IMAGE AND FACE CLASSIFICATION USING, SVM, REGRESSION TREES AND LOGISTIC REGRESSION

By
Clyde Clarke

Outline

Problem Statement

- Object Dataset
- Face Dataset

Image Classification

Feature Extraction

- PCA
- Hough
- Color Histogram
- Harlick

Classification Methods

- SVM
- Regression Trees
- CNN

Evaluation

The Object and Face Datasets

- 1) Caltech 101
- Pictures of objects belonging to 101 categories. About 40 to 800 images per category. Most categories have about 50 images. Collected in September 2003 by Fei-Fei Li, Marco Andreetto, and Marc 'Aurelio Ranzato. The size of each image is roughly 300 x 200 pixels.
- 2) Labeled Faces in the Wild
- Database of face photographs The data set contains more than 13,000 images of faces collected from the web. Each face has been labeled with the name of the person pictured.

Image Classification

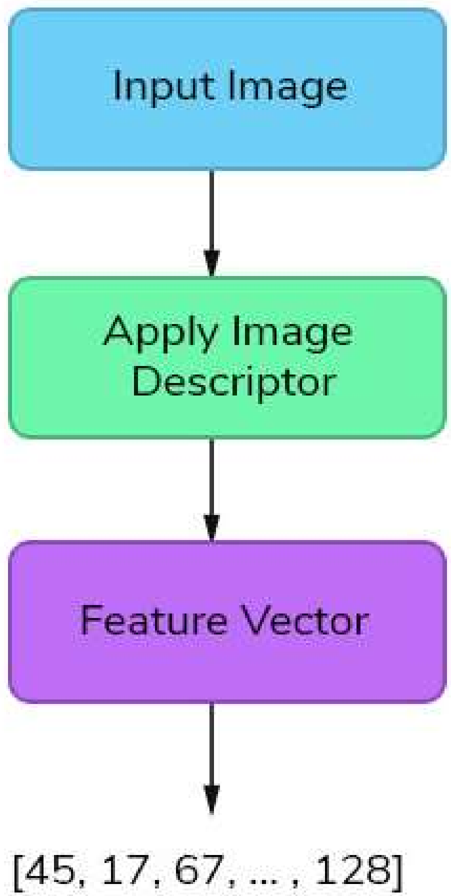
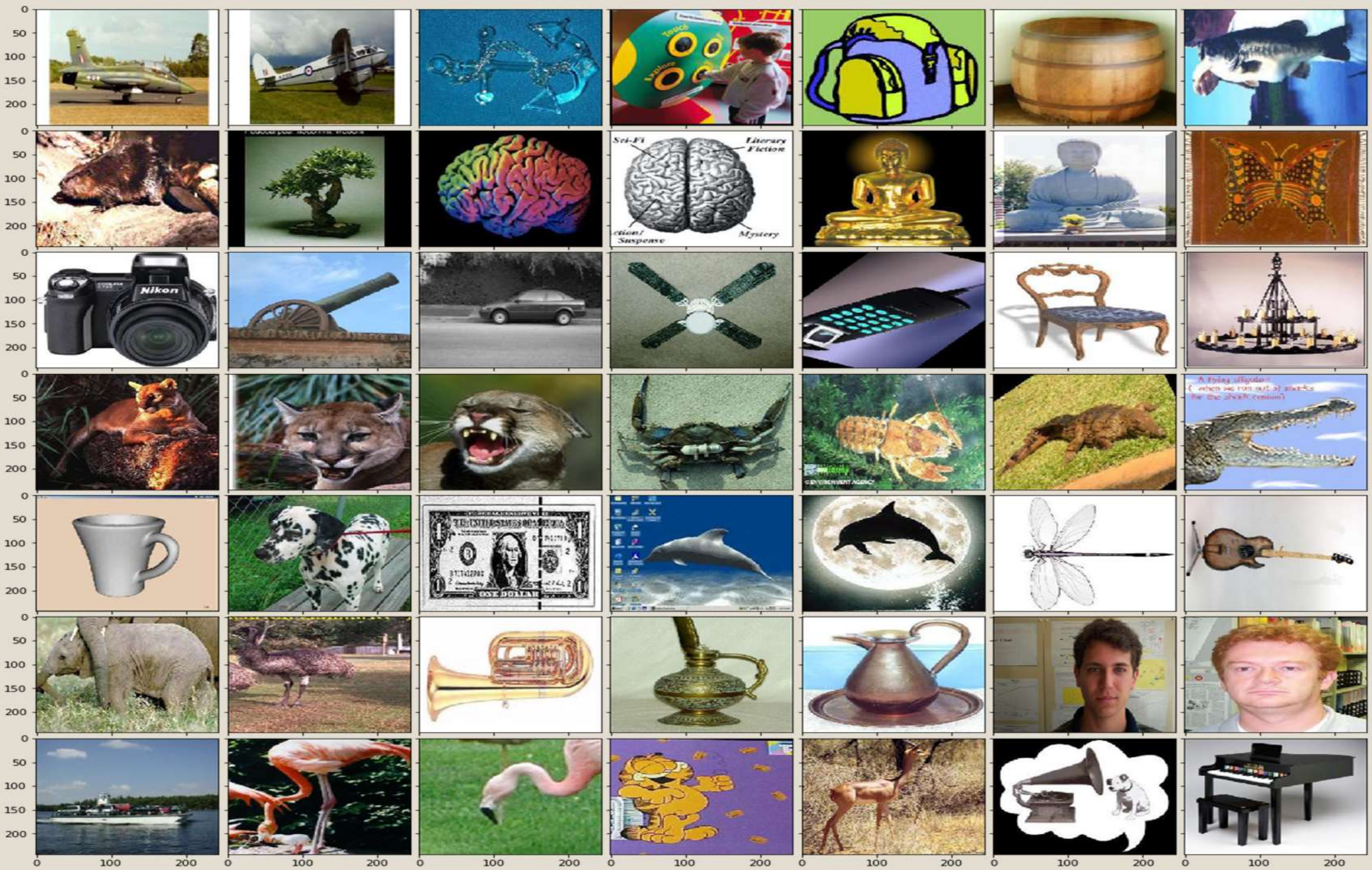
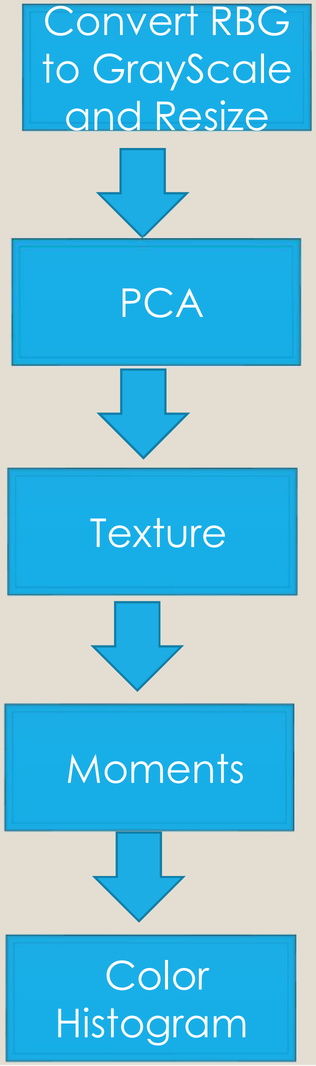


Image Classes Objects (50 +)

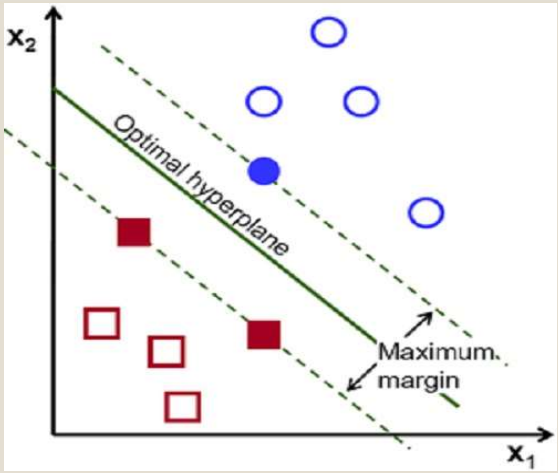


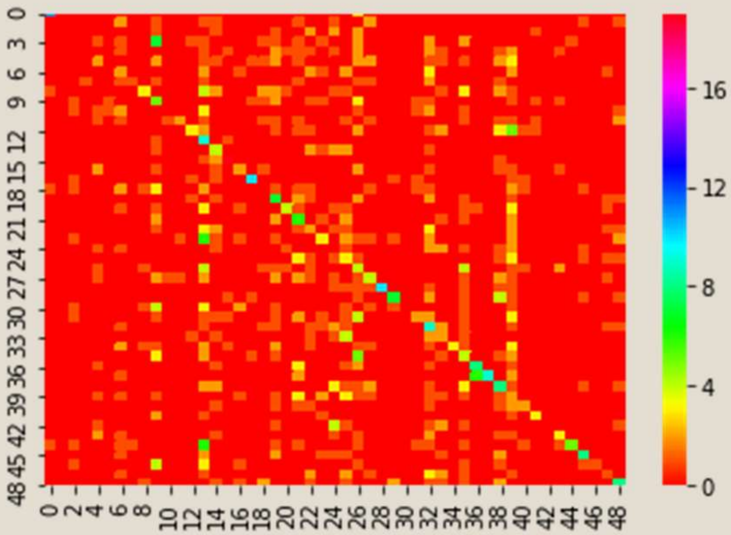


Feature Extraction

- PCA
  - Flatten image into vector array and apply PCA for first 150 components
- Texture
  - Done to prevent gibbs phenomena at frame edges that would corrupt spectral analysis
- Hue Moments
  - Find centers of color throughout the image
    -

Grayscale Images





SVM

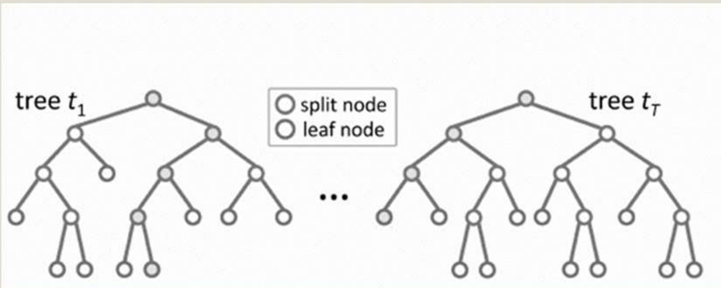
- Large Margin Classifier
- Maximizes the distance between decision boundary and support vectors
- Kernel
  - Radial Basis Function
- Accuracy
  - 43% on training data and
  - 26% accuracy on test data

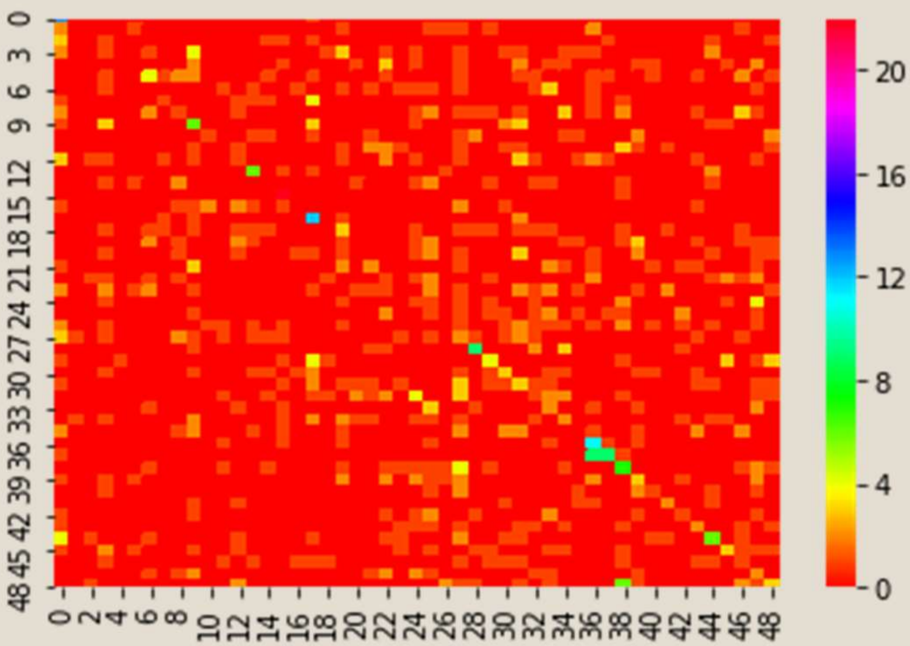
SVM Class Test Scores

.ing object names on the SVM test set

	precision	recall	f1-score	support
airplanes	1.00	0.79	0.88	14
anchor	1.00	0.09	0.17	11
ant	0.00	0.00	0.00	10
UND_Google	0.00	0.00	0.00	22
barrel	0.00	0.00	0.00	18
bass	0.33	0.04	0.08	23
beaver	0.06	0.06	0.06	16
binocular	0.33	0.09	0.14	11
bonsai	0.20	0.04	0.07	24
brain	0.11	0.29	0.16	21
ontosaurus	0.00	0.00	0.00	14
buddha	0.29	0.12	0.17	17
butterfly	0.38	0.13	0.19	23
camera	0.15	0.82	0.25	11
cannon	0.22	0.31	0.26	13
car_side	0.76	0.83	0.79	23
eiling_fan	0.20	0.15	0.17	13
cellphone	0.73	0.65	0.69	17
chair	0.04	0.05	0.04	20
chandelier	0.17	0.35	0.23	20
ougar_body	0.23	0.16	0.19	19
ougar_face	0.12	0.06	0.08	17
crab	0.05	0.11	0.07	18
crayfish	0.57	0.16	0.25	25

	precision	recall	f1-score	support
crocodile	0.00	0.00	0.00	14
crocodile_head	0.09	0.38	0.15	16
cup	0.09	0.09	0.09	22
dalmatian	0.31	0.22	0.26	23
dollar_bill	0.91	0.56	0.69	18
dolphin	0.60	0.30	0.40	20
dragonfly	0.00	0.00	0.00	17
electric_guitar	0.50	0.05	0.08	22
elephant	0.14	0.46	0.21	24
emu	0.13	0.23	0.17	13
euphonium	0.75	0.15	0.25	20
ewer	0.16	0.20	0.18	20
Faces	0.56	0.56	0.56	18
Faces_easy	0.91	0.50	0.65	20
ferry	0.22	0.27	0.24	26
flamingo	0.05	0.10	0.07	21
flamingo_head	0.13	0.30	0.18	10
garfield	0.38	0.38	0.38	8
gerenuk	0.00	0.00	0.00	12
gramophone	0.43	0.23	0.30	13
grand_piano	0.78	0.30	0.44	23
hawksbill	0.42	0.42	0.42	19
headphone	1.00	0.07	0.13	14
hedgehog	0.12	0.07	0.09	15
helicopter	0.36	0.41	0.38	22
accuracy			0.24	870
macro avg	0.33	0.23	0.23	870
weighted avg	0.33	0.24	0.23	870





Random Forest

- Random forests (RF) are a combination of tree predictors
- Each tree depends on the values of a random vector sampled in dependently
- The generalization error depends on the strength of the individual trees and the correlation between them
  - Accuracy
    - .87 on training data
    - .17 on test data

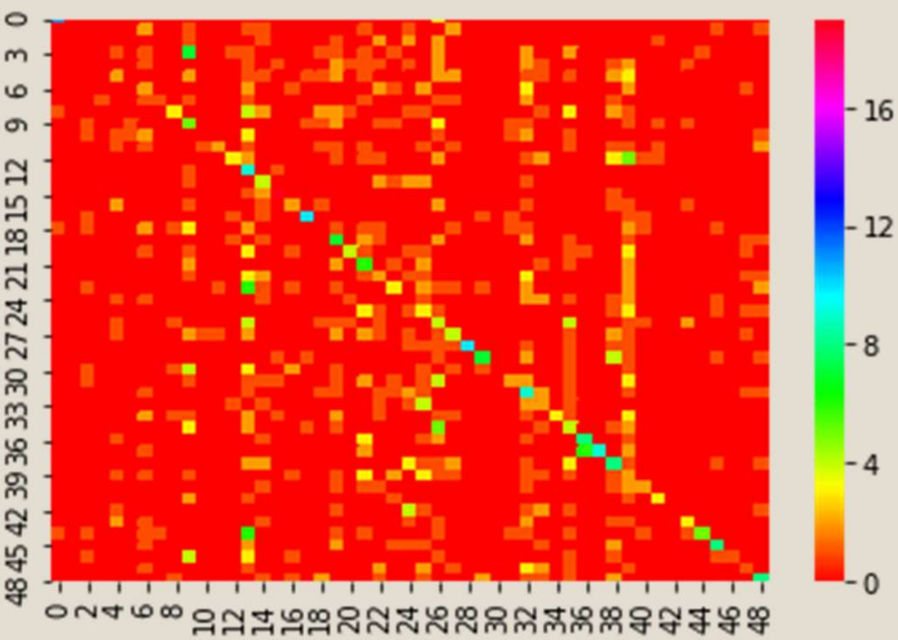
Random Forest Class Scores

Predicting object names on the test set

	precision	recall	f1-score	support
airplanes	0.32	0.86	0.47	14
anchor	0.22	0.18	0.20	11
ant	0.00	0.00	0.00	10
BACKGROUND_Google	0.24	0.18	0.21	22
barrel	0.00	0.00	0.00	18
bass	0.00	0.00	0.00	23
beaver	0.14	0.06	0.09	16
binocular	0.40	0.18	0.25	11
bonsai	0.09	0.04	0.06	24
brain	0.44	0.52	0.48	21
brontosaurus	0.00	0.00	0.00	14
buddha	0.00	0.00	0.00	17
butterfly	0.33	0.09	0.14	23
camera	0.18	0.27	0.21	11
cannon	0.00	0.00	0.00	13
car_side	0.68	0.91	0.78	23
ceiling_fan	0.00	0.00	0.00	13
cellphone	0.38	0.47	0.42	17
chair	0.33	0.05	0.09	20
chandelier	0.10	0.10	0.10	20
cougar_body	0.00	0.00	0.00	19
cougar_face	0.14	0.18	0.15	17
crab	0.05	0.06	0.05	18
crayfish	0.00	0.00	0.00	25

	precision	recall	f1-score
crocodile	0.00	0.00	0.00
crocodile_head	0.05	0.06	0.05
cup	0.04	0.05	0.04
dalmatian	0.07	0.09	0.08
dollar_bill	0.19	0.50	0.27
dolphin	0.21	0.30	0.25
dragonfly	0.09	0.12	0.10
electric_guitar	0.00	0.00	0.00
elephant	0.13	0.12	0.13
emu	0.03	0.08	0.04
euphonium	0.16	0.30	0.21
ewer	0.10	0.05	0.07
Faces	0.26	0.33	0.29
Faces_easy	0.54	0.65	0.59
ferry	0.22	0.15	0.18
flamingo	0.14	0.24	0.18
flamingo_head	0.00	0.00	0.00
garfield	0.13	0.25	0.17
gerenuk	0.00	0.00	0.00
gramophone	0.00	0.00	0.00
grand_piano	0.33	0.26	0.29
hawksbill	0.25	0.26	0.26
headphone	0.31	0.29	0.30
hedgehog	0.10	0.13	0.11
helicopter	0.21	0.23	0.22

accuracy			0.17	870
macro avg	0.14	0.17	0.14	870
weighted avg	0.14	0.17	0.14	870



Logistic Regression

- Logistic regression is used to describe data and to explain the relationship between one dependent binary variable and one or more nominal, ordinal, interval or ratio-level independent variables.
- A one vs rest classifier is used in combination with logistic regression to assign a probability that each image object will be assigned to a class

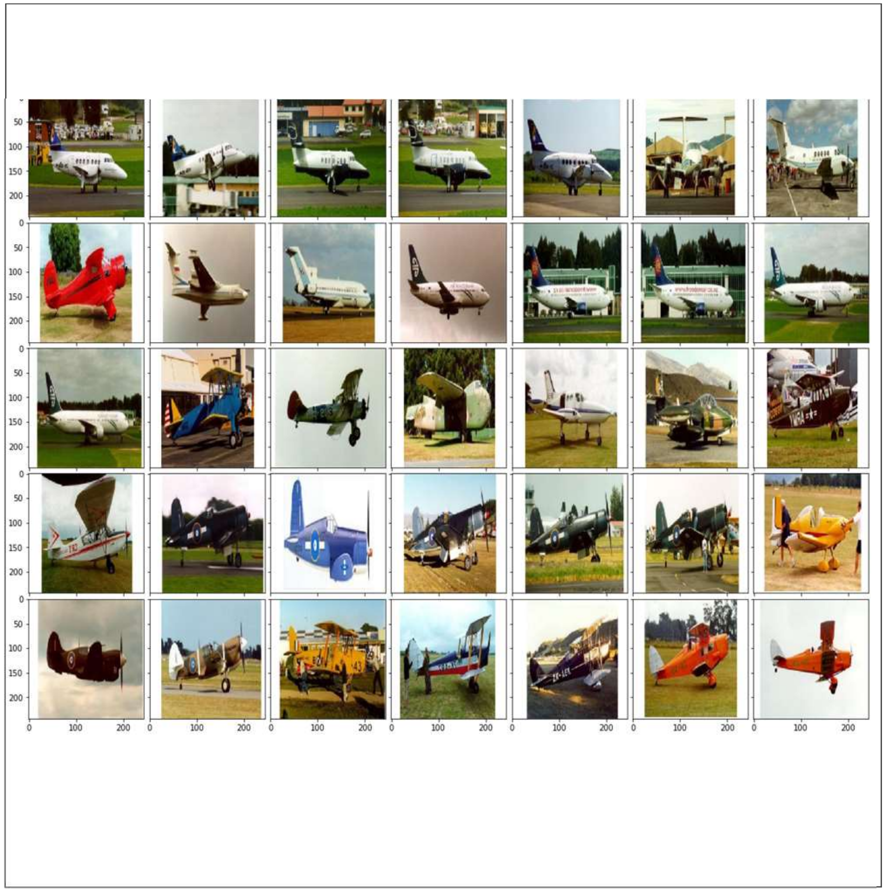
◦ Accuracy
  ◦ 54 on training data
  ◦ .26 on test data

Logistic Regression Class Test Scores

Predicting object names on the Logistic Regression test set

	precision	recall	f1-score	suppor
airplanes	0.41	1.00	0.58	1
anchor	0.33	0.09	0.14	1
ant	0.17	0.10	0.12	1
BACKGROUND_Google	0.11	0.09	0.10	2
barrel	0.00	0.00	0.00	1
bass	0.50	0.04	0.08	2
beaver	0.00	0.00	0.00	1
binocular	0.50	0.09	0.15	1
bonsai	0.17	0.04	0.07	2
brain	0.24	0.38	0.30	2
brontosaurus	0.00	0.00	0.00	1
buddha	0.28	0.29	0.29	1
butterfly	0.38	0.13	0.19	2
camera	0.35	0.55	0.43	1
cannon	0.60	0.23	0.33	1
car_side	0.61	1.00	0.75	2
ceiling_fan	0.25	0.15	0.19	1
cellphone	0.58	0.65	0.61	1
chair	0.09	0.20	0.12	2
chandelier	0.23	0.25	0.24	2
cougar_body	0.00	0.00	0.00	1
cougar_face	0.20	0.06	0.09	1
crab	0.19	0.28	0.23	1
crayfish	1.00	0.08	0.15	2

	precision	recall	f1-score	support
crocodile	0.00	0.00	0.00	14
crocodile_head	0.11	0.06	0.08	16
cup	0.00	0.00	0.00	22
dalmatian	0.35	0.30	0.33	23
dollar_bill	0.46	0.67	0.55	18
dolphin	0.37	0.35	0.36	20
dragonfly	0.67	0.12	0.20	17
electric_guitar	0.25	0.14	0.18	22
elephant	0.15	0.46	0.22	24
emu	0.04	0.31	0.07	13
euphonium	0.80	0.20	0.32	20
ewer	0.12	0.25	0.16	20
Faces	0.56	0.78	0.65	18
Faces_easy	0.86	0.90	0.88	20
ferry	0.24	0.23	0.24	26
flamingo	0.03	0.05	0.04	21
flamingo_head	0.14	0.30	0.19	10
garfield	0.23	0.38	0.29	8
gerenuk	0.00	0.00	0.00	12
gramophone	0.29	0.15	0.20	13
grand_piano	0.75	0.39	0.51	23
hawksbill	0.36	0.42	0.39	19
headphone	0.75	0.21	0.33	14
hedgehog	0.00	0.00	0.00	15
helicopter	0.33	0.36	0.35	22
accuracy			0.26	870
macro avg	0.31	0.26	0.24	870
weighted avg	0.32	0.26	0.24	870

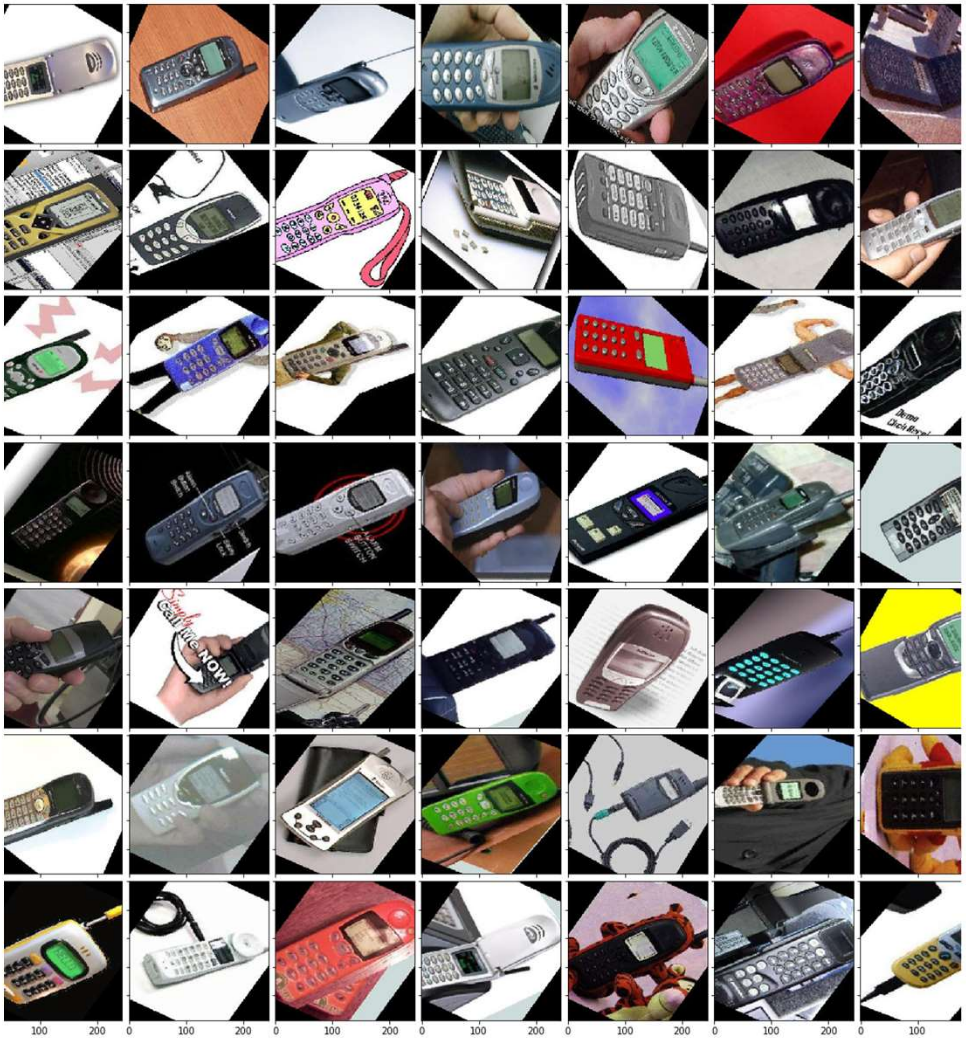


Classification Good

- Specific classes from the data were scored with high accuracy
- The airplanes were scored with 89% accuracy
  - Image background and orientation are mostly consistent within this class of the dataset
- This is true regardless of the classification model used with higher accuracy for classes with consistent image orientation and background

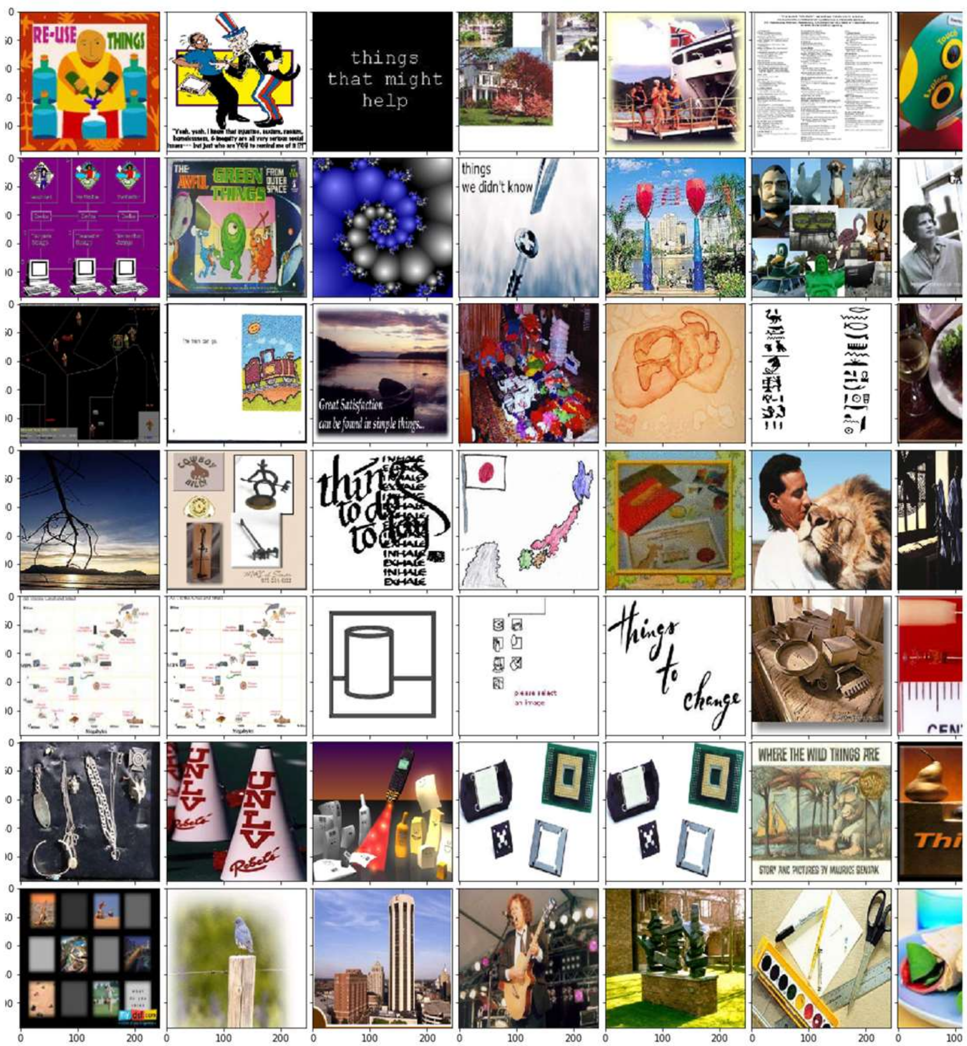
Classification Good

- As seen the following class of cars has a consistent angle of image as well as a reasonably consistent background
  - The best performing model was able to obtain a 79% accuracy score on this class of data



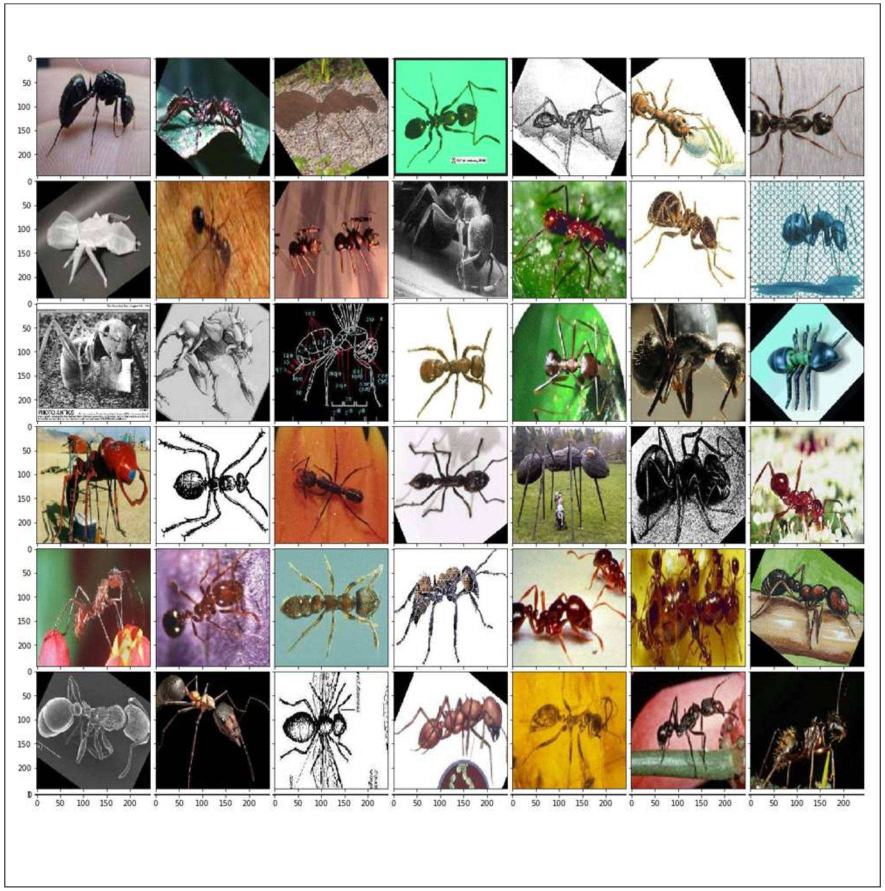
Classification Good

- As seen the following class of cell has a consistent angle of image as well as a reasonably consistent background
  - The best performing model was able to obtain a 61% accuracy score on this class of data in SVM Training



Classification Bad

- Poorly performing classes had either inconsistent image content, orientation or background
  - Accuracy for most models range between 0-7% accuracy for this class labeled ‘Google Backgrounds”



Classification Bad

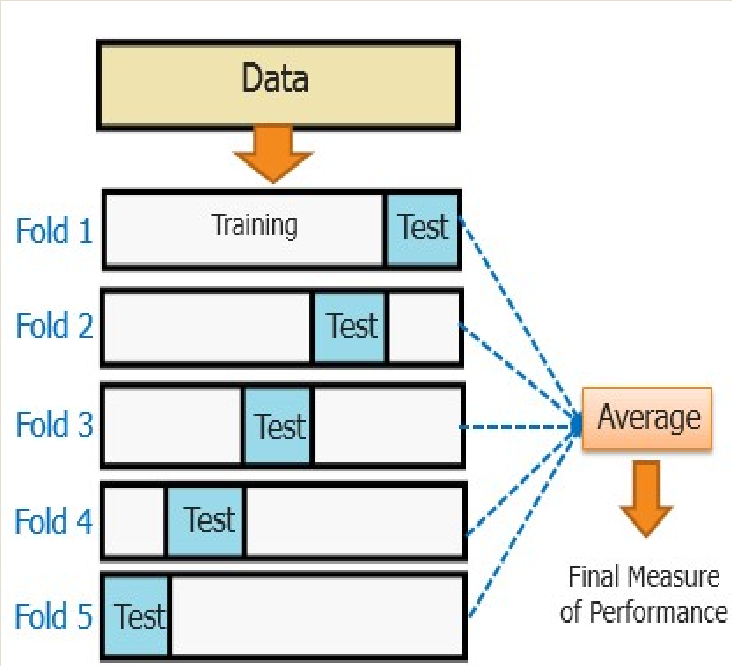
- Poorly performing classes had either inconsistent image content, orientation or background
  - Accuracy for most models range between 0-7% accuracy for this class labeled ‘Google Backgrounds”

LABELED FACES IN THE WILD

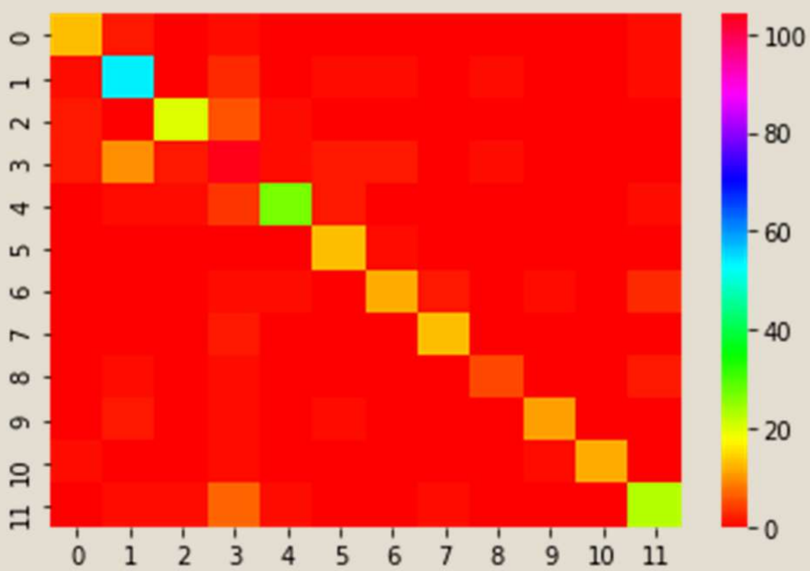
Dataset consists of grayscale images with consistent backgrounds

Training and Processing

- 5 Fold Crossvalidation was done using an 80/20 split of the data
- PCA was performed along with the aforementioned feature extractions
- A batch size of 32 and 100 epoch was used to train two convnet architectures

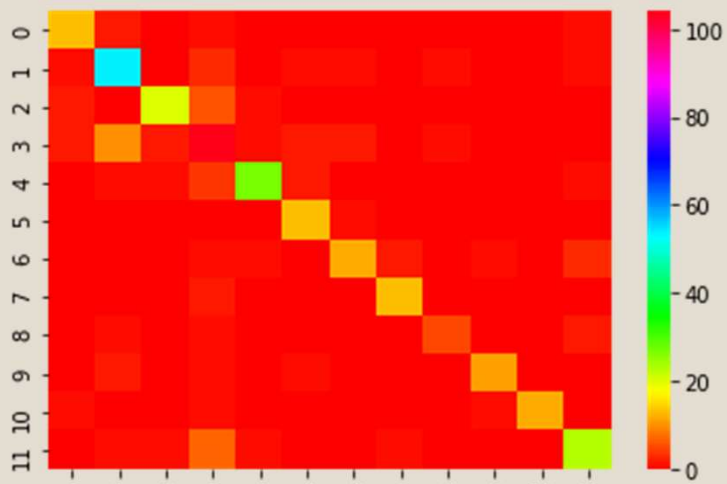


SVM Classification



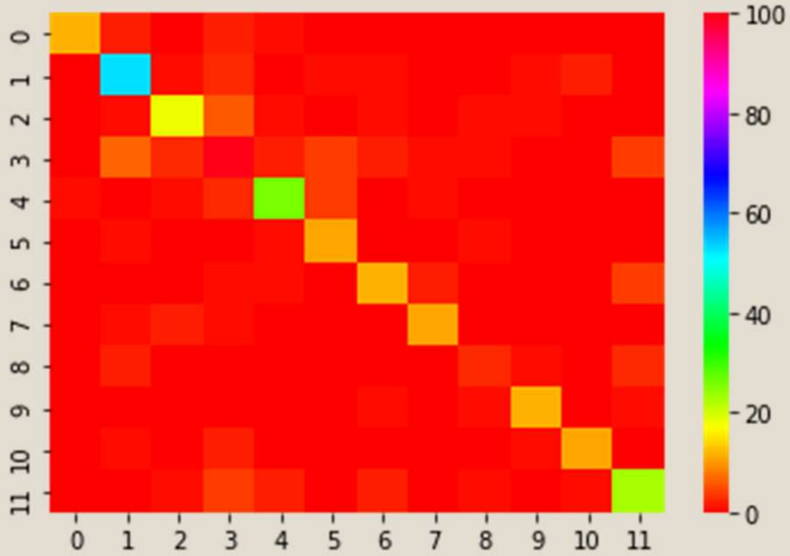
Predicting people's names on the test set				
done in 0.529s				
	precision	recall	f1-score	support
Ariel Sharon	0.68	0.76	0.72	17
Colin Powell	0.76	0.87	0.81	62
Donald Rumsfeld	0.83	0.69	0.75	29
George W Bush	0.79	0.84	0.82	124
Gerhard Schroeder	0.87	0.75	0.81	36
Hugo Chavez	0.68	0.93	0.79	14
Jacques Chirac	0.75	0.60	0.67	20
Jean Chretien	0.81	0.87	0.84	15
John Ashcroft	0.71	0.56	0.63	9
Junichiro Koizumi	0.85	0.73	0.79	15
Serena Williams	1.00	0.80	0.89	15
Tony Blair	0.74	0.68	0.71	34
accuracy			0.79	390
macro avg	0.79	0.76	0.77	390
weighted avg	0.79	0.79	0.79	390

Random Forest Classification



	precision	recall	f1-score	support
Ariel Sharon	0.56	0.29	0.38	17
Colin Powell	0.63	0.63	0.63	62
Donald Rumsfeld	0.29	0.14	0.19	29
George W Bush	0.52	0.90	0.66	124
Gerhard Schroeder	0.57	0.22	0.32	36
Hugo Chavez	0.41	0.50	0.45	14
Jacques Chirac	0.00	0.00	0.00	20
Jean Chretien	0.60	0.20	0.30	15
John Ashcroft	0.50	0.11	0.18	9
Junichiro Koizumi	0.67	0.27	0.38	15
Serena Williams	0.76	0.87	0.81	15
Tony Blair	0.43	0.38	0.41	34
accuracy			0.54	390
macro avg	0.50	0.38	0.39	390
weighted avg	0.51	0.54	0.48	390

Logistic Regression Classifier



	precision	recall	f1-score	support
Ariel Sharon	0.92	0.71	0.80	17
Colin Powell	0.78	0.85	0.82	62
Donald Rumsfeld	0.69	0.62	0.65	29
George W Bush	0.82	0.81	0.81	124
Gerhard Schroeder	0.76	0.72	0.74	36
Hugo Chavez	0.55	0.79	0.65	14
Jacques Chirac	0.63	0.60	0.62	20
Jean Chretien	0.73	0.73	0.73	15
John Ashcroft	0.38	0.33	0.35	9
Junichiro Koizumi	0.75	0.80	0.77	15
Serena Williams	0.79	0.73	0.76	15
Tony Blair	0.66	0.68	0.67	34
accuracy			0.75	390
macro avg	0.71	0.70	0.70	390
weighted avg	0.75	0.75	0.75	390

Conclusions

- Feature Extraction was insufficient to deal with large variations in dataset
- Choosing a dataset with less variation improved accuracy
- Perhaps other features that account for object orientation and background could improve performance of supervised learning techinques