

# Deep Learning guinea pig image classification using NVidia DIGITS and GoogLeNet

Lukasz Zmudzinski

University of Warmia and Mazury in Olsztyn

lukasz@zmudzinski.me

**Abstract**—In this paper guinea pig classification using deep learning imaging methods was performed on the Nvidia DIGITS 6. Models capable of distinguishing skinny, abyssinian and crested fur types were created in the process. To increase the classification accuracy empty images (with only the background) were added to the data set. Upon evaluation, the created model for the imported data set produced an accuracy of 98% with inference time of around 3 ms on an p2.xlarge DIGITS instance.

**Index Terms**—classification, deep learning, animal recognition, robotics, DIGITS.

## 1 INTRODUCTION

ROBOTIC systems are nowadays increasingly introduced in various animal care facilities like farms, daries and shelters [1], [2]. With the growing need of automating work more methods of image classification and decision making are needed.

In this work, deep learning techniques were utilized to create a classification model of guinea pigs in different home environments, to explore the possibilities of bringing such systems into the world of household animals.

### 1.1 Nvidia DIGITS

DIGITS (NVIDIA Deep Learning GPU Training System) was used in this project. It allows to rapidly train deep neural networks (DNNs) for image classification, segmentation and object detection tasks [3].

The solution comes with pre-trained models for example:

- GoogLeNet,
- AlexNet,
- UNET and more.

In this paper GoogLeNet was used as the model of choice.

## 2 BACKGROUND / FORMULATION

Deep Learning is growing in popularity over the past few years. It is connected to the fact that it becomes more useful than before thanks to the amount of available training data and advances in computer hardware/software [4].

Another problem with neural networks in general was vanishing gradients and unpredictable weight initialization [5]. This becomes solved by the use of RELUs (Rectified Linear Units) [6] and more recently the GoogLeNet deep learning model [7].

The model is divided into four classes:

- **None** - when there is no guinea pig in the image,
- **Ayssinian**, **Crested** and **Skinny** - fur type.

## 3 DATA ACQUISITION

The data was collected by recording a square video of each guinea pig over a period of 30 seconds in different environments and then extracting each frame as an image. Each frame was then lowered in resolution to 256px in order to fit GoogLeNet requirements. Example images can be seen in Fig. 1.

To ensure good accuracy of the model, each set had to cover front, side and back angles of the guinea pig body. Moreover empty images (without guinea pigs) were added of the provided backgrounds, to increase classification accuracy (and to distinguish background from objects that should be recognized).

Guinea pigs used in the experiment:

- Fifi (4 years, abyssinian),
- Rey (2 years, crested),
- Asajj (2 years, skinny).

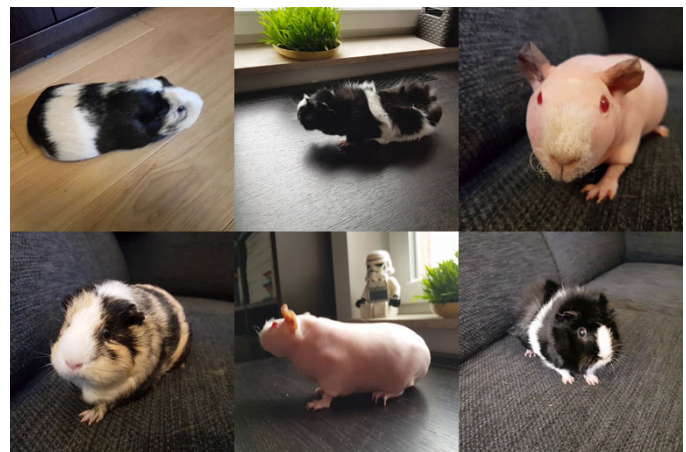


Fig. 1. Example images for abyssinian, crested and skinny guinea pigs from the provided data set.

4 RESULTS

The final classification model was gathered from running the described image set over 15 epochs of GoogLeNet model on Nvidia DIGITS. As you can see on Fig. 2 the accuracy after 15 epochs rose to over 90%, while the final value loss was around 0.3 with a learning rate of 0.01. Using more epochs and a different learning was tested, but it led to overfitting of the data.

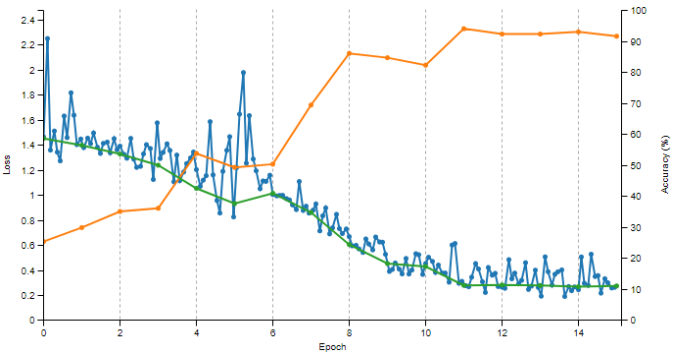


Fig. 2. Graph showing the value accuracy (orange), test data loss (blue) and value loss (green) after running the GoogLeNet model.

After the model was created, a series of tests were performed to check, if the inference is correct. The first set of tests consisted of images from the previously gathered data set. The prediction was always right for provided images, with the value between 70%-95%. You can see an example result on Fig. 3.

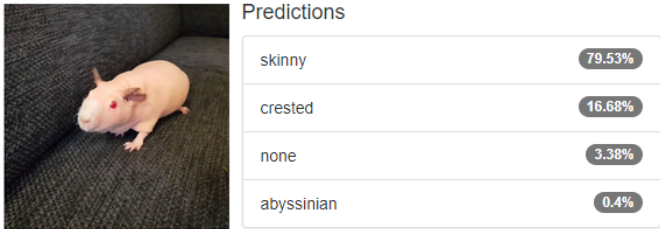


Fig. 3. Example correct prediction results (79.53%) using previously captured images in selected environments using GoogLeNet.

After acquiring enough information on the predictions, tests on new images were performed (different environments, same guinea pigs) giving satisfactory results - guinea pigs were classified correctly on each provided picture containing the animal. Example result on Fig. 4.

One failed prediction was encountered, when a picture of a background with a cat was used. The cat was badly classified as an abyssinian guinea pig.

5 DISCUSSION

TODO

6 CONCLUSION / FUTURE WORK

The created model proves that guinea pig fur recognition for robotic systems is possible. The project gave good results with accuracy of 98% and inference under 10ms. Low inference is especially important for robotic systems

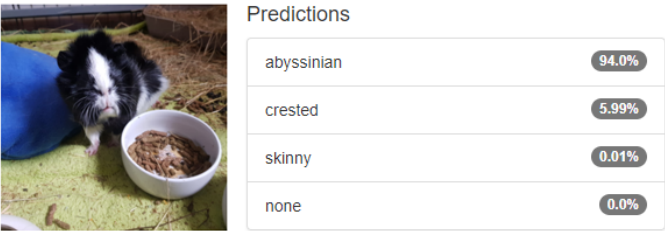


Fig. 4. Example correct prediction results (94.0%) using post-model captured images in animal cage environment using GoogLeNet.

that deal with live animals, because the reaction times need to be low.

Future work might include robotic systems that monitor the state of specific animals, adjust food distribution depending on image readings or alert when the guinea pig suffers from any kind of illness.

REFERENCES

[1] J. J. Roldán, J. del Cerro, D. Garzón-Ramos, P. Garcia-Aunon, M. Garzón, J. de León, and A. Barrientos, "Robots in agriculture: State of art and practical experiences," *Service Robots Antonio Neves, IntechOpen*, 2018. DOI: 10.5772/intechopen.69874. Available from: <https://www.intechopen.com/books/service-robots/robots-in-agriculture-state-of-art-and-practical-experiences>.

[2] M. BD, S. Adil, and S. Ranvir, "Robotics: An emerging technology in dairy and food industry: Review," *International Journal of Chemical Studies*, 2018.

[3] "Nvidia DIGITS." <https://developer.nvidia.com/digits>.

[4] I. Goodfellow and Y. Bengio, *Deep Learning*. Cambridge, Massachusetts: The MIT Press, 2016.

[5] "GoogLeNet presentation," Available from: <https://goo.gl/pPgJGn>.

[6] X. Glorot, A. Bordes, and Y. Bengio, "Deep sparse rectifier networks," *In Proceedings of the 14th International Conference on Artificial Intelligence and Statistics. JMLR W&CP Volume (Vol. 15, pp. 315-323)*.

[7] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Deeper with convolutions," Available from: <http://arxiv.org/abs/1409.4842>.