Classifying LEGO bricks – A building block of a LEGO sorting machine

LEGO is one of the most popular toys for children and adults around the world. After the company was founded 1932 in Denmark, it took over 25 years until the LEGO bricks got patented in 1958. From there on LEGO conquered the world as creative toys for young and older children as well as adults. In 2020, the LEGO Group had a revenue of 5.7 billion EUR and a pre-tax profit of 1.6 billion EUR [9]. Beside the new products every year, LEGO has a lively second hand market. The LEGO group participates and promotes this market since it has acquired the LEGO-tailored online store BrickLink in the end of 2019 [1]. The regional Excalibur Store [6] focuses on selling used lego bricks and other toys. The owner told us

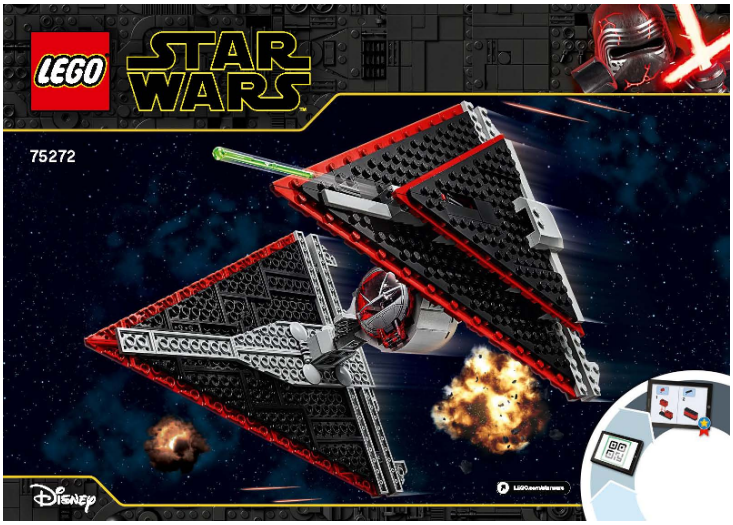

Figure 1: Sith TIE Fighter: The LEGO set to be classified

that retailers specializing on the second hand market face the challenge to sort LEGO bricks. Doing it manually by hand is very time consuming and therefore an automatic sorting solution would be an huge improvement. Therefore, we identified an exciting business case for a classification task. Retailers buy original sets to resell individual bricks that are needed in community-created LEGO sets. A well-maintained public database for lego bricks already exists and can be used as a building block for the classification of bricks. An Image classification method to differentiate parts in one LEGO set is a first step towards an automated sorting of LEGO bricks to support retailers. For this, Excalibur store provided us the LEGO set “Sith TIE Fighter” (set ID 75272) for the project in the course MLME. The set consists of 458 parts. Every single brick can be identified with a unique LEGO ID. The Element ID is unique for one brick. It specifies the design and color of the brick (e.g. a red 2x4 brick). The Design ID, however, only identifies the species regardless of the color (e.g. all 2x4 bricks). The main task of the project is the classification of LEGO bricks in the provided set regarding their Design ID. The students start with a small dataset of approximately 300 images that were recorded in a controlled environment and differ significantly.

Task description

The main task is the classification of the Design ID of lego bricks by convolutional neural networks (CNNs). For this, the focus of the project is on methods to increase the data set size, since the data basis for each class is rather small. Suitable methods shall be selected and discussed and limitations shall also be identified. [7]
We recommend the application of the deep Learning frameworks Keras and Tensorflow. Beside that, the image augmentation library imgaug [3] is a useful addition for the project and we recommend to also have a look at the classical image analysis toolbox OpenCV [2].

Mandatory tasks

The following tasks **have to be completed** in order to pass the project.

- Understand and analyze the provided data
- Apply suitable image augmentation to the data set [7]
- Implement a CNN to train a classifier for the LEGO bricks Design ID for the provided subset (15 classes) of LEGO bricks, we recommend to use mobile net as starting point [5]
- Discuss to what extend image augmentation can improve the classification
- Discuss the quality of the classification in respect to properties of the classes, e.g. texture, geometry, etc.

Additional tasks

Students should complete **some** of these **or other** additional tasks to obtain good or excellent grading for their projects.

- Classification of as many Design IDs as possible with an accuracy $\geq xy\%$
- Apply transfer learning with different networks [4, 8] as basic
- Transfer to a different image set (Internet or photos taken at University)
- Parameter study: Improve net structure (better performance) or training methods (faster training times)
- Extend the classifier to choose the Element ID.

We want to emphasize that the above points are only suggestions for potential further research questions.

Deliverables

The following materials have to be submitted in order to pass the semester project:

- **Recorded final presentation** (video screencast). The presentation must be **5-7 minutes** (for the entire group) and the file should not exceed 200 mb. Highlight on the slides which group member(s) are responsible.
- **Written report** to present and discuss the obtained results. You must use the supplied template on Moodle and write no more than **4-6 pages** (for the entire group). Highlight which group member worked on which section.
- **Source code** of your project. Please ensure that the code is executable and optionally add a short explanation of the structure (readme).

Please submit the files on Moodle. Deadlines for the submissions can be found on Moodle.

Responsible tutor

Please address questions to:

Name	Contact	Role
Tim Janus	tim.janus@tu-dortmund.de	Tutor
Janine Lins	janine.lins@tu-dortmund.de	Tutor
Asha Choudhary	asha.choudhary@tu-dortmund.de	Student assistant
Ali Tabesh	ali.tabesh@tu-dortmund.de	Student assistant

References

[1] Team BrickLink. ***BrickLink will join the LEGO Group to empower the creativity of AFOLs and fuel future innovation***. Nov. 2019. URL: https://www.bricklink.com/v2/community/newsview.page?msgid=1166622.

[2] ***Computer Vision Library OpenCV***. Feb. 2021. URL: https://opencv.org/.

[3] Alexander Jung. ***Image Augmentation Library imgaug - Documentation***. Dec. 2020. URL: https://imgaug.readthedocs.io/en/latest.

[4] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. "ImageNet Classification with Deep Convolutional Neural Networks". In: ***Commun. ACM*** 60.6 (May 2017), pp. 84–90. ISSN: 0001-0782. DOI: 10.1145/3065386. URL: https://doi.org/10.1145/3065386.

[5] Mark Sandler et al. ***MobileNetV2: Inverted Residuals and Linear Bottlenecks***. 2019. arXiv: 1801.04381 [cs.CV].

[6] Oliver Semp. ***LEGO store from the Ruhrarea***. Nov. 2019. URL: https://www.excalibur-store.de/.

[7] Connor Shorten and Taghi M. Khoshgoftaar. "A survey on Image Data Augmentation for Deep Learning". In: ***Journal of Big Data*** 6.60 (2019). ISSN: 2196-1115. URL: https://journalofbigdata.springeropen.com/articles/10.1186/s40537-019-0197-0.

[8] Karen Simonyan and Andrew Zisserman. ***Very Deep Convolutional Networks for Large-Scale Image Recognition***. 2015. arXiv: 1409.1556 [cs.CV].

[9] ***The LEGO Group annual report 2020***. Feb. 2021. URL: https://www.lego.com/cdn/cs/aboutus/assets/bltdee53df07dda2287/Annual_Report_2020_ENG.pdf.