PD.TVZ.HR
POLYTECHNIC & DESIGN Vol. 8, No. 1, 2020.
DOI: 10.19279/TVZ.PD.2020-8-1-18

AUTOMOBILE CLASSIFICATION USING TRANSFER LEARNING ON RESNET NEURAL NETWORK ARCHITECTURE

KLASIFIKACIJA AUTOMOBILA KORISTEĆI TRANSFERIRANO UČENJE NA RESNET ARHITEKTURI NEURONSKE MREŽE

Stjepan Ložnjak[1], Tin Kramberger[2], Ivan Cesar[2], Renata Kramberger[2]

[1]Tehničko veleučilište u Zagrebu, Vrbik 8, Zagreb, Hrvatska, Student
[2]Tehničko veleučilište u Zagrebu, Vrbik 8, Zagreb, Hrvatska

ABSTRACT

Classification is one of the most common problems that neural networks are used for. In the case of higher resolution image classification, convolutional neural networks are commonly used. Due to the reason that convolutional neural networks are so often used in classification, there are many pretrained models that can be adapted for new domains using a technique called transfer learning. This paper shows how excellent results in classification accuracy can be achieved by applying transfer learning to pretrained convolution neural network. This paper presents the results of the learning transfer of the ResNet-152 convolution neural network on the Stanford Cars dataset. The results show accuracy over 88% only by training the last fully connected layer.

Keywords: *Transfer learning, ResNet, Stanford Car dataset*

SAŽETAK

Klasifikacija je jedan od najčešćih problema za koji se koriste neuronske mreže. U slučaju klasifikacije slika veće razlučivosti, najčešće se koriste konvolucijske neuronske mreže. Iz razloga što se konvolucijske neuronske mreže izuzetno često koriste za klasifikaciju postoji mnogo inačica unaprijed istreniranih mreža koje je moguće uz određene preinake upotrijebiti i dotrenirati za rješavanje novih domena koristeći tehniku transferiranog učenja. Ovaj rad prikazuje način na koji je moguće dobiti izvrsne rezultate s obzirom na točnost klasifikacije pomoću upotrebe transferiranog učenja i unaprijed trenirane konvolucijske neuronske mreže. Rad prezentira rezultate transferiranog učenja ResNet-152 konvolucijske neuronske mreže na Stanford Car setu podataka. Rezultati pokazuju točnost od preko 88% točnosti samo pomoću treniranja zadnjeg potpuno povezanog sloja neuronske mreže.

Ključne riječi: *Transfer learning, ResNet, Stanford Car dataset*

1. INTRODUCTION
1. UVOD

In recent years with the development of better and faster GPUs that can cope with numerous calculations the field of deep learning emerged. Better hardware yielded better and deeper neural networks which easily exceed 100 layers of depth which have led to a series of breakthroughs for image classification [1]–[4]. These kinds of deep convolutional neural networks are very complex and need a vast amount of training data and require long training time on conventional computers with high end standard GPUs. Due to the long training time and availability of pretrained ResNet models on the COCO dataset [5], it was decided that the best approach is to incorporate the transfer learning technique to the existing pretrained neural network. To classify cars, the network must be trained on images of cars that provide information on car classification. Currently, the best dataset for car classification is the Stanford Car dataset [6]. It consists of 197 car models on 16.185 images covering sedans, SUVs, coupes, convertibles, pickups, hatchbacks and station wagons.

2. RESNET-152 NEURAL NETWORK

2. RESNET-152 NEURONSKA MREŽA

ResNet neural network [4] was specifically developed for ImageNet Large Scale Visual Recognition Challenge 2015 [7] which it won that year. The network, depending on its purpose, can consist of 18 to several hundreds of layers. Such traditional deep convolutional neural network can expect many problems during training and testing one of which is degradation. With an increase of network depth, accuracy gets saturated after which it rapidly degrades which is unexpectedly not caused by overfitting, and adding more layers leads to higher training error [8], [9]. To overcome this problem, the ResNet neural network uses residual connectors between layers. Residual connectors are basically short connections between several layers which turn the neural network into a residual neural network which can be seen in Figure 1.

3. STANFORD CAR DATASET

3. STANFORD CAR SET PODATAKA

As was mentioned in the introduction, the Stanford car dataset consists of 16.185 car images in 197 categories of cars. The dataset is devised in two parts: training part which consists of 8.144 images used for training the neural network and testing part which consists of 8.041 images that are used for testing the neural network. Of course, the dataset provides the data which image belongs to which class in the form of "CSV" data files which can be seen in Figure 2. The files "anno_test.csv" and "anno_train.csv" consist of 6 columns. The first column is the name of the image, from the second to the fifth column are four values of pixels that determine the exact position of the car in the image. The sixth column contains information that classifies the car into classes. The list of all classes is in the "classes.csv" file. The data from the "anno_train.csv" table was used to obtain training images that were adjusted in such a way that the background was as small as possible and that the car was in focus which can be seen in Figure 3.

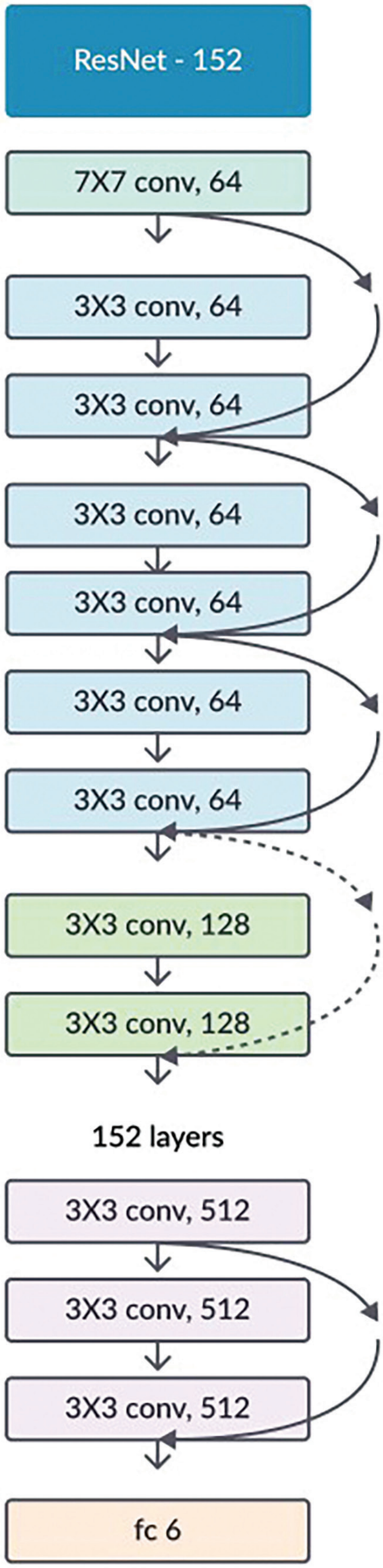
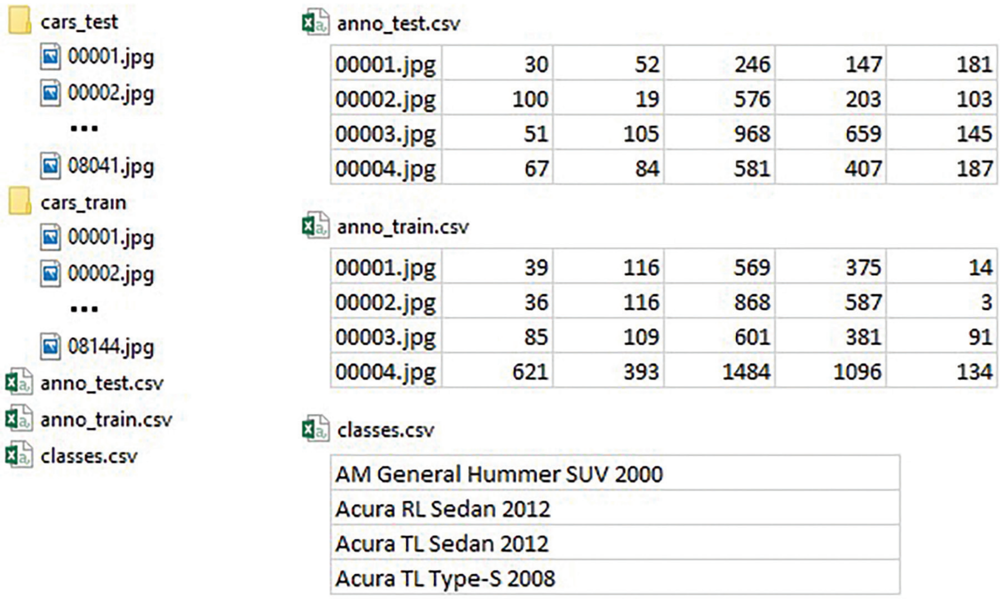


Figure 1 *ResNet-152 neural network architecture*

Slika 1 *Arhitektura ResNet-152 neuronske mreže*



00001.jpg	30	52	246	147	181
00002.jpg	100	19	576	203	103
00003.jpg	51	105	968	659	145
00004.jpg	67	84	581	407	187

00001.jpg	39	116	569	375	14
00002.jpg	36	116	868	587	3
00003.jpg	85	109	601	381	91
00004.jpg	621	393	1484	1096	134

AM General Hummer SUV 2000
Acura RL Sedan 2012
Acura TL Sedan 2012
Acura TL Type-S 2008

Figure 2 *Dataset and csv files structures*

Slika 2 *Struktura seta podataka i csv datoteka*

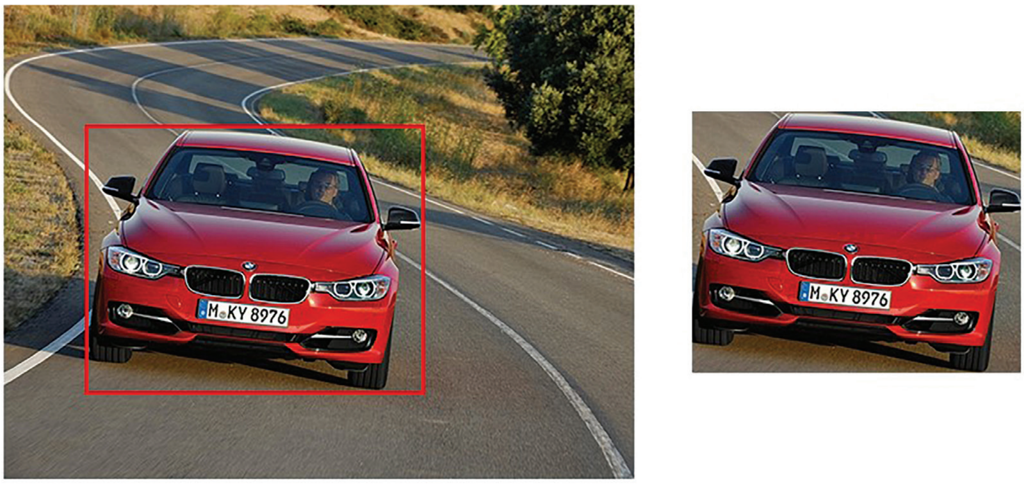


Figure 3 *The image with the correct position of the automobile (left) and cropped image (right)*

Slika 3 *Slika s točnom pozicijom automobila (lijevo) i izrezana slika (desno)*

4. COMPUTER METHODS AND EXPERIMENTAL SETUP

4. RAČUNALNE METODE I EKSPERIMENTALNA POSTAVA

In the experimental part, we want to demonstrate the benefits of the transfer learning technique and how it is beneficial on a larger neural network model that was pretrained on a different dataset. The ResNet-152 model was trained using a transfer learning method to retrain only the last fully connected layer. Finally, we will present the results of training the neural network in such a manner.

The neural network was trained using the training set of images of size 224x224 and three channels (RGB). The last fully connected layer was set to an output vector of size 196 using softmax activation function [10] which is the most common in classification [11]. The ResNet-152 model was trained on a GTX 1070 8GB GPU. For the software platform CUDA 10.0, Tensorflow 1.15.0 [12] and Keras 2.3.0 [13] were used as frameworks on which the training was performed. Augmentations were made to the training dataset in order to achieve better results. All augmentations are described in Table 1.

Tablica 1. *Dataset augmentations parameters*

Table 1. *Parametri povećavanja seta podataka*

Rotation range	20
Width shift	0.1
Height shift	0.1
Zoom range	0.2
Horizontal flip	True

5. RESULTS
5. REZULTATI

In our experiment, the training took approximately 1 hour and 8 minutes to complete the 10 epochs of training the fully connected layer with the batch size of 16 images per batch. The training was repeated but the neural network was trained on 20 epochs instead which took 2 hours and 16 minutes.

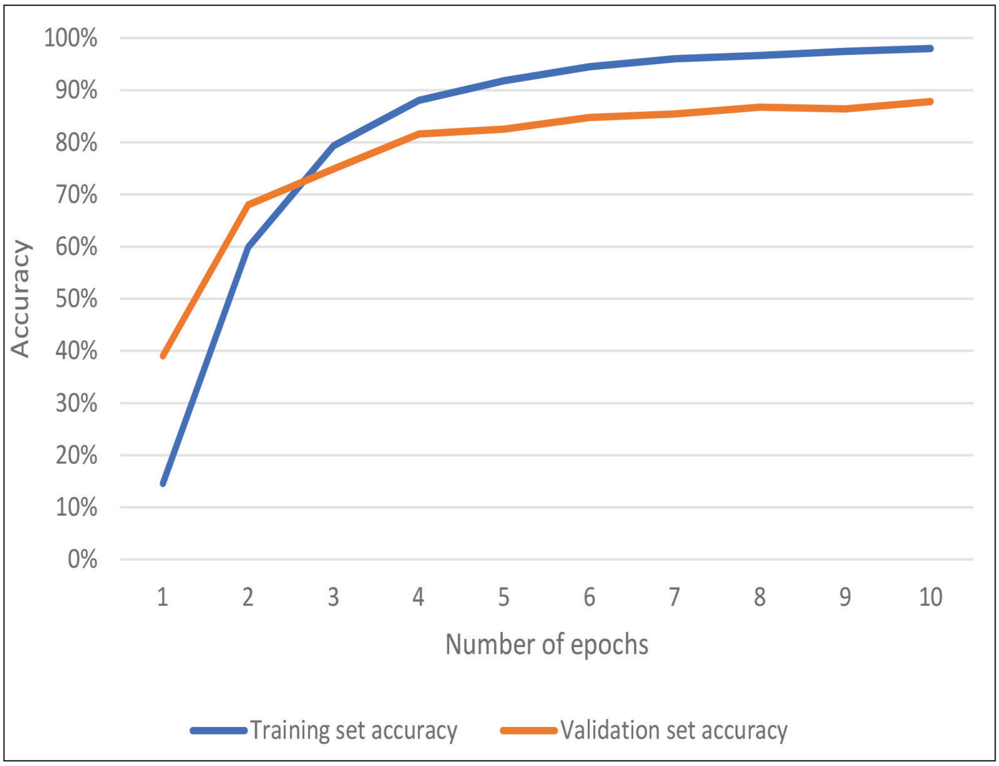


Figure 4 *Results of training after 10 epochs*

Slika 4 *Rezultati treniranja nakon 10 epoha*

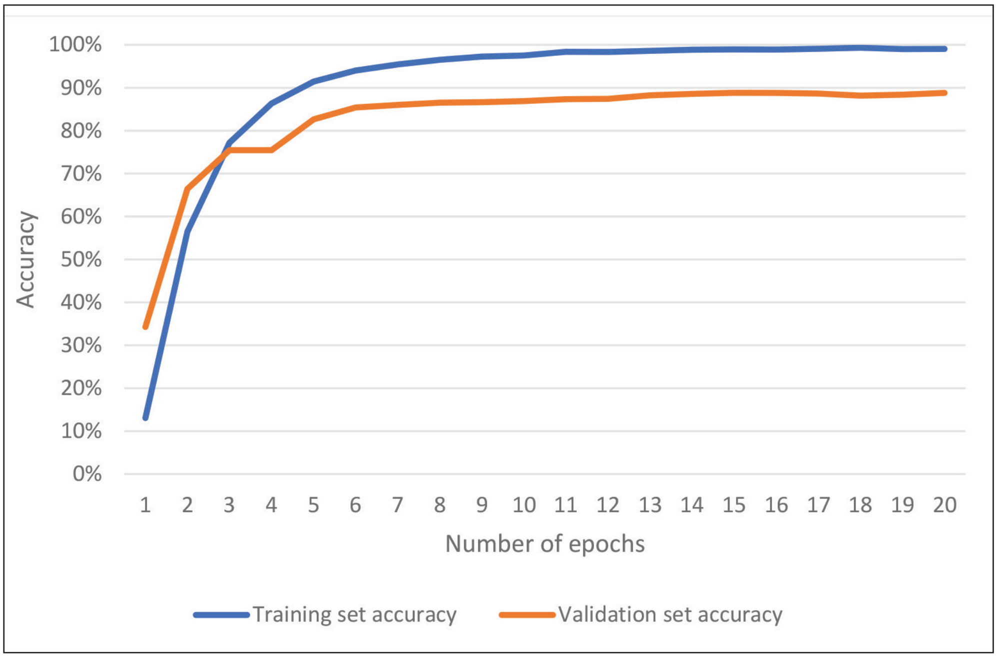


Figure 5 *Results of training after 20 epochs*

Slika 5 *Rezultati treniranja nakon 20 epoha*

The accuracy of the trained model has not changed much between trainings. The accuracy after 20 epochs is 88.81% on the validation part of the dataset. The results can be seen in Figures 4 and 5.

The graphs in Figures 4 and 5 show that there are no major differences between 10 and 20 epochs of training in terms of accuracy. This can be explained in a way that the neural network already converged in the first 10 epochs, and that last layer cannot converge any more to learn new features of this specific dataset. The best achieved accuracy was on the 15th epoch measuring 88.82% on the validation set.

6. CONCLUSION
6. ZAKLJUČAK

In this paper, we have shown how to retrain an existing neural network in the easiest possible way using a dataset which neural network has not been trained on. In only a few hours, the existing neural network can be used to classify completely different objects than it was trained to classify in the first place. For transfer learning, we have used the existing ResNet-152 neural network and retrained only last fully connected layer to classify automobiles from the Stanford car dataset with 88.82% accuracy. This type of training has a lot of room for further improvement. To achieve better results, we could use ResNet-200. Also, if the network was completely trained on the Stanford car dataset the results should be better due to specific shapes of cars, badges, design, etc. The paper shows how using a simple and fast method of transfer learning can yield great results in terms of classification on another dataset.

7. REFERENCE
7. REFERENCES

[1.] M. D. Zeiler and R. Fergus, “LNCS 8689 - Visualizing and Understanding Convolutional Networks,” in ECCV, 2014, pp. 818–833.

[2.] P. Sermanet, D. Eigen, X. Zhang, M. Mathieu, R. Fergus, and Y. LeCun, “Overfeat: Integrated recognition, localization and detection using convolutional networks,” in 2nd International Conference on Learning Representations, ICLR 2014 - Conference Track Proceedings, 2014.

[3.] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” Commun. ACM, vol. 60, no. 6, pp. 84–90, 2017, doi: 10.1145/3065386 ISSN: 15577317.

[4.] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2016, vol. 2016-Decem, pp. 770–778, doi: 10.1109/CVPR.2016.90 ISBN: 9781467388504ISSN: 10636919.

[5.] T. Y. Lin et al., “Microsoft COCO: Common objects in context,” in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2014, vol. 8693 LNCS, no. PART 5, pp. 740–755, doi: 10.1007/978-3-319-10602-1_48 ISSN: 16113349.

[6.] J. Krause, M. Stark, J. Deng, and L. Fei-Fei, “3D object representations for fine-grained categorization,” in Proceedings of the IEEE International Conference on Computer Vision, 2013, pp. 554–561, doi: 10.1109/ICCVW.2013.77 ISBN: 9781479930227.

[7.] O. Russakovsky et al., “ImageNet Large Scale Visual Recognition Challenge,” Int. J. Comput. Vis., vol. 115, no. 3, pp. 211–252, Dec. 2015, doi: 10.1007/s11263-015-0816-y ISSN: 15731405.

[8.] K. He and J. Sun, “Convolutional neural networks at constrained time cost,” in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2015, vol. 07-12-June, pp. 5353–5360, doi: 10.1109/CVPR.2015.7299173 ISBN: 9781467369640ISSN: 10636919.

[9.] R. K. Srivastava, K. Greff, and J. Schmidhuber, “Highway Networks,” arXiv:1505.00387, 2015.

[10.] J. S. Bridle, "Probabilistic Interpretation of Feedforward Classification Network Outputs, with Relationships to Statistical Pattern Recognition," in Neurocomputing, Berlin, Heidelberg: Springer Berlin Heidelberg, 1990, pp. 227–236.

[11.] T. Krambeger, B. Nožica, I. Dodig, and D. Cafuta, "Pregled tehnologija u neuronskim mrežama," Polytech. Des., vol. 7, no. 1, 2019, doi: 10.19279/TVZ.PD.2019-7-1-04

[12.] M. Abadi et al., "TensorFlow: A system for large-scale machine learning," 12th Symp. Oper. Syst. Des. Implement., pp. 265–283, 2016.

[13.] F. Chollet and others, "Keras," 2015. [Online]. Available: https://keras.io.

[14.] Stjepan Ložnjak, rođen je 23. prosinca 1995. godine u Zagrebu. Preddiplomski studij računarstva, smjer programsko inženjerstvo, završio je 2017. godine. Specijalistički diplomski stručni studij završio je 2019. godine, također na Tehničkom veleučilištu u Zagrebu. Tijekom studija radio je kao razvojni inženjer za jednu informatičku tvrtku, a nakon studija odlučio je pokrenuti vlastiti posao. Suosnivač je tvrtke koja se bavi razvojem informacijskih sustava i web aplikacijama. Interesi su mu digitalizacija poslovanja te korištenje informacijskih tehnologija u svakodnevnom životu.

AUTORI · *AUTHORS*

● Stjepan Ložnjak
Rođen je 23. prosinca 1995. godine u Zagrebu. Preddiplomski studij računarstva, smjer programsko inženjerstvo, završio je 2017. godine. Specijalistički diplomski stručni studij završio je 2019. godine, također na Tehničkom veleučilištu u Zagrebu. Tijekom studija radio je kao razvojni inženjer za jednu informatičku tvrtku, a nakon studija odlučio je pokrenuti vlastiti posao. Suosnivač je tvrtke koja se bavi razvojem informacijskih sustava i web aplikacijama. Interesi su mu digitalizacija poslovanja te korištenje informacijskih tehnologija u svakodnevnom životu.

Korespondencija · ***Correspondence***
sloznjak@tvz.hr

● Tin Kramberger - biograpgy can be found in the Polytechnic & Design Vol. 5, No. 1, 2017.

Correspondence · ***Korespondencija***
tkramberg@tvz.hr

● Ivan Cesar - biograpgy can be found in the Polytechnic & Design Vol. 7, No. 2, 2019.

Correspondence · ***Korespondencija***
icesar@tvz.hr

● Renata Kramberger - biograpgy can be found in the Polytechnic & Design Vol. 5, No. 1, 2017.

Correspondence · ***Korespondencija***
rkramberg@tvz.hr