

[About](#) [Citation Policy](#) [Donate a Data Set](#) [Contact](#)☒ Repository ☐ Web [View ALL Data Sets](#)

# Statlog (Vehicle Silhouettes) Data Set

*Download:* [Data Folder](#), [Data Set Description](#)

**Abstract:** 3D objects within a 2D image by application of an ensemble of shape feature extractors to the 2D silhouettes of the objects.



<b>Data Set Characteristics:</b>	Multivariate	<b>Number of Instances:</b>	946	<b>Area:</b>	N/A
<b>Attribute Characteristics:</b>	Integer	<b>Number of Attributes:</b>	18	<b>Date Donated</b>	N/A
<b>Associated Tasks:</b>	Classification	<b>Missing Values?</b>	N/A	<b>Number of Web Hits:</b>	73015

## Source:

SOURCE:

Drs.Pete Mowforth and Barry Shepherd  
Turing Institute  
George House  
36 North Hanover St.  
Glasgow  
G1 2AD

CONTACT:

Alistair Sutherland  
Statistics Dept.  
Strathclyde University  
Livingstone Tower  
26 Richmond St.  
GLASGOW G1 1XH  
Great Britain

Tel: 041 552 4400 x3033

Fax: 041 552 4711

e-mail: [alistair '@' uk.ac.strathclyde.stams](mailto:alistair '@' uk.ac.strathclyde.stams)

## Data Set Information:

The purpose is to classify a given silhouette as one of four types of vehicle, using a set of features extracted from the silhouette. The vehicle may be viewed from one of many different angles.

#### HISTORY:

This data was originally gathered at the TI in 1986-87 by JP Siebert. It was partially financed by Barr and Stroud Ltd. The original purpose was to find a method of distinguishing 3D objects within a 2D image by application of an ensemble of shape feature extractors to the 2D silhouettes of the objects. Measures of shape features extracted from example silhouettes of objects to be discriminated were used to generate a classification rule tree by means of computer induction.

This object recognition strategy was successfully used to discriminate between silhouettes of model cars, vans and buses viewed from constrained elevation but all angles of rotation.

The rule tree classification performance compared favourably to MDC (Minimum Distance Classifier) and k-NN (k-Nearest Neighbour) statistical classifiers in terms of both error rate and computational efficiency. An investigation of these rule trees generated by example indicated that the tree structure was heavily influenced by the orientation of the objects, and grouped similar object views into single decisions.

#### DESCRIPTION:

The features were extracted from the silhouettes by the HIPS (Hierarchical Image Processing System) extension BINATTS, which extracts a combination of scale independent features utilising both classical moments based measures such as scaled variance, skewness and kurtosis about the major/minor axes and heuristic measures such as hollows, circularity, rectangularity and compactness.

Four "Corgie" model vehicles were used for the experiment: a double decker bus, Cheverolet van, Saab 9000 and an Opel Manta 400. This particular combination of vehicles was chosen with the expectation that the bus, van and either one of the cars would be readily distinguishable, but it would be more difficult to distinguish between the cars.

The images were acquired by a camera looking downwards at the model vehicle from a fixed angle of elevation (34.2 degrees to the horizontal). The vehicles were placed on a diffuse backlit surface (lightbox). The vehicles were painted matte black to minimise highlights. The images were captured using a CRS4000 framestore connected to a vax 750. All images were captured with a spatial resolution of 128x128 pixels quantised to 64 greylevels. These images were thresholded to produce binary vehicle silhouettes, negated (to comply with the processing requirements of BINATTS) and thereafter subjected to shrink-expand-expand-shrink HIPS modules to remove "salt and pepper" image noise.

The vehicles were rotated and their angle of orientation was measured using a radial graticule beneath the vehicle. 0 and 180 degrees corresponded to "head on" and "rear" views respectively while 90 and 270 corresponded to profiles in opposite directions. Two sets of 60 images, each set covering a full 360 degree rotation, were captured for each vehicle. The vehicle was rotated by a fixed angle between images. These datasets are known as e2 and e3 respectively.

A further two sets of images, e4 and e5, were captured with the camera at elevations of 37.5 degs and 30.8 degs respectively. These sets also contain 60 images per vehicle apart from e4.van which contains only 46 owing to the difficulty of containing the van in the image at some orientations.

#### Attribute Information:

##### ATTRIBUTES

COMPACTNESS (average perim)\*\*2/area

CIRCULARITY (average radius)\*\*2/area

DISTANCE CIRCULARITY area/(av.distance from border)\*\*2

RADIUS RATIO (max.rad-min.rad)/av.radius

PR.AXIS ASPECT RATIO (minor axis)/(major axis)

MAX.LENGTH ASPECT RATIO (length perp. max length)/(max length)

SCATTER RATIO (inertia about minor axis)/(inertia about major axis)

ELONGATEDNESS area/(shrink width)\*\*2

PR.AXIS RECTANGULARITY area/(pr.axis length\*pr.axis width)

MAX.LENGTH RECTANGULARITY area/(max.length\*length perp. to this)

SCALED VARIANCE (2nd order moment about minor axis)/area  
ALONG MAJOR AXIS

SCALED VARIANCE (2nd order moment about major axis)/area  
ALONG MINOR AXIS

SCALED RADIUS OF GYRATION (mavar+mivar)/area

SKEWNESS ABOUT (3rd order moment about major axis)/sigma\_min\*\*3  
MAJOR AXIS

SKEWNESS ABOUT (3rd order moment about minor axis)/sigma\_maj\*\*3  
MINOR AXIS

KURTOSIS ABOUT (4th order moment about major axis)/sigma\_min\*\*4  
MINOR AXIS

KURTOSIS ABOUT (4th order moment about minor axis)/sigma\_maj\*\*4  
MAJOR AXIS

HOLLOWS RATIO (area of hollows)/(area of bounding polygon)

Where sigma\_maj\*\*2 is the variance along the major axis and sigma\_min\*\*2 is the variance along the minor axis,  
and

area of hollows= area of bounding poly-area of object

The area of the bounding polygon is found as a side result of the computation to find the maximum length. Each individual length computation yields a pair of calipers to the object orientated at every 5 degrees. The object is propagated into an image containing the union of these calipers to obtain an image of the bounding polygon.

NUMBER OF CLASSES

4 OPEL, SAAB, BUS, VAN

## Relevant Papers:

Turing Institute Research Memorandum TIRM-87-018 "Vehicle Recognition Using Rule Based Methods" by Siebert,JP (March 1987)  
[[Web Link](#)]

## Papers That Cite This Data Set<sup>1</sup>:



Ken Tang and Ponnuthurai N. Suganthan and Xi Yao and A. Kai Qin. [Linear dimensionality reduction using relevance weighted LDA](#). School of Electrical and Electronic Engineering Nanyang Technological University. 2005. [\[View Context\]](#).

Ping Zhong and Masao Fukushima. [A Regularized Nonsmooth Newton Method for Multi-class Support Vector Machines](#). 2005. [\[View Context\]](#).

Remco R. Bouckaert and Eibe Frank. [Evaluating the Replicability of Significance Tests for Comparing Learning Algorithms](#). PAKDD. 2004. [\[View Context\]](#).

Dmitry Pavlov and Alexandrin Popescul and David M. Pennock and Lyle H. Ungar. [Mixtures of Conditional Maximum Entropy Models](#). ICML. 2003. [\[View Context\]](#).

Gisele L. Pappa and Alex Alves Freitas and Celso A A Kaestner. [Attribute Selection with a Multi-objective Genetic Algorithm](#). SBIA. 2002. [\[View Context\]](#).

James Bailey and Thomas Manoukian and Kotagiri Ramamohanarao. [Fast Algorithms for Mining Emerging Patterns](#). PKDD. 2002. [\[View Context\]](#).

Robi Polikar and L. Upda and S. S. Upda and Vasant Honavar. [Learn++: an incremental learning algorithm for supervised neural networks](#). IEEE Transactions on Systems, Man, and Cybernetics, Part C, 31. 2001. [\[View Context\]](#).

Thomas G. Dietterich. [An Experimental Comparison of Three Methods for Constructing Ensembles of Decision Trees: Bagging, Boosting, and Randomization](#). Machine Learning, 40. 2000. [\[View Context\]](#).

Thierry Denoeux. [A neural network classifier based on Dempster-Shafer theory](#). IEEE Transactions on Systems, Man, and Cybernetics, Part A, 30. 2000. [\[View Context\]](#).

Richard Maclin. [Boosting Classifiers Regionally](#). AAAI/IAAI. 1998. [\[View Context\]](#).

Robert E. Schapire and Yoav Freund and Peter Bartlett and Wee Sun Lee. [The Annals of Statistics, to appear. Boosting the Margin: A New Explanation for the Effectiveness of Voting Methods](#). AT&T Labs. 1998. [\[View Context\]](#).

Ron Kohavi and Mehran Sahami. [Error-Based and Entropy-Based Discretization of Continuous Features](#). KDD. 1996. [\[View Context\]](#).

Ron Kohavi. [A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection](#). IJCAI. 1995. [\[View Context\]](#).

Jeffrey P. Bradford and Clayton Kunz and Ron Kohavi and Clifford Brunk and Carla Brodley. [Appears in ECML-98 as a research note Pruning Decision Trees with Misclassification Costs](#). School of Electrical Engineering. [\[View Context\]](#).

Gisele L. Pappa and Alex Alves Freitas and Celso A A Kaestner. [AMultiobjective Genetic Algorithm for Attribute Selection](#). Computing Laboratory Pontificia Universidade Catolica do Parana University of Kent at Canterbury. [\[View Context\]](#).

Chih-Wei Hsu and Cheng-Ru Lin. [A Comparison of Methods for Multi-class Support Vector Machines](#). Department of Computer Science and Information Engineering National Taiwan University. [\[View Context\]](#).

Yin Zhang and W. Nick Street. [Bagging with Adaptive Costs](#). Management Sciences Department University of Iowa Iowa City. [\[View Context\]](#).

H. Altay Guvenir. [A Classification Learning Algorithm Robust to Irrelevant Features](#). Bilkent University, Department of Computer Engineering and Information Science. [\[View Context\]](#).

Alexander K. Seewald. [Dissertation Towards Understanding Stacking Studies of a General Ensemble Learning Scheme ausgeführt zum Zwecke der Erlangung des akademischen Grades eines Doktors der technischen Naturwissenschaften](#). [\[View Context\]](#).

Adil M. Bagirov and Alex Rubinov and A. N. Soukhojak and John Yearwood. [Unsupervised and supervised data classification via nonsmooth and global optimization](#). School of Information Technology and Mathematical Sciences, The University of Ballarat. [\[View Context\]](#).

Ron Kohavi and George H. John. [Automatic Parameter Selection by Minimizing Estimated Error](#). Computer Science Dept. Stanford University. [\[View Context\]](#).

Rajesh Parekh and Jihoon Yang and Vasant Honavar. [Constructive Neural-Network Learning Algorithms for Pattern Classification](#). [\[View Context\]](#).

Vikas Sindhwani and P. Bhattacharya and Subrata Rakshit. [Information Theoretic Feature Crediting in Multiclass Support Vector Machines](#). [\[View Context\]](#).

Maria Salamo and Elisabet Golobardes. [Analysing Rough Sets weighting methods for Case-Based Reasoning Systems](#). Enginyeria i Arquitectura La Salle. [\[View Context\]](#).

Ronaldo C. Prati and Peter A. Flach. [ROCCER: an Algorithm for Rule Learning Based on ROC Analysis](#). Institute of Mathematics and Computer Science University of São Paulo. [\[View Context\]](#).

## Citation Request:

This dataset comes from the Turing Institute, Glasgow, Scotland. If you use this dataset in any publication you must acknowledge this source.

---

[1] Papers were automatically harvested and associated with this data set, in collaboration with [Rexa.info](#)



In Collaboration With:



[About](#) || [Citation Policy](#) || [Donation Policy](#) || [Contact](#) || [CML](#)