Final Paper: Individual Animal Identification from Video

Deadline: 28th May 2022, 23:59

1 Preliminaries

Weight: 60% of the total mark of the module.

Total number of points: 100

Deadline: 28th May 2022, 23:59. The submission point will close automatically thereafter.

Submission: Submit one pdf file with your paper in the *Final Paper* submission point. Then submit separate file or files with your Python code in the *Final Paper Code* submission point. The paper itself should not contain pieces of code.

2 Content

Your task is to write a report which contains components of a scientific paper. Your report should be titled "Individual Animal Identification from Video".

The problem of animal identification and reidentification from video frames and image collections is gaining importance in the context of climate change. For your task, you will use data from five annotated videos containing pigeons, koi fish and pigs. The videos are available here: `https://github.com/LucyKuncheva/Animal-Identification-from-Video`. The animals in each video have been given names. The training data is the first half of the video. The task is to recognise the individual animals in the second half of the video.

Each image (frame of the video) is manually annotated with bounding boxes. Each bounding box contains one animal. Examples of bounding boxes from the videos are shown in Figure 1.

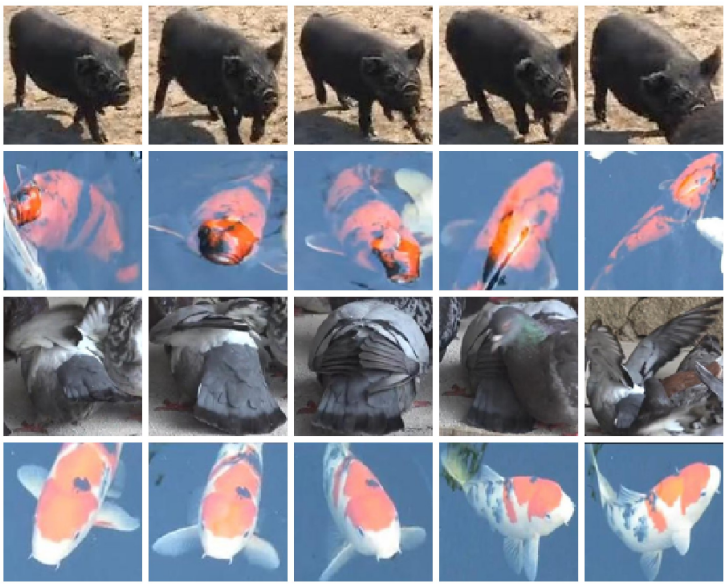

Figure 1: Examples of clips from four animals.

3 Data

The following videos are included in the dataset:

Video name	Training size	Testing size
Pigs_49651_960_540_500f	2710	3212
Koi_5652_952_540	916	719
Pigeons_8234_1280_720	2268	2291
Pigeons_4927_960_540_600f	1574	1303
Pigeons_29033_960_540_300f	2148	2241

The data has been preprocessed for you. The bounding boxes have been stored. Each bounding box is an object in the data set of the respective video. Four feature representations have been extracted: RGB features (colour), H10 - hue histogram features (colour), HOG features (shape), and LBP features (texture). The class labels are provided in separate csv files. For example, the class labels for video `Pigs_49651_960_540_500f` are in file `Pigs_49651_960_540_500f_Labels.csv`. The four data representations for the same video are in the respective csv files:

```
Pigs_49651_960_540_500f_DataRGB.csv
Pigs_49651_960_540_500f_DataHOG.csv
Pigs_49651_960_540_500f_DataH10.csv
Pigs_49651_960_540_500f_DataLBP.csv
```

The data in the files (features and labels) are arranged so that the first part is the training data, and the second part is the test data. However, there is a caveat! The objects in the data should not be split in the middle. The *video frames* are split in the middle. Hence, the training and testing sizes shown above.

4 Research task and structure of your paper

Your task is to find out how different classifiers can recognise individual animals in the given data collection.

4.1 Introduction and Related Work

A zip file with references is uploaded to Blackboard. You should educate yourself on the topic of animal identification from images and video. The literature I have given you is excessive but does not cover everything. You are welcome to look for more literature, especially newer titles. You DON'T have to read all papers! You don't need to read any paper from cover to cover! The papers are there for you to see the general trends - what people have done, how they do animal identification, what software is available, what animals can be identified, how important the topic is, etc. Using this literature and other literature that you find yourself, you should prepare a literature review.

Your paper should contain an *Introduction* section, which will explain the importance of the problem and the aims and objectives of your paper. Next, you will have a section *Related Work*, in which you will write your literature review. While writing the review, bear the following points in mind:

- Choose a structure - sort and present the papers by topics, not one-by-one.
- Try to relate your review to the task of this paper! Choose the references that are most relevant for your project.
- The more papers you include, the better. You should have seen each paper you cite.
- Write SHORT sentences - about 15 words. Long sentences are NOT "more scientific"

4.2 Data description

Make sure that you describe the parameters of your data sets (relevant sizes). The details can be obtained from the Github site and from the data files. Check whether the data is balanced or not.

4.3 Classification methods

Explain the classification methods that you will be comparing. Include the following classifiers (you can use more, if you like):

- Largest Prior classifier (don't use a Python library for this one)
- Linear Discriminant Analysis
- 3-nn
- Decision Tree
- SVM
- Bagging
- Random Forest

4.4 Results

Choose a nice way to present the results. Pages full of numbers will not be useful. Consider showing tables and confusion matrices.

4.5 Conclusion

In a conclusion section of about half a page, recap the work you have done and what the outcome is. Include also future possibilities for improving the results, if applicable.

4.6 Formatting and typesetting

Use LaTeX for your report. Do not include code in the report. Write the report as a scientific document so that a reader with your level of expertise understands it. If you like, use the class article with the following preamble:

```
\documentclass[12pt]{article}

\usepackage[cmex10]{amsmath}
\usepackage{amssymb,color}
\usepackage{lineno,hyperref}
\usepackage{url}
\usepackage{amsfonts}
\usepackage[pdftex]{graphicx}
\usepackage{amsmath}
\usepackage{iftex}
\ifPDFTeX
\usepackage[T1]{fontenc}
```

```
\usepackage{mathpazo}
\else
\usepackage{fontspec}
\fi
\usepackage{geometry}
\geometry{verbose,tmargin=1in,bmargin=1in,lmargin=1in,rmargin=1in}
\setlength{\parindent}{0pt}
\setlength{\parskip}{8pt}
```

You can use the ShareLaTeX system of our School, which is a version of Overleaf. Cameron Gray maintains the system. You can ask her to set up an account for you. You will get a message from the system with a username and password. Please login asap because the login will expire within one week. Check in your Junk folder if you don't receive the system invite. When you set up your project in the system, you can edit it from anywhere. There is an option to check the number of words you currently have in the document. If you like, you can share the document with me (from the Share button on the top right). I will be able to see your progress, answer questions, and fix LaTeX things that may not be working properly.

It is expected that your document will contain approximately 2000 words. Deviations from this number, in either direction, are acceptable.

5 Evaluation

Your work will be evaluated on the basis of your text and your technical expertise.

1	Introduction	[5]
2	Related Work	[15]
3	Methodology	[15]
4	Data description	[12]
5	Experiment and results (code)	[33]
6	Conclusion	[4]
7	References	[6]
8	English, writing style and grammar	[6]
9	LaTeX and appearance	[4]
	Total	[100]

Best of luck!