

Detection of cats in images

ANNA MARIA WALACH (s121540)

Technical University of Denmark

02238 Biometric Systems

The L^AT_EX **acmtrans** document style formats articles in the style of the ACM transactions. Users who have prepared their document with L^AT_EX can, with very little effort, produce camera-ready copy for these journals.

Categories and Subject Descriptors: D.2.7 [**Software Engineering**]: Distribution and Maintenance—*documentation*; H.4.0 [**Information Systems Applications**]: General

General Terms: Algorithms, Experimentation

Additional Key Words and Phrases: cat, detection, images

1. INTRODUCTION

There are more and more projects that uses live streaming and social media power to help and safe feral cats. Some of them are meant to control the population in the area [Society 2015], other focus on raising awareness about feral cats, spaying and neutering importance [Kittens 2015] or just to increase changes of finding a new home for homeless cats and kittens [Room 2015].

People are also more interested in health and safety of their own pets. Lower prices of video technologies and smarthone popularity helped to create various apps for controlling your pet lifestyle. [Zillians 2015] created an automatic feeder with cat identification system and connected it to the smart-phone app (product can be seen in fig. 1. It allows owner to supervise the amount of dry food and water his cats are consuming everyday, to receive alarms in case of abnormalities and even spy on them while eating.



Fig. 1. CatFi product used by a cat.

The PiP [Company 2015] app is designated to help people find their lost cats and dogs. According to American Human Society, almost 3.5 million pets are lost each year and about 80% of dogs and 98% of cats are never re-united with their families. The authors of the smarthone app allows the owners to made profiles for their pets (including pictures) and in case the pet is missing, you launch the "alarm" on that pet, together with last seen location. If anyone finds a homeless animal, he may send the photo of it to the PiP app, which automatically try to match one of the lost pets with found pet.

Increasing number of pets' photographs also encouraged camera producers to implement appropriate support for it in their devices. Right now, Fujifilm company

introduced Auto Dog / Cat Detection function [FujiFilm 2015] in some of their cameras that allows to automatically detect pet's face on the image and put auto-focus on it.

1.1 Problem description

Almost all of the applications and initiatives described in previous section already use or would benefit from the cats detection and/or identification. In this research I would like to focus on topic of cat detection in images. The main interest is on the use case of detecting cats presence in shots from live cameras and streams. Such images are characterized by diversified, often bad lighting settings, with possible presence of another animals or human beings. I will perform a state-of-the-art survey and then perform a comparison of existing, available algorithms.

2. THE STATE OF THE ART

Reliable comparison of existing software for detecting cats in pictures is hard to performed. It is caused by a few factors:

- black-box algorithms** - many algorithms, that claim to have high accuracy, are part of the hardware product [Zillians 2015] or closed-source app [Company 2015]. This way it is impossible to make a theoretical comparison of the algorithms and usually significantly hampers possibility of performing experiments with test data, as the products' interfaces are not adjusted to bulk operations.
- high diversity of use cases** - although there are a lot of studies focusing on cat face detection, usually the classification or testing environmental does not check robustness of the algorithm in typical real-time scenarios. The classification is e.g. made as cat vs still life [Fleuret and Geman 2007; Akihiko Yamada et al. 2011] or as cat/dog breed detection [Omkar M. Parkhi et al. 2012].
- different test sets** - even for algorithms with similar use cases, e.g. cat vs still life, different test set are in use to estimate the accuracy of algorithm, which makes it significantly harder to perform direct comparisons. Usually articles use their own data sets [Omkar M. Parkhi et al. 2012; Fleuret and Geman 2007; Akihiko Yamada et al. 2011] and only some of them compare to standard data bases, like Pascal 2007 cats database [Omkar M. Parkhi et al. 2012].

As the main idea of this research is to find an algorithm that will allow for highly accurate detection of cat faces in the pictures, only articles presenting "cat vs still life" classification will be taken under consideration.

2.1 Unsupervised image categorization

One of the more interesting approaches to this problem was presented in [Le et al. 2012]. Researchers from Google and Stanford university constructed large scale unsupervised learning tool. It was a 9-layered neural network build on a cluster that consisted of 1 000 machines with total of 16 000 cores. The network was trained with 10 million unlabelled images, all of them had the same size and local contrast normalization was applied on them. The model consisted of 1 bilion connections and it took 3 days to train the classifiers. Final results are very promising for overall image categorization issue, because authors claims to have 15.8% accuracy

for ImageNet dataset (22 000 object categories), which is a 70% improvement in comparison to the previous state of the art.

However, the most interesting part of [Le et al. 2012] for this research is that after performing face-detection experiment, authors tried to use trained classifiers to detect cats. The results are very promising - the random guess had 64.8% chance, while this algorithm was able to detect cats with 74.8% accuracy.

This algorithm and use of deep learning networks seems to be a promising technique for classification of images categories, however seemed to be lacking needed accuracy for binary classification (cat vs not-cat).

2.2 Pet detection method for digital camera

Another interesting project was performed by [Akihiko Yamada et al. 2011]. They tried to construct a cat and dogs face detection mechanism for use in camera auto-focus and auto-exposure. They were not satisfied with current method, because camera has a very high constraints about calculation time and power consumption, were most of the studies focus on high recall and small false-detection.

Authors decided to use four direction features as an input for the classifier. In order to speed-up the process, classification of an image window was hierarchical - if coarse classifier did not detect a sign of face in the window, it was automatically rejected and not send to more detailed and feature-rich fine classifier.

The results are very promising. As described in fig. 2, with the average recall on 90%, the false positive count is around 20. Although the FPC may seem high (the sample size was 400, with 200 of non-animal pictures), it is still on user-acceptable level. What is more important, the processing time was only 245ms on ARM9 device with 166MHz core. It is impressive result, especially because this solution turned out to be also very robust to illumination changes. Authors measured how color and intensity fluctuation caused by illumination changes influence the recall rate. Intensity changes didn't cause the recall rate to drop under 80% even on very high intensity values. With 50% color fluctuation recall rate still oscillate between 80%-90%.

Although the results are not perfect (usually recall rate do not exceed 90%), short processing time allows to use this algorithm in real-time cases, with may be important case with live streaming observation. This solution can be used as a part of pet-identification algorithm, that informs appropriate person when the cat appears in the camera (e.g. if it is known it is pregnant and needs to be caught and checked) and in such cases time is always a crucial value. It could be an interesting experiment to see if a little bit of power efficiency (i.e. better processing unit) could significantly improve the recall rate.

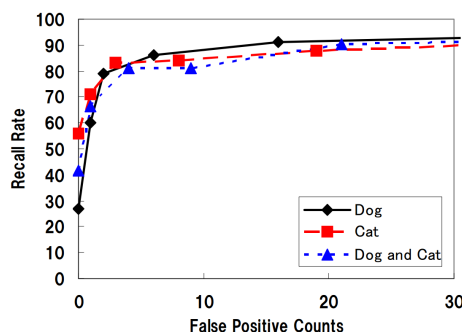


Fig. 2. ROC curves achieved as a result of [Akihiko Yamada et al. 2011].

2.3 Stationary features and cat detection

2.4 Shape and texture features for cat face detection

The solution presented in [Zhang et al. 2008] is the only one with existing available implementation. Also, as part of it, the on-line interface was provide for testing [Heather 2015]. Authors of [Zhang et al. 2008] decided, that both texture and shape are important features of cat face detection. Their solution has two steps - in first, they create separately trained shape and texture detector, both based on Haar-like features. As a second step they are jointly training fusion classifier (joint shape and texture). Both shape and texture classifier were trained on differently normalized data. Their features are based on oriented gradients.

Important part of that study was creating a database of 10 000 annotated cat pictures [Zhang et al. 2015] (annotations of 9 points in the cat face, like ears triangles, nose, mouth). This pictures will be the part of this research dataset.

The results of that study are presented in fig. 3. The recall rate oscillate around 90% with 100 FPC for the joint texture and shape detector. Because huge focus of [Zhang et al. 2008] was on feature engineering and comparing their features with different standard approaches, authors also made a direct comparison, using Pascal 2007 data set. They claim that Average Precision (AP) of their approach is 0.364, when best reported method has AP equalled to 0.24.

The study seems to find interesting and quite straight forward method of detecting cats face. Coming from human face detection, they adjusted their feature choice with respect to differences appearing in cat face and then tested different type of gradient features to picked the most suitable one for this aim. However, it should be noted that this article is 7 years old and probably some of the studies presented in previous sections may have better results, but it is hard to estimate, as they are using different databases for testing.

2.5 Summation

3. CUSTOM FRAMEWORK

shape - comparison of different approaches

fujifilm - photo restrictions

it is important that such algorithm is robust to pose variation and illumination level.

3.1 Problem analysis

3.2 Proposed method

3.2.1 Features.

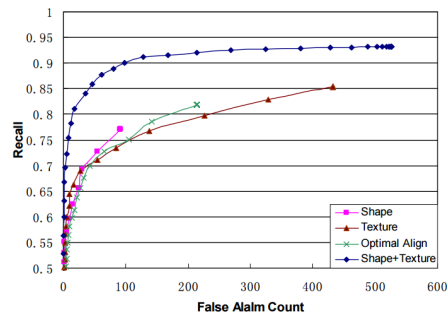


Fig. 3. ROC curves achieved as a result of [Zhang et al. 2008].

3.2.2 Classifier.

3.3 Bibliography

REFERENCES

- AKIHIKO YAMADA, KAZUHIRO KOJIMA, J. K., OKAMOTO, M., AND MURATA, H. 2011. Directional edge-based dog and cat face detection method for digital camera. *IEEE International Conference on Consumer Electronics*.
- COMPANY, T. P. R. 2013 (accessed June 22, 2015). *PiP - reuniting lost pets with their families*. <http://www.petrecognition.com/>.
- FLEURET, F. AND GEMAN, D. 2007. Stationary features and cat detection. *Idiap Research Report*.
- FUJIFILM. 2009 (accessed June 22, 2015). *Auto Dogs & Cats Detection*. http://www.fujifilm.com/products/digital_cameras/pet/.
- HEATHER, A. 2013 (accessed June 22, 2015). *Kittydar - face detection for cat*. <http://harthur.github.io/kittydar/>.
- KITTENS, T. 2015 (accessed June 22, 2015). *Rescue Kitten Reality TV*. <http://www.tinykittens.com/>.
- LE, Q. V., RANZATO, M. A., MONGA, R., DEVIN, M., CHEN, K., CORRADO, G. S., DEAN, J., AND NG, A. Y. 2012. Building high-level features using large scale unsupervised learning.
- OMKAR M. PARKHI, A. V., ZISSERMAN, A., AND JAWAHAR, C. V. 2012. Cats and dogs.
- ROOM, T. C. 2015 (accessed June 22, 2015). *Foster Kitten Cam*. <https://livestream.com/FosterKittenCam>.
- SOCIETY, L. A. P. 2015 (accessed June 22, 2015). *Trap-Neuter-Return (TNR) Pilot*. <http://www.tinykittens.com/cats>.
- ZHANG, W., SUN, J., AND TANG, X. 2008. Cat head detection - how to effectively exploit shape and texture features. *Proc. of European Conf. Computer Vision 4*, 802–816.
- ZHANG, W., SUN, J., AND TANG, X. 2008 (accessed June 22, 2015). *Cat database*. <http://137.189.35.203/WebUI/CatDatabase/catData.html>.
- ZILLIANS. 2015 (accessed June 22, 2015). *CatFi - Smart Feeder*. <http://catfi.com/>.

Received June 2015;