

Data Science Capstone Project

Lego Figurine Classification

1. Dataset

(1/2)

Description

The dataset comprises 327 images of LEGO figurines from different LEGO series (e.g. STAR WARS, SPIDERMAN and MARVEL). The figurines are positioned against different backgrounds and in different positions. The dataset has already been separated in TRAIN and VALID sets.



1. Dataset

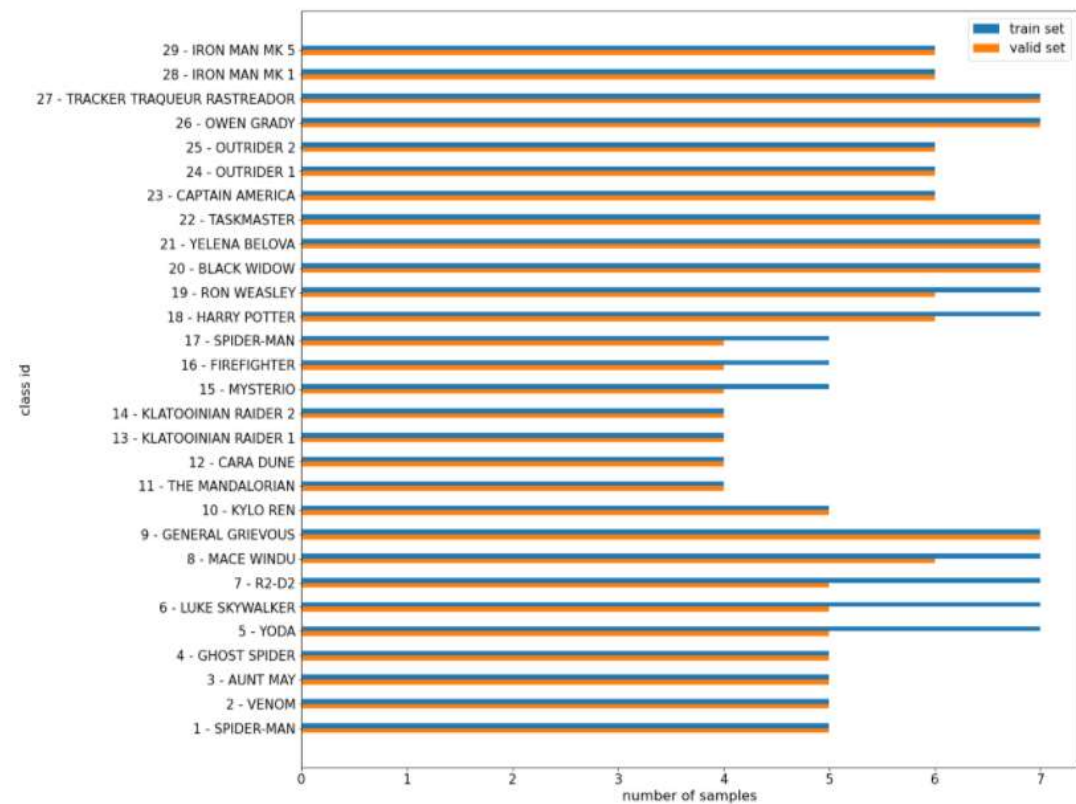
(2/2)

Descriptive Statistics

Table1. Dataset Characteristics

	Unnamed: 0	class_id
count	352.000000	352.000000
mean	175.500000	16.656250
std	101.757883	9.029456
min	0.000000	1.000000
25%	87.750000	8.000000
50%	175.500000	18.000000
75%	263.250000	24.250000
max	351.000000	31.000000

Image1. "Samples" per "class id" (split by TRAIN/VALID).



2. Usecase

CURRENT SCENARIO

- LEGO receives queries from customers about broken parts or parts not fitting LEGO figurines.
- Different teams handle queries on different LEGO series.
- Images sent by customers are checked manually before the query can be sent to the respective service department.

OPTIMAL SCENARIO

- An image classification algorithm could:
 - Recognize the LEGO series of the received service query;
 - Send the service query to the respective department to handle the query for that particular LEGO series.

PROBLEM STATEMENT

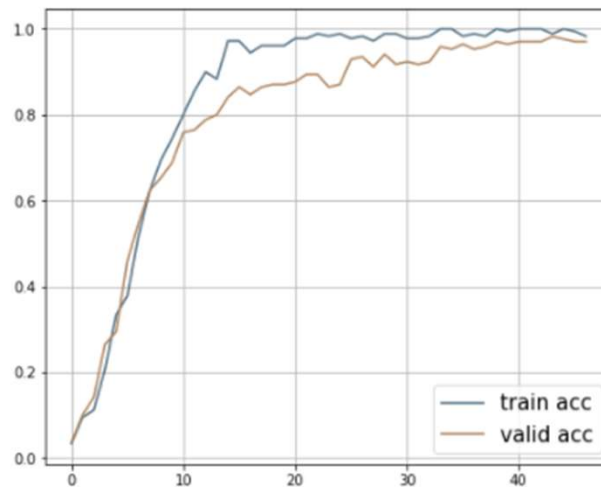
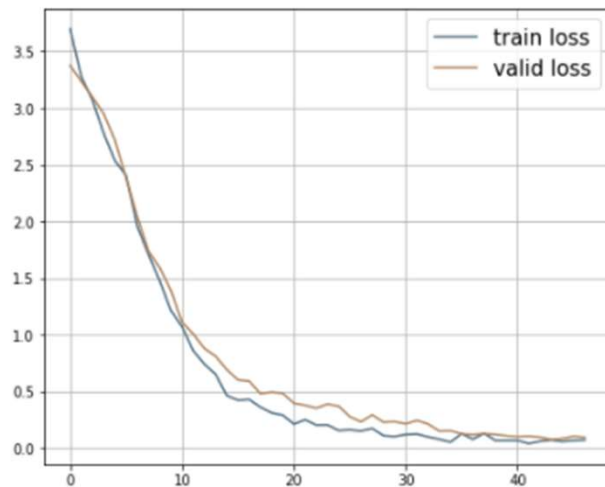
“Can a Convolutional Neural Network analyse images of LEGO figurines to accurately classify the image’s LEGO series?”

3. Solution

CONCLUSION

- The created model is quite successful at correctly classifying the LEGO figurines included in validity data subset (accuracy validation data is 98,2%).
- A convolutional neural network can be successfully applied to automatically classify LEGO figures to handle services queries.
- More data will make this model even more accurate in the future.

Image2. Training vs. Valid Losses



3. Solution

Image3. Confusion Matrix

