Car image recognition using convolutional neural network with efficient met architecture

Serhiy Balovsyak*1,†*, Olga Kroitor*1,†*, Khrystyna Odaiska*1,†*, Abdel-Badeeh M. Salem*2,†* and Serhii Stets*1,∗,†*

1 Yuriy Fedkovych Chernivtsi National University, Kotsiubynsky 2, 58012, Chernivtsi, Ukraine
2 Ain Shams University, El-Khalyfa El-Mamoun Street Abbasya, Cairo, Egypt

Abstract

A convolutional neural network for car image recognition has been developed. The neural network model is implemented with the EfficientNet architecture. The software implementation of the neural network was written in Python using the Keras library. The neural network model was trained on the basis of the publicly available Vehicle Detection Image Set dataset, which made it possible to compare the accuracy of the developed model with analogues. The structure of the developed model with the basic architecture of EfficientNet was improved, which increased the accuracy of model training. The improvement consisted of adjusting the network parameters, unfreezing some of its layers, and creating stop conditions to avoid overtraining. These additional measures helped to improve the model's accuracy. The images used to train the model were pre-processed by scaling and randomly rotating them (augmentation). Due to the model improvement and data preprocessing, a high model training accuracy (99.98%) was obtained, which exceeds the training accuracy for the best analog model (99.63%) on the used dataset. The developed convolutional neural network can be used for car image recognition and localization.

1. Introduction

In today's world, there is a need to develop Artificial Neural Networks (ANN) designed to recognize digital images. In particular, in practice, the task of recognizing car images often arises. This task is quite effectively solved by Convolutional Neural Networks (CNN) [1-3], which use deep learning technologies. CNN are particularly effective in image recognition, as they take into account the geometry of images, the peculiarities of human visual perception, and use multilayer signal processing for a large number of examples. The name