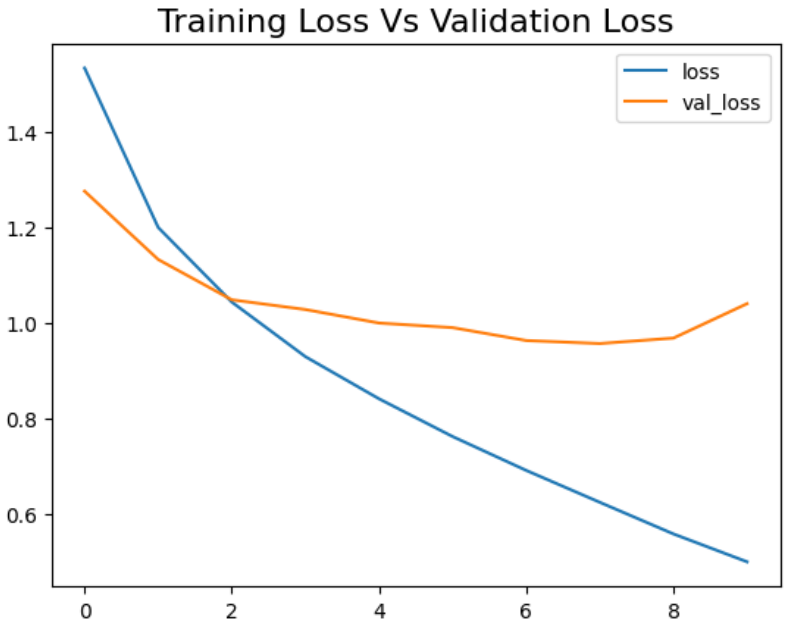

Figure 1: Initial set up

Figure 2: After adjustments

Neural network design – Choice of CNN

Convolutional Layer:
Filters extract localized features from an image.

Pooling Layer:
Down-samples data, emphasizing dominant patterns

Dense Layer:
Interprets features for final image classifications.

Neural network design – enhancing CNN efficiency

ReLU Activation:
Adds non-linearity, aiding in intricate pattern recognition.

Adam Optimizer:
Adapts learning rates for swift and optimal training.

Neural network design – comparison of models

Attribute	Model 1	Model 2
Input Shape	32×32×332×32×3	32×32×332×32×3
Convolutional Layers	2	4
Dropout Layers	None	4 (20%, 30%, 40%, 50%)
Batch Normalization	None	7 (After each convolutional layer)
Pooling Layers	2 max-pooling (2×22×2)	4 max-pooling (2×22×2)
Dense Layers	1 (256 neurons)	1 (128 neurons)
Output Neurons	10	10
Total Parameters	225,610	717,994

Conclusion

Initial model

Proposed model

Recommendations to improve the model

References

- Adam: A method for stochastic optimization - semantic scholar. Available at: https://www.semanticscholar.org/paper/Adam%3A-A-Method-for-Stochastic-Optimization-Kingma-Ba/a6cb366736791bcccc5c8639de5a8f9636bf87e8 (Accessed: 21 October 2023).
- Akbar, S., Peikari, M., Salama, S., Nofech-Mozes, S. & Martel, A. (2017), Transitioning between Convolutional and Fully Connected Layers in Neural Networks, Cornell University Library, Available from: https://arxiv.org/pdf/1707.05743.pdf
- Albawi, S., Mohammed, T.A. and Al-Zawi, S. (2017) 'Understanding of a convolutional neural network', 2017 International Conference on Engineering and Technology (ICET) [Preprint]. doi:10.1109/icengtechnol.2017.8308186.
- Ali, L., Alnajjar, F., Jassmi, H., Gocho, M., Khan, W. & Serhani, M. A. (2021) Performance Evaluation of Deep CNN-Based Crack Detection and Localization Techniques for Concrete Structures. *Sensors* 21(5):1688. DOI: https://doi.org/10.3390/s21051688 .
- Aloysius, N., Geetha, M. (2017) "A review on deep convolutional neural networks," 2017 International Conference on Communication and Signal Processing (ICCSP), Chennai, India, , pp. 0588-0592, doi: 10.1109/ICCSP.2017.8286426.
- Alzubaidi, L. et al. (2021) 'Review of Deep Learning: Concepts, CNN Architectures, challenges, applications, Future Directions', Journal of Big Data, 8(1). doi:10.1186/s40537-021-00444-8.
- Baeldung. (n.d.) What is a learning curve in Machine Learning? Available from: https://www.baeldung.com/cs/learning-curve-ml#:~:text=One%20of%20the%20most%20widely,the%20model%20fits%20new%20data. [Accessed 12 October 2023]
- Brownlee, J. (2018). Cross-validation: A Simple Explanation.https://machinelearningmastery.com/k-fold-cross-validation/ [Accessed 13 October 2023]
- Brownlee, J. (2020) What is the difference between test and validation datasets. Available from: https://machinelearningmastery.com/difference-test-validation-datasets/#:~:text=Validation%20Dataset%3A%20The%20sample%20of,incorporated%20into%20the%20model%20configuration. [Accessed 4 October 2023]
- Brownlee, J. (2021). Choosing the Right Cross-Validation Splitting Strategy. Available from: https://machinelearningmastery.com/k-fold-cross-validation/ [Accessed 13 October 2023]

References

- Convolutional neural network for CIFAR-10 Dataset Image Classification. Available at: https://www.researchgate.net/publication/337240963_Convolutional_Neural_Network_for_CIFAR-10_Dataset_Image_Classification (Accessed: 21 October 2023).
- Goel, T., Murugan, R. & Mirjalili, S.(2021) OptCoNet: an optimized convolutional neural network for an automatic diagnosis of COVID-19. *Applied Intelligence* 51(1):1351–1366. DOI: https://doi.org/10.1007/s10489-020-01904-z .
- Goodfellow, I., Bengio, Y. and Courville, A. (2017) Deep learning. Cambridge, MA: MIT Press.
- Holbrook, R. & Cook, A . (2019) Convolution and ReLU. Available from: https://www.kaggle.com/code/ryanholbrook/convolution-and-relu/tutorial
- Holbrook, R. & Cook, A. (2019) The Convolutional Classifier. Available From: https://www.kaggle.com/code/ryanholbrook/the-convolutional-classifier
- Holbrook, R. & Cook, A. (2019) Maximum Pooling. Available From: https://www.kaggle.com/code/ryanholbrook/maximum-pooling
- Howard, R. (2017). How (and why) to create a good validation set. Available from: https://www.fast.ai/2017/11/13/validation-sets/ [Accessed 13 October 2023]
- Jain, V. (2019) Everything you need to know about “Activation Functions” in Deep learning models. Available from: https://towardsdatascience.com/everything-you-need-to-know-about-activation-functions-in-deep-learning-models-84ba9f82c253
- LeCun, Y. and Bengio, Y., (1995). Convolutional networks for images, speech, and time series. The handbook of brain theory and neural networks, 3361(10), Available from: https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=e26cc4a1c717653f323715d751c8dea7461aa105
- Nurnoby, M. F., El-Alfy, E.-S. M. & Luqman, H. (2020) Evaluation of CNN Models with Transfer Learning for Recognition of Sign Language Alphabets with Complex Background. *2020 International Conference on Innovation and Intelligence for Informatics, Computing and Technologies (3ICT)* 1(1):1-6. DOI:https://doi.org/10.1109/3ICT51146.2020.9311989

References

- Obla, S., Gong, X., Aloufi, A., Hu, P. and Takabi, D. (2020) "Effective Activation Functions for Homomorphic Evaluation of Deep Neural Networks," in IEEE Access, vol. 8, pp. 153098-153112, doi: 10.1109/ACCESS.2020.3017436.
- Pramoditha, R. (2022) Why do we need a validation set in addition to training and test sets? Available from: https://towardsdatascience.com/why-do-we-need-a-validation-set-in-addition-to-training-and-test-sets-5cf4a65550e0 [Accessed 4 October 2023]
- Sharma, E. (2020) Simple Cifar10 CNN Keras code with 88% accuracy. Available from: https://www.kaggle.com/code/ektasharma/simple-cifar10-cnn-keras-code-with-88-accuracy/notebook
- Soon, F. C., Khaw, H. Y., Chuah, J. H. & Kanesan, J. (2018) Hyper-parameters optimisation of deep CNN architecture for vehicle logo recognition. IET Intelligent Transport Systems 12(8):939-946. DOI:https://doi.org/10.1049/iet-its.2018.5127 .
- Stanford University (n.d.) Softmax Regression. Available from http://deeplearning.stanford.edu/tutorial/supervised/SoftmaxRegression/#:~:text=Softmax%20regression%20(or%20multinomial%20logistic,kinds%20of%20hand%2Dwritten%20digits
- Tan, P.N., Steinbach, M., & Kumar, V. (2006). Introduction to Data Mining. In "Introduction to Data Mining" (Chapter 5. https://wwwusers.cse.umn.edu/~kumar001/dmbook/ch5_association_analysis.pdf
- Teco, K. (2019) Fully-Connected Layer With Dynamic Input Shape, Medium. Available From: https://medium.com/@tecokids.monastir/fully-connected-layer-with-dynamic-input-shape-70c869ae71af
- TensorFlow (n.d.) tf.keras.losses.CategoricalCrossentropy. Available from: https://www.tensorflow.org/api_docs/python/tf/keras/losses/CategoricalCrossentropy [Accessed 7 October 2023]
- Voulodimos, A., Doulamis, N., Doulamis, A., Protopapadakis, E. (2018). "Deep Learning for Computer Vision: A Brief Review", Computational Intelligence and Neuroscience. Available From: https://doi.org/10.1155/2018/7068349
- Yathish, V. (2022) Loss functions and their use in neural networks. Available from: https://towardsdatascience.com/loss-functions-and-their-use-in-neural-networks-a470e703f1e9 [Accessed 12 October 2023]