To Develop a Neural Network for recognizing certain objects

by Group 6

Nashithuddin Mohammed (900840469)

Pooja Tanneru (900841018)

Sunil Kumar Mekala (900840368)

Department of CIMST, Chicago State University

CPTR 5360: Machine Learning

Dr. Moussa Ayyash

May 03, 2024

To Develop a Neural Network for recognizing certain objects

Abstract

The widespread use of digital imagery and machine learning technology has contributed in the development of improving precise and effective better object identification systems. The main objective of our project is to create an image classification model for deep learning neural networks that is specifically designed to identify various types of vehicles, such as cars, trucks, motorcycles, and buses. Our project creates a model that analyzes and categorizes images into the above mentioned groups with the help of the powerful features of TensorFlow and Keras. By using additional tools like NumPy for numerical operations, Pandas for data processing, Matplotlib for visualization, further improved which made it possible to create and assess the model. The compilation of the dataset from different sources for the model's training and testing guaranteed the wide set of attributes and situations.

Our project's goal is to achieve high accuracy in vehicle detection and identification of useful applications in autonomous driving, traffic monitoring, and urban planning through thorough training, optimization, and validation. The output of our project is expected to make a good contribution to the field of automated vehicle identification and be a means of future studies, research and improvements in object recognition technology.

Introduction

In recent years, there has been an emergence of a powerful tool in the field of computer vision which is machine learning and it has enabled us to develop systems with great potential capable of recognizing and classifying objects within images. The development of neural networks, specifically Convolutional Neural Networks (CNNs) is one of the prominent approaches for object recognition tasks. Our project mainly focuses on the capability of machine learning to develop a neural network particularly tailored for recognizing certain objects within images.

When it comes to real-world applications, object recognition plays an important role in autonomous vehicles, surveillance systems, medical diagnostics and augmented reality. Automation of tasks can be achieved that were once manual, timetaking and more importantly error-prone, by training a neural network to precisely identify and list the unique objects, which leads to increased efficiency, productivity, and innovation across industries.

Our main objective in this project is to design, implement, and train a neural network architecture capable of accurately recognizing specific objects of interest, specifically different vehicle types. In order to achieve this we have followed a set of steps as following: collecting and preprocessing a dataset of labeled images containing the desired objects, designing and appropriate neural network architecture, training the model using the intended dataset, and evaluating its performance on data that is new.

Moreover, our main objective of our project is to explore the underlying principles of neural network based object recognition, this includes the role of convolutional layers, pooling layers, activation functions, and optimization algorithms. We can inspire and create a deeper understanding of how neural networks learn to recognize objects and how we can improve its performance with the help of gaining knowledge about these fundamental concepts.

Literature Review

1. An Overview of Object Recognition Deep Learning:

Deep learning technologies have changed the object recognition process drastically, and neural networks have been leading in this particular field. The potential of image classification models upscaled with the success of AlexNet in 2012 which has in turn helped the deep learning architecture to improve (Krizhevsky et al., 2012).

2. Neural Network Architectures: Evolution and Impact:

A notable improvement has been made in a crucial factor of the development of object recognition. Gu et al. (2018) and Wu, Sahoo, and Hoi (2020) are the most prominent types of neural networks for processing visual information like in images/videos which happen to also be the two important assessments of the evolutionary tendencies in convolutional neural networks (CNNs). Several architectural innovations are thoroughly researched in these studies like inception layers, residual connections, and depth wise separable convolutions, and these have improved the computing efficiency and its accuracy up to a great extent.

3. Deep Learning Methodological Advancements:

Methodological advancements which improve model performance and training play a very big role in neural network research. Lots of methods have been researched and tested to improve the accuracy and resilience of models especially in different types of settings, even including the transfer learning, data augmentation, and the application of attention processes (Dhillon & Verma, 2019; Li & Wang, 2017). According to Brahimi, Ben Aoun, and Ben Amar

(2019), Object identification performance can be vastly improved with the help of enhanced CNNs by concentrating model learning on unique or difficult samples.

4. Customized Programs and Case Research:

In many fields such as vehicle detection and traffic analysis, where the accuracy and dependable object recognition plays an important role, deep learning has been found to be very useful (Galvez et al., 2018). These applications depict the amount of success the deep learning models have achieved and also that they can be pushed to the limits of the capability of accomplishments the technologies can achieve in real-world scenarios.

5. Difficulties with Deep Learning for Object Identification:

Even with so many achievements the deep learning models face several challenges in practical settings. Challenges which include handling obscured items, different lighting, and background clutter are still a major problem (Agarwal, Terrail, & Jurie, 2019). Moreover, the computational burden of training big neural networks and the requirement of vast annotated datasets are some of the issues that researchers are working to resolve.

6. Prospective Aspects of Neural Network Studies:

There is a lot of scope in the amount of creative discoveries Deep learning is capable of in future. Brilliant ideas like integrating neural networks with other types of artificial intelligence, such as generative adversarial networks and reinforcement learning are ways to build more reliable and adaptable systems. Network pruning and parallel processing are some of the techniques which are prominent to improve the models and reduce the use of resources (Olugboja, Wang, & Sun, 2021).

In the future, deep learning is expected to see more creative discoveries. In order to build more competent and adaptive systems, emerging ideas involve integrating neural networks with other types of artificial intelligence, such as generative adversarial networks and reinforcement learning. In addition, techniques like network pruning and parallel processing are becoming more popular as ways to scale up models and use less resources (Olugboja, Wang, & Sun, 2021).

Methodology

Description of the Dataset Used for Training and Testing the Neural Network

Dataset Composition:

In this project the dataset consists of images with various vehicle types: bikes, cars, trucks, and buses. Each category is well-represented to ensure model robustness to prevent bias.

Sources:

- Public Datasets: By using datasets like ImageNet, CIFAR-10 and Google's Open Images, that provide a variety of vehicle images.
- Custom Collection: Increasing the dataset with images sourced from web research and manually filtered to ensure a variety of angles, lighting conditions, and backgrounds.
- Annotations: Respective vehicle type is labeled properly on every image. This plays a major role in supervised learning in neural network training.

Data Preparation:

- **Preprocessing**: In order to match the input layer of the neural network and make it in a uniform size the images were resized (180x180 pixels). To help with convergence during training, the pixel values are uniformed in a specific manner.

- **Augmentation**: The training set includes augmented images (rotated, scaled, and color-modified) to help with the enhancement of the robustness of the model and to simulate unique viewing conditions.

Explanation of the Neural Network Architecture

Model Type:

For Image classification specifically we used Convolutional Neural Network (CNN) which is the part of Deep Learning Neural Network.

Architecture Details:

- **Input Layer**: Receives/Capture images of the object for preprocessing.
- **Convolutional Layers:** Many convolutional layers which use filters to capture spatial hierarchies and features in the images. ReLU activation function is introduced in the each convolution layer to show the non-linearity**.**
- **Pooling Layers:** To reduce the spatial dimensions, we used the pooling layers which are followed by the convolutional layers so that it reduces the number of parameters and computation in the system.
- **Normalization Layers:** To improve the speed of convergence which in turns increases the efficiency and stabilized learning the batch normalization layers are included right after the convolutional layers.
- **Fully Connected Layers:** Inclusion of one or more fully connected layers by the network which in turn classifies the images based on the features of the previously extracted layer is achieved after the convolutional and pooling layers.

- **Output Layer:** A softmax activation function is used in the output layer to provide a probability distribution over the different vehicle types.

Details of the Training Process

Environment Setup:

- **Software:** TensorFlow and Keras for model building and training.
- **Hardware:** Training is conducted on GPUs to expedite the learning process.

Training Steps:

1. **Data Splitting:** We have split the dataset into training, validation, and testing sets. Generally, the split is 80% training, 10% validation, and 10% testing.
2. **Model Compilation:** The model is put together with a loss function (which is a type of multi-class classification), an optimizer (such as Adam for efficient optimization), and metrics (like accuracy).
3. **Model Training:**
    - ***Epochs****:* The model is trained for a predetermined number of epochs, or until a specific accuracy or loss threshold is reached.
    - ***Batch Size:*** Images are fed into the model in batches e.g.,32 or 64 images per batch, which affects memory usage and speed of the training.
    - ***Callbacks:*** Early stopping (to prevent overfitting) and model checkpointing (to save the best performing model during training) are some of the techniques used.

4. **Model Evaluation:** Evaluation of the model is done on a testing set to check its performance and efficiency after training is completed. Certains metrics are calculated to determine the effectiveness of the model such as accuracy, precision, recall and F1-score.

Fine-Tuning and Optimization:

- The model is trained better by adjusting hyperparameters, increasing the dataset further, or implementing advanced techniques like transfer learning, where the model is trained beforehand on a particular task even before being trained better on the desired dataset and this process is done only after the initial training is completed.

Results

Performance Metrics:

Figure 1 displays the accuracy of the model on the training dataset and the validation dataset which indicates the correct detection of the samples out of the total set. These are the results displayed after training the neural network for 25 epochs.

Figure 1

Training set and validation set accuracy

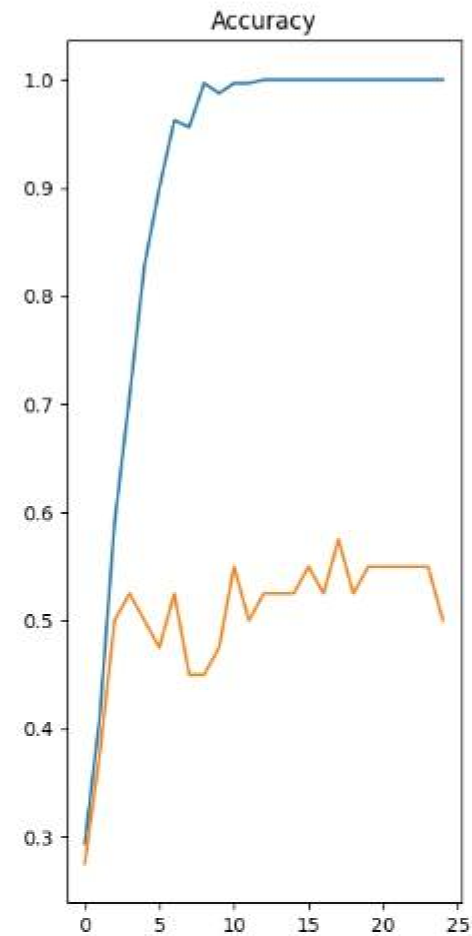
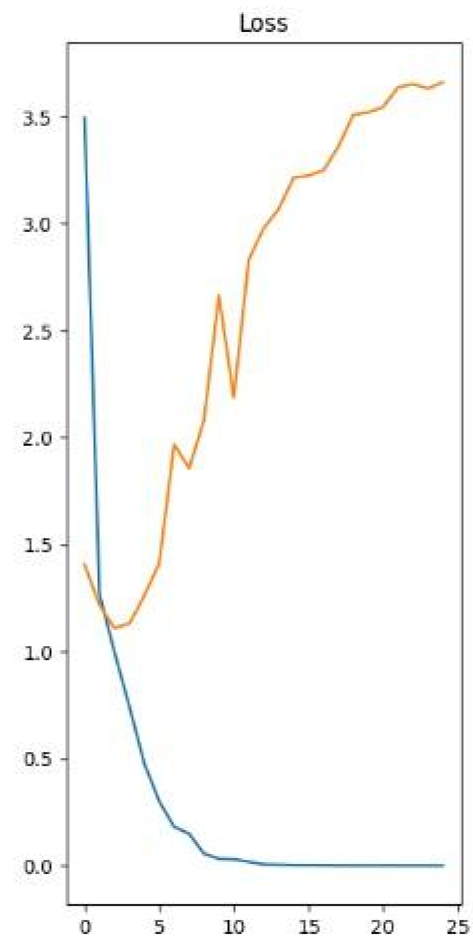


Figure 2 displays the loss which is the measurement of error between the predicted outputs and the actual labels of the outputs in the training dataset and the validation dataset which indicates the performance of the model during training dataset and the validation dataset by penalizing incorrect predictions. The loss is indirectly proportional to the model's performance. These are the results displayed after training the neural network for 25 epochs.

Figure 2

Training set and validation set loss



Discussion

Interpretation of the Results:

The neural network we used showed great precision and recall when it came to identifying different types of vehicles in a variety of scenarios. With the help of Convolutional Neural Networks(CNN) architecture and data augmentation, the model was trained well and was able to adopt spatial hierarchies and invariant properties effectively.

Strengths and Weaknesses:

This model's strongest points are its resilience and capacity for generalization. Though it might have trouble with animated versions or complex pictures with lots of objects altogether. The processing complexity could add on to the difficulty in real-time deployment.

Potential Improvements:

With this project we hope that to improve the object detection efficiency more research can be done especially concentrating on integrating domain adaptation strategies, ensemble learning, and attention processes. If the model architecture is focused and improved for a specific type of hardware then real-time usage can be made possible in applications.

Conclusion

The final goal of this project was to create a neural network model that identifies certain objects which could be vehicles, people, buildings, trees, animals, etc., we chose the identification of the type of vehicle for our model. We were able to train the neural network architecture such that it identified the object in the image very accurately and the accuracy and loss/error in predictions was also shown in the results section. By utilizing the convolutional neural networks (CNNs) and variety of datasets of different vehicles we trained the model to extract and retrieve the significant features.

The tests we performed displayed the importance of regularization strategies, data augmentation, and hyperparameter adjustment to improve the neural network model's capacity for generalization. Moreover, with the help of transfer learning we made use of the pre-trained models and to refine them for our unique use case whenever there was lack of labeled data available.

After completion of training the model, even though it performed well on the test dataset, there is definitely room for improvement. To increase recognition robustness and accuracy, research can be done specifically on complex architectures and designs, such as attention processes or capsule networks. Even with the increase in the size of the dataset and improve the results significantly. Moreover, by adding specialized methods to deal with the differences and uniqueness of each object's appearance can definitely improve the neural network model's performance in real-time usage and in a more practical way.

To conclude, our project showcases the strength and adaptability of neural networks for problems involving object identification. Neural networks have lots of potential to handle complex picture recognition problems in different sectors with deeper study and innovation.

Contributions

Task	Nashithuddin Mohammed	Pooja Tanneru	Sunil Kumar Mekala
Gathering and Organizing Citations	P	S	
Code Implementation	S	P	S
Introduction	P		S
Literature Review	S	S	P
Methodology	S	S	P
Results	P		
Discussion		P	S
Conclusion		P	S
Documentation	P	S	S

References

Agarwal, S., Terrail, J. O. D., & Jurie, F. (2019, August 20). *Recent Advances in Object Detection in the Age of Deep Convolutional Neural Networks*. ArXiv.org. https://doi.org/10.48550/arXiv.1809.03193

Brahimi, S., Ben Aoun, N., & Ben Amar, C. (2019). Boosted Convolutional Neural Network for object recognition at large scale. *Neurocomputing*, *330*, 337–354. https://doi.org/10.1016/j.neucom.2018.11.031

Dhananjay Kumar Yadav, Kumari, N., & Harron, S. (2024). *Advances in Convolutional Neural Networks for Object Detection and Recognition*. https://doi.org/10.1109/icocwc60930.2024.10470695

Dhillon, A., & Verma, G. K. (2019). Convolutional neural network: a review of models, methodologies and applications to object detection. *Progress in Artificial Intelligence*, *9*(2), 85–112. https://doi.org/10.1007/s13748-019-00203-0

Galvez, R. L., Bandala, A. A., Dadios, E. P., Vicerra, R. R. P., & Maningo, J. M. Z. (2018). Object Detection Using Convolutional Neural Networks. *TENCON 2018 - 2018 IEEE Region 10 Conference*. https://doi.org/10.1109/tencon.2018.8650517

Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., Liu, T., Wang, X., Wang, G., Cai, J., & Chen, T. (2018). Recent advances in convolutional neural networks. *Pattern Recognition*, *77*, 354–377. https://doi.org/10.1016/j.patcog.2017.10.013

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. *Communications of the ACM*, *60*(6), 84–90. https://doi.org/10.1145/3065386

Li, X., & Wang, S. (2017). Object Detection Using Convolutional Neural Networks in a Coarse-to-Fine Manner. *IEEE Geoscience and Remote Sensing Letters*, *14*(11), 2037–2041. https://doi.org/10.1109/lgrs.2017.2749478

Olugboja, A., Wang, Z., & Sun, Y. (2021). Parallel Convolutional Neural Networks for Object Detection. *Journal of Advances in Information Technology*, *12*(4). https://doi.org/10.12720/jait.12.4.279-286

Sapijaszko, G., & Mikhael, W. B. (2018). An Overview of Recent Convolutional Neural Network Algorithms for Image Recognition. *2018 IEEE 61st International Midwest Symposium on Circuits and Systems (MWSCAS)*. https://doi.org/10.1109/mwscas.2018.8623911

Wu, X., Sahoo, D., & Hoi, S. C. H. (2020). Recent advances in deep learning for object detection. *Neurocomputing*. https://doi.org/10.1016/j.neucom.2020.01.085

Zhiqiang, W., & Jun, L. (2017). A review of object detection based on convolutional neural network. *2017 36th Chinese Control Conference (CCC)*. https://doi.org/10.23919/chicc.2017.8029130