From Aardvark to Zorro: A Benchmark of Mammal Images

Michael Fink
Interdisciplinary Center for Neural Computation
The Hebrew University of Jerusalem
Jerusalem 91904, Israel
fink@huji.ac.il

Shimon Ullman
Faculty of Mathematics and Computer Science
Weizmann Institute of Science
Rehovot 76100, Israel
shimon.ullman@weizmann.ac.il

Abstract

Current object recognition systems aim at achieving two challenging goals: recognizing numerous object classes and learning new object classes from a small number of examples. This paper provides a benchmark for evaluating progress on these fundamental tasks. Several methods have recently proposed to utilize the commonalities between object classes in order to effectively acquire new object classes. Such methods can be termed *interclass transfer* techniques. However, it is currently difficult to asses which of the proposed methods maximally utilizes the shared structure of related classes. To facilitate the development, as well as the assessment of methods for dealing with multiple related classes, the proposed benchmark provides labeled images of over 400 mammal classes. The images are organized in five levels of variability, and their labels provide information on the objects identity, location and pose. A labeled benchmark containing a large number of related classes is crucial for assessing how well proposed solutions scale up, and which methods better utilize interclass transfer.

1 Introduction

This paper describes a benchmark for evaluating methods aimed at recognizing numerous object classes. In the last few years a significant improvement is observed in the capabilities of object recognition methods. Reliable systems have been proposed for recognizing *individual* objects (e.g. a specific mug), tuned to overcome occlusions, changing illumination conditions and pose variations [10, 8, 12, 1]. Moreover, for certain objects (like faces and cars), reliable recognition on the *class* level has been demonstrated as well (see for example [13]). However, handling the within class variability in other object classes (e.g. recognizing all mugs), remains a challenging task. The proliferation of machine learning

Figure 1: Example images from the Aardvark search separated to the five database tiers

techniques into the computer vision community is manifested by the fact that many current object recognition systems rely on an extensive training stage to handle class variability. A systematic evaluation of these methods must examine both recognition performance and training requirements. Training requirements are typically assessed by the number of examples needed for reaching a certain performance measure and by the type of labeling assumed as input (e.g. image label or bounding box). Two intertwined goals have recently emerged in the frontier of object recognition systems:

1. learning to recognize numerous object classes
2. learning to recognize an object class from very few examples

The need for recognizing many objects emerges naturally in settings such as query by image content and robotic vision applications. While, learning from few examples is motivated by the fact that training data is often expensive to acquire or otherwise scarce. The human perceptual system has the remarkable capacity to recognize numerous objects, often learning to reliably classify a novel object from just a short exposure to a single example. Inspired by the human capabilities, it has been hypothesized that to accomplish such tasks, recognition systems must rely on interclass transfer. It could be said that a recognition system benefits from interclass transfer, if the multiple target classification tasks share common underlying structures that can be utilized to facilitate training and augment detection. The machine learning literature provides several past attempts at formalizing interclass transfer [2, 15]. Recently, the object recognition community has discovered the value of interclass transfer. By applying interclass transfer methods, a significant improvement over single object recognition systems has been reported by several researchers [7, 11, 4, 14, 9, 6]. It should be noted that besides the common emphasis on interclass transfer, these methods are widely diverse. While some are based on generative modeling, others are purely discriminative. The information transferred between the classes varies accordingly, from priors on object configurations to reusable discriminative features. Regrettably, unlike other object recognition domains (e.g. face detection), the object recognition community lacks the appropriate data sets required for evaluating the merits and drawbacks of proposed interclass transfer methods. Both in the generative setting proposed by [4] and in the discriminative model proposed by [14] the information transfered between classes seems to be quite restricted (priors on object configurations in the former and edges and corner locations in the latter). However, this effect does not seems to truly characterize the proposed algorithms, but rather emerges due to the fact that the evaluated classes were only weakly related (e.g. cars and street signs). It is clear that if the target classes are not related, little if any interclass transfer might be observed. It could be expected that interclass transfer might be fully manifested, only in experiments training on many classes that share significant common structures. Thus, the interesting question is which of the current approaches performs best in scenarios where object classes do share significant common structures. In order to enable a systematic comparison of current and future methods aimed at exploiting interclass transfer, a designated benchmark must be available. This need is addressed by providing a benchmark of training and testing images of several hundred related object classes (mammals). By providing such a benchmark it will hopefully be possible to assess how well proposed solutions scale up and which better exploit interclass transfer.

2 Benchmark Generation

As stated above, the contribution of selecting an interclass transfer approach increases with the degree of common structure shared by the target object classes. However it is unclear how might one assess a-priori whether certain classes do, or do not, share common structure. Clearly, if an interclass transfer method contributes significantly to the recognition accuracy it could be assumed that common structure exists. Yet, a benchmark database must provide an a-priori justification for that belief. In [5] it was claimed that recognition of multiple characters within a certain writing system should exhibit interclass transfer. This claim was justified by the reusable patterns within writing systems described by [3]. The selection of mammals is analogously justified by two reasons. First, the phylogenetic origin of mammals entails a hierarchy of common structures. Second, even genetically distant mammals often share common physical structures due to evolutionary convergence caused by similar survival constraints. For example, the Aardvark and Hare have a similar ear structure, despite their genetic distance. These a-priori reasons justify the selection of mammal images as a suitable benchmark for interclass transfer object recognition methods. In addition, images of many mammals (like the Nabarlek) are rare, thus naturally necessitating interclass transfer techniques. It should be noted that while common structures exist, the variability of the object classes is immense, thus progress in this benchmark is a truly challenging task.

2.1 Image Acquisition

The compiled list of mammals is available in Appendix I. Since the benchmark goal is perceptual recognition, the provided list is by no means Zoologically verified and is known to be deficient on many bat and rodent species. Images were downloaded utilizing all URLs provided by the Google Images search engine and then converted to JPEG format.

2.2 Image Labeling

Images were manually filtered in a five tier paradigm:

1. irrelevant images including semantic noise (e.g. the fiction figure Zorro rather than the desert fox Zorro), duplicate images, non-typical versions of an animal (e.g. a baby Tapir) etc.
2. images of the animal that are not color photographs: including statues, paintings, sketches, computer generated animations and black and white images.
3. color images including a cropped version of the animal (e.g. only the head appears)
4. color images including the full animal in a non-standard pose or view
5. color images including the full animal in a standard pose

Example images of these five tiers are provided in Fig. 1. Naturally, all the images in the database are associated with the appropriate mammal name, however, the images of the fifth tier (Fig. 2) were further labeled by:

1. upright vs. four-legged pose
2. left profile, right profile or frontal view
3. bounding box parameters (top left corner coordinates, width and hight in pixels)

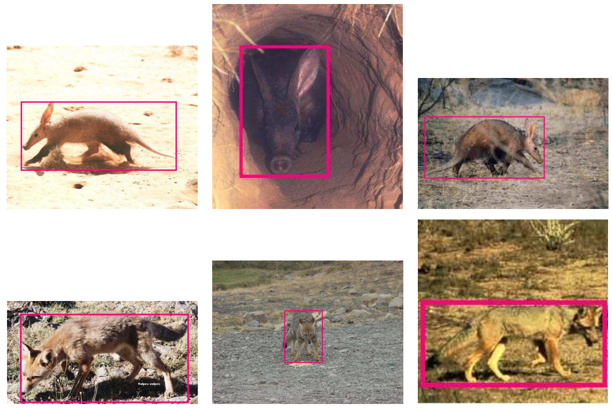

Figure 2: Examples of the mammals Aardvark and Zorro labeled with pose and bounding box

3 Database Description

A total of 272,486 images (˜7GB) have been downloaded from the web on March, 24, 2005. Labeling all images of a single animal demands approximately 30 minutes, thus a total of 300 labeling hours were required for the entire mammal database. This task was performed by a group of 15 undergraduate students who were paid for labeling the images. It should be noted that approximately 75% of the downloaded images were filtered as irrelevant due to semantic noise, duplicate appearances, non-typical versions of an animal etc. A histogram of the remaining four tiers is provided in Fig. 3. Assuming that additional information might be provided in the future (including pose parameters, pixel annotations etc.) labels are provided in XML format (in addition to binary, ASCII and matlab files). The benchmark and labeling applications are freely provided in order to emphasize an *open ended* effort, in which additional labels or class families (vehicles, trees etc.) could be incorporated in the future. The database and a web interface for viewing and introducing additional labeled images could be found at www.cs.huji.ac.il/˜fink/mammals/.

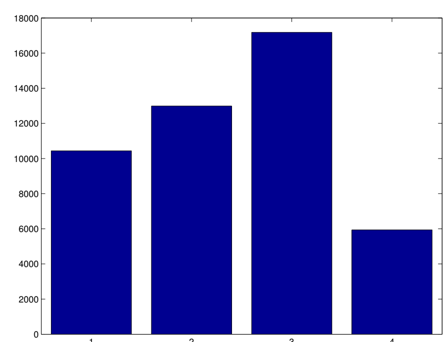

Figure 3: Histogram of the four category tiers: non-color-picture, cropped-image, non-standard-pose and standard-pose

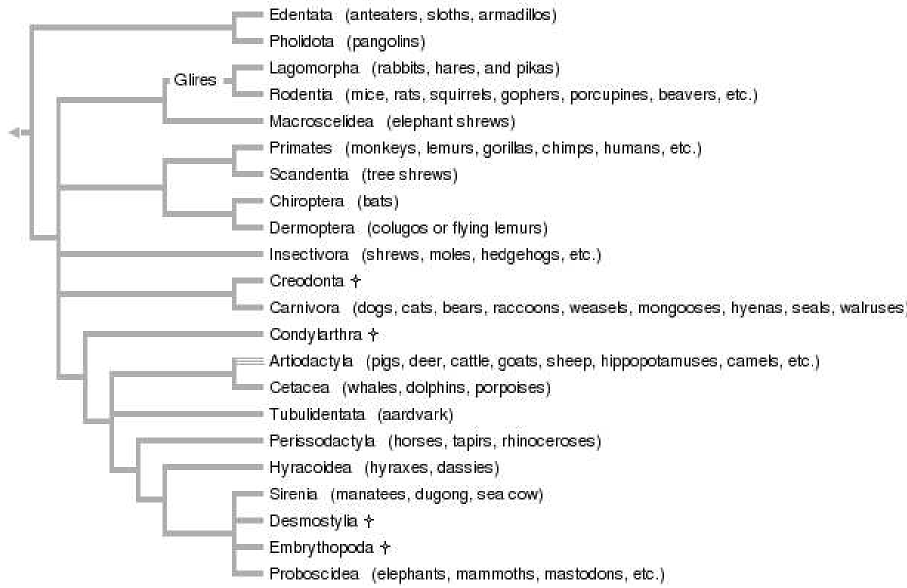

Figure 4: The Mammalia include Monotremata (Platypus, Echidnas), Marsupialia (Oppossums, Kangaroos etc.) and the variety of Eutheria depicted here in their phylogenetic tree structure (adapted from [16]). Recognition of the 400 non-extinct terrestrial mammals represented in the database, could be explored in context of their underlying common structures.

4 Research Benefits of the benchmark

The availability of a large labeled database containing multiple related classes offers distinct advantages for the study of several basic aspects of object recognition. As discussed above, one basic issue is the use of interclass transfer in the learning of new classes. A second issue is learning to deal with multiple classes with different levels of similarity. The recognition of highly similar classes that differ in only small details may require additional mechanisms compared with the recognition of more distant classes, and it will be of interest to test systematically the capacity of different methods to deal with different degrees of interclass similarity. Another issue is dealing with deformations that are specific to some classes but not to others. For example, in the mammals database, some classes are characterized by specific deformations due to typical pose variations (e.g., grazers' heads appearing either up or down). Mammal classes are organized in rich structures like regions of habitant, environmental role (e.g. grazers or hunters) and phylogenetic trees (Fig. 4). It will be of interest to study the relationships between these categories and visually-based classification.

Another general issue is the need for multi-channel processing and the discovery of the useful dimensions for different classification tasks. For example, the relative role of shape, color, and texture depends on the classes to be recognized. Occasionally, a particular texture may be helpful in recognizing a particular mammal class, yet often different classes share similar texture patterns (Fig. 5). Finally, since the mammal images were stored uncropped, they can be used to further investigate the contributing factors of image context. These issues are interesting as criteria for assessing the strengths and weaknesses of recognition approaches. In addition to overall performance measures such as ROC curves, it will be of interest to assess more specifically different criteria, such as the ability to deal successfully with highly similar classes.

Figure 5: The need for characterizing common multi-channel information is demonstrated by the prevalence of shared textures, like the stripe patterns of Amur Tigers, Aardwolves and Banded Duikers.

5 Appendix I

References

[1] S. Belongie, J. Malik and J. Puzicha. Matching Shapes. *ICCV*, 2001.

[2] S. Ben-David and R. Schuller. Exploiting Task Relatedness for Multitask Learning. *COLT*, 2003.

[3] M.A. Chanagizi and S. Shimojo. Complexity and redundancy of writing systems, and implications for letter perception. *under review*, 2004.

[4] L. Fei-Fei, R. Fergus and P. Perona. Learning generative visual models from few training examples. *CVPR*, 2004.

[5] M. Fink. Object classification from a single example utilizing class relevance metrics. *NIPS*, 2004.

[6] M. Fink and S. Shalev-Shwartz. Online object-recognition. submitted to *ICCV*, 2005.

[7] S. Krempp, D. Geman and Y. Amit. Sequential learning of reusable parts for object detection. *Technical report, CS Johns Hopkins*, 2002.

[8] S. Lazebnik, C. Schmid and J. Ponce. Affine-Invariant Local Descriptors and Neighborhood Statistics for Texture Recognition. *ICCV*, 2003.

[9] K. Levi and M. Fink. Learning From a Small Number of Training Examples by Exploiting Object Categories. *LCVPR04 workshop on Learning in Computer Vision*, 2004.

[10] D. Lowe. Distinctive image features from scale-invariant keypoints. *IJCV*, 60(2), pp. 91-110, 2004.

[11] E. Miller, N. Matsakis and P. Viola. Learning from One Example through Shared Densities on Transforms. *CVPR*, 2000.

[12] B. Schiele and J. Crowley. Recognition without Correspondence using Multidimensional Receptive Field Histograms. *IJCV*, 36(1), pp. 31-50, 2000.

[13] H. Schneiderman and T. Kanade. A statistical model for 3D object detection applied to faces and cars. *CVPR*, 2000.

[14] A. Torralba, K. Murphy and W. Freeman. Sharing features: efficient boosting procedures for multiclass object detection. *CVPR*, 2004.

[15] S. Thrun and L. Pratt. Learning to Learn. *Kluwer Academic Publishers*, 1997.

[16] Tree of life project. *http://tolweb.org/*, 1995.

Aardvark	Aardwolf	Addax	Addra
African Elephant	African Lion	African Wild Cat	African Wild Dog
Agile Antechinus	Agile Wallaby	Agouti	Agrant Shrew
Alaska Vole	Allied Rock-Wallaby	Alpaca	American Badger
American Bison	American Mink	American Porcupine	Ampurta
Amur Tiger	Andrew's Beaked Whale	Anteater	Antelope
Ape	Arctic Fox	Arctic Ground Squirrel	Arctic Hare
Arctic Wolf	Armadillo	Arnoux's Beaked Whale	Arsinoitherium
Artiodactyls	Asian Elephant	Asian Lion	Asiatic Black Bear
Atherton Antechinus	Atlantic White sided Dolphin	Aye-Aye	Babirusa
Baboon	Bactrian Camel	Badger	Bahamonde's Beaked Whale
Baiji	Baleen Whale	Banded Duiker	Banded Hare-Wallaby
Banded Mongoose	Bandicoot	Bat	Bay Duiker
Bearded seal	Beaver	Beluga Whale	Bengal tiger
Bennett's Tree-Kangaroo	Big Brown Bat	Bighorn Sheep	Bilby
Binturong	Bison	Black Bear	Black Bear Hamster
Black Footed Rock Wallaby	Black Lemur	Black Rat	Black Tailed Jack Rabbit
Black Wallaroo	Blackbuck	Black-footed Ferret	Black-striped Wallaby
Black-tailed Jack Rabbit	Blue Whale	Bobcat	Bongo
Boto	Bottle-nosed Dolphin	Bowhead Whale	Brazilian Free-Tailed Bat
Bridled nailtail wallaby	Broad Faced Potoroo	Brown Antechinus	Brown Bear
Brown Rat	Brush Tailed Rock Wallaby	Brush-tailed Bettong	Brush-tailed Phascogale
Bryde's Whale	Buffalo	Burmeister's Porpoise	Burmese Cat
Burrowing Bettong	Bush Dog	Bush Pig	Bushy-tailed Woodrat
Butler's Dunnart	California Myotis	California Sea Lion	Camel
Canyon Mouse	Cape Buffalo	Cape Hunting Dog	Cape York Rock Wallaby
Capuchin	Capybara	Caracal	Caribou
Carpentarian Pseudantechinus	Cat	Chamois	Cheetah
Chestnut Dunnart	Chevrotain	Chilean Dolphin	Chimpanzee
Chinchilla	Chipmunk	Cinnamon Antechinus	Civet
Cliff Chipmunk	Clouded Leopard	Clymene Dolphin	Coati
Collared Lemming	Collared Peccary	Commerson's Dolphin	Common Brushtail Possum
Common Duiker	Common Dunnart	Common Planigale	Common Ringtail Possum
Common Spotted Cuscus	Common Wallaroo	Common Wombat	Cougar
Cow	Coyote	Coypu	Crabeater Seal
Cuvier's Beaked Whale	Daintree River Ringtail Possum	Dall Sheep	Dark Kangaroo Mouse
Deer	Deer Mouse	Dense-beaked Whale	Desert Cottontail
Desert rat-kangaroo	Desert Woodrat	Dhole	Dingo
Dog	Dolphin	Donkey	Dromedary Camel
Duck-billed Platypus	Dugong	Dusky Antechinus	Dusky Dolphin
Eastern Barred Bandicoot	Eastern Chipmunk	Eastern Cougar	Eastern Grey Kangaroo
Eastern Grey Squirrel	Eastern Mole	Eastern Pipistrelle	Eastern Pygmy Possum
Eastern Quoll	Eastern Small-footed Bat	Eastern Tarsier	Echidna
Egyptian Mongoose	Ekaltadeta	Eland	Elephant
Elephant seal	Elk	Ermine	Eurasian Otter
European Hare	European Hedgehog	European Mole	European Rabbot
Evening Bat	Fallow Deer	False Killer Whale	Fanaloka
Fat-tailed Dunnart	Fat-Tailed Pseudantechinus	Fawn Antechinus	Feathertail Glider
Fennec Fox	Feral pig	Ferret	Fin Whale
Finless Porpoise	Fisher	Florida Manatee	Florida Mastiff Bat
Flying Squirrel	Fossa	Fox	Franciscana
Fraser's Dolphin	Free-Tailed Bat	Fruit Bat	Ganges River Dolphin
Gaur	Gazelle	Gemsbok	Genet
Gerbil	Gerenuk	Gervais' Beaked Whale	Giant Anteater
Giant Armadillo	Giant Otter	Giant Panda	Gibbon
Gilbert's Dunnart	Giles' Planigale	Ginkgo-toothed Beaked Whale	Giraffe
Gnu	Goat	Godman's Rock Wallaby	Golden Bandicoot
Golden Lion Tamarin	Golden-manteled Ground Squirrel	Goose-beaked Whale	Gopher
Goral	Gorilla	Grant's Gazelle	Gray Fox
Gray seal	Gray Whale	Gray's Beaked Whale	Great Basin Kangaroo Rat
Great Basin Pocket Mouse	Greater Glider	Green Ringtail Possum	Grey Bellied Dunnart
Grey Whale	Grison	Grizzly Bear	Groundhog
Guadalupe fur seal	Guanaco	Hairy Footed Dunnart	Hairy-tailed Mole
Hamsters	Harbor Porpoise	Harbor seal	Hare
Harp Seal	Hartebeest	Heaviside's Dolphin	Hector's Beaked Whale
Hector's Dolphin	Hedgehog	Herbert River Ringtail Possum	Herbert's Rock Wallaby
Hippopotamus	Hoary Bat	Hoary Marmot	Hog Badger
Honey Badger	Honey Possum	Hooded Seal	Horse
Hourglass Dolphin	House Cat	House Mouse	Howler Monkey
Hubbs' Beaked Whale	Humpback Whale	Hyena	Ibex
Impala	Indian Rhinoceros	Indiana Bat	Indo-Pacific Humpbacked Dolphin
Indus River Dolphin	Irrawaddy Dolphin	Jaguar	Jaguarundi
Javelina	Julia Creek Dunnart	Kakadu Dunnart	Kangaroo
Kangaroo Island Dunnart	Kangaroo Rat	Karakul	Killer Whale
Kinkajou	Kirk's Dik-dik	Kit fox	Klipspringer
Koala	Kowari	Kudu	Kultarr

Figure 6: List of mammals A-K represented in the benchmark database

Leadbeater's Possum	Least Chipmunk	Least Weasel	Lechwe
Lemming	Lemur	Lemuroid Ringtail Possum	Leopard
Lesser Hairy Footed Dunnart	Linsang	Lion	Little Pocket Mouse
Little Pygmy Possum	Little-Long tailed Dunnart	Llama	Long Footed Potoroo
Long Nosed Potoroo	Long-beaked Common Dolphin	Long-eared Myotis	Long-Finned Pilot Whale
Longman's Beaked Whale	Long-tailed Dunnart	Long-tailed Planigale	Long-tailed Pocket Mouse
Long-tailed Pygmy Possum	Long-tailed Vole	Long-tailed weasel	Loris
Lumholtz's Tree-kangaroo	Lynx	Manatee	Maned Wolf
Manul	Markhor	Marmot	Marsupial Mole
Marten	Meadow Vole	Mediterranean Monk	Meerkat
Melon-Headed Whale	Merriam's Shrew	Mink	Minke Whale
Mole	Mongoose	Monjon	Monkey
Montane Shrew	Montane Vole	Moose	Mouflon
Mountain Brushtail Possum	Mountain Goat	Mountain Gorilla	Mountain Hare
Mountain Lion	Mountain Pygmy Possum	Mouse	Mouse Deer
Mule	Mule Deer	Mulgara	Musk Ox
Muskrat	Musky Rat-kangaroo	Mustang	Nabarlek
Naked Mole-rat	Narbalek	Narwhal	New England Cottontail
Nine Banded Armadillo	Ningbing Pseudantechinus	Norrow-nosed Planigale	North Atlantic Beaked Whale
Norther Dibbler	Northern Atl. Bottle-nosed Whale	Northern Bettong	Northern Black Right Whale
Northern Brown Bandicoot	Northern Elephant Seal	Northern Flying Squirrel	Northern Fur Seal
Northern Grasshopper Mouse	Northern Hairy-nosed Wombat	Northern Nailtail Wallaby	Northern Quoll
Northern Rightwhale Dolphin	Northern Short-tailed shrew	Northern Yellow Bat	Norway Rat
Numbat	Nutall Cottontail	Nutria	Ocelot
Okapi	Old World Badger	Ooldea Dunnart	Opossum
Orangutan	Orca	Ord Kangaroo Rat	Oribi
Oryx	Otter	Pacific Water Shrew	Pacific White-Sided Dolphin
Pallid Bat	Palm Civet	Panda	Pangolin
Panther	Pantropical Spotted Dolphin	Paraguayan Fox	Parma Wallaby
Peale's Dolphin	Peary Caribou	Persian Cat	Pig
Pika	Pilbara Ningaui	Pine Marten	Pinyon Mouse
Platypus	Polar Bear	Polecat	Porcupine
Prairie dog	Preble's Shrew	Pronghorn	Pronghorn Antelope
Proserpine Rock-Wallaby	Przewalski's Horse	Puma	Pygmy Cottontail
Pygmy Hippopotamus	Pygmy Sperm Whale	Quokka	Quoll
Rabbit	Raccoon	Raccoon-Dog	Rat
Red Bat	Red Bellied Pademelon	Red Fox	Red Kangaroo
Red Legged Pademelon	Red Necked Pademelon	Red Panda	Red Squirrel
Red Wolf	Red-cheeked Dunnart	Red-Necked Wallaby	Red-tailed Phascogale
Reedbuck	Reindeer	Rhinoceros	Ribbon seal
Richardson Ground	Right Whale	Ringtail Cat	Ringtail Possum
Ring-tailed Lemur	Risso's Dolphin	River Otter	Rock Ringtail Possum
Rock Squirrel	Roe Deer	Rough-toothed Dolphin	Royal Antelope
Rufous Bettong	Rufous Hare Wallaby	Rufous Spiny Bandicoot	Sable Antelope
Saddle-backed Dolphin	Sagebrush Vole	Saiga Antelope	Salt's Dik-dik
Sandhill Dunnart	Scaly-tailed Possum	Scrub Hare	Sea Otter
Seal	Sei Whale	Seminole Bat	Serow
Serval	Sheep	Shepherd's Beaked Whale	Short Earned Rock Wallaby
Short-Finned Pilot Whale	Siamang	Siberian Tiger	Sika Deer
Simian Jackal	Siver-naired Bat	Skunk	Sloth
Sloth Bear	Smoky Shrew	Snow Leopard	Snowshoe Hare
Southeastern Bat	Southern Bog Lemming	Southern Brown Bandicoot	Southern Common Cuscus
Southern Dibbler	Southern Flying Squirrel	Southern Hairy-nosed Wombat	Southern Ningaui
Southern Pocket Gopher	Sowerby's Beaked Whale	Spectacled Hare-Wallaby	Spectacled Porpoise
Sperm Whale	Spider Monkey	Spinner Dolphin	Spotted Hyena
Spotted Skunk	Spotted-tailed Quoll	Springbok	Squirrel
Squirrel Glider	Squirrel Monkey	Star-nosed Mole	Steenbok
Stejneger's Beaked Whale	Steller Sea Lion	Straptoothed Whale	Striped Dolphin
Striped Possum	Striped Skunk	Stripe-faced Dunnart	Sugar Glider
Sun Bear	Swamp Antechinus	Swamp Wallaby	Takin
Tammar Wallaby	Tapir	Tarsier	Tasmanian Bettong
Tasmanian Devil	Tasmanian Tiger	Tetra	Thomson's Gazelle
Three-toed Sloth	Thylacine	Tiger	Topi
Townsend's Ground Squirrel	Trowbridge's Shrew	True's Beaked Whale	Tucuxi
Tundra Hare	Tundra Red-back vole	Twilight Bats	Uinta Chipmunk
Unadorned Rock Wallaby	Vagrant Shrew	Vampire Bat	Vancouver Island
Vaquita	Vicuna	Virginia Opossum	Wallaby
Walrus	Warthog	Water Shrew	Waterbuck
Weasel	Weddell Seal	Wester Quoll	Western Barred Bandicoot
Western Brush Wallaby	Western Grey Kangaroo	Western Havest Mouse	Western Jumping Mouse
Western Pygmy Possum	Western Ringtail Possum	Western small-footed Myotis	Whale
Whiptail Wallaby	White Rhinoceros	White Tailed Antelope	White Whale
White-footed Dunnart	White-footed Mouse	White-tailed Deer	White-tailed Dunnart
Wild Ass	Wild Boar	Wild dog	Wild horse
Wild Yak	Wildebeest	Wolf	Wolverine
Wombat	Wongai Ningaui	Woodland Caribou	Woodland Vole
Woolleys' Pseudantechinus	Yak	Yellow Footed Rock Wallaby	Yellow-bellied glider
Yellow-bellied Marmot	Yellow-footed Antechinus	Zebra	Zorilla
Zorro			

Figure 7: List of mammals L-Z represented in the benchmark database