# Assignment 3

DS405B - Practical Deep Learning from Visual Data

Dr. Aseem Behl
School of Business and Economics
University of Tübingen
Monday, 4. July 2022

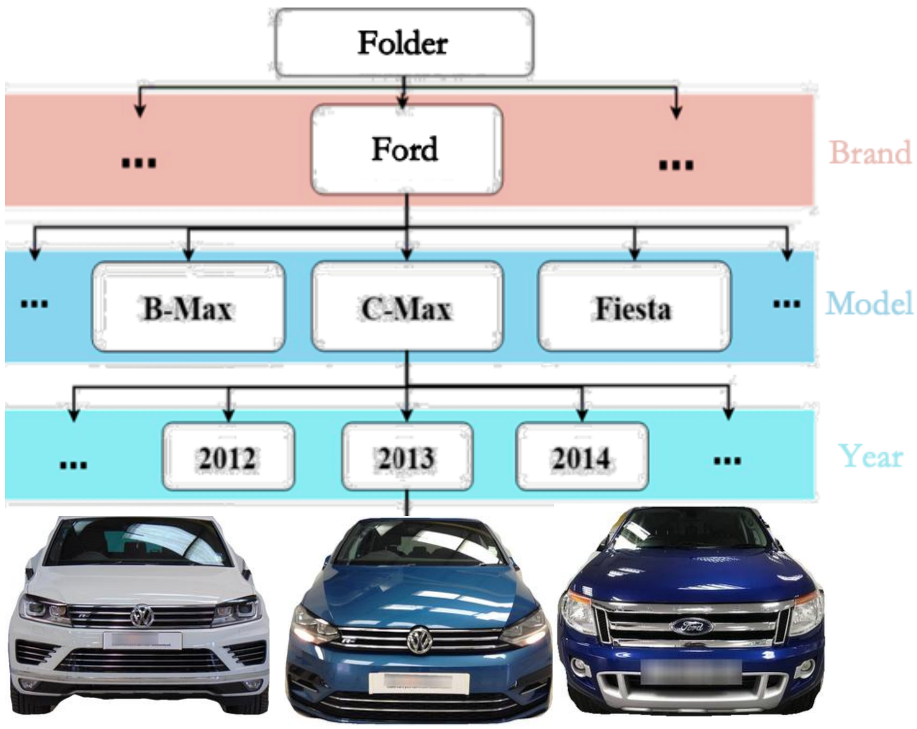
# Introduction

The goal of this assignment is to learn how to employ transfer learning with Convolutional Neural Networks for downstream business and economics applications. Specifically, in this assignment, we are interested in applying deep learning models to extract visual attributes from car images to improve our understanding of automotive exterior aesthetic design. While not in the scope of this assignment, marketing practitioners can further exploit these visual attributes to understand consumer preferences and model sales.

# Data

We will use the *DVM-CAR dataset (A Large-Scale Automotive Dataset for Visual Marketing Research and Applications)* for this task.

Download the image dataset here. Each image in the dataset is named with the following format `Brand_Name$$Model_Name$$Launch_Year$$Color$$Model_ID$$Advertiser_ID$$Image_ID.jpg`. Following figure illustrates the directory structure of the DVM-CAR dataset.



# Automotive Design Modernity

| **Defender 110**<br>Modernity score: 3.0 | **Range Rover**<br>Modernity score: 3.4 | **Discovery 4**<br>Modernity score: 4.1 | **Discovery Sport**<br>Modernity score: 5.4 |
|---|---|---|---|

The first visual characteristic we will learn to infer from images is Automotive Design Modernity. In order to predict design modernity, we will employ ResNet18 network pre-tained on ImageNet dataset. We then use transfer learning adapt our network to the task classification of car images to production year categories. For example, we can specify year ranges 2000-2003, 2006-2008, 2009-2011, 2012-2014 and 2015-2017 to be production year categories 0, 1, 2, 3 and 4 respectively. The design modernity score for a car model image can then be defined as a weighted sum of the the production year category labels (weighted by the respective predicted probabilities for that category). Feel free to adapt the categories to finer or coarser ranges based on your observations. The above figure (source: *DVM-CAR dataset*) shows an example of how Land Rover models which were sold in the same year can have different modernity score, indicating difference in design language.

To complete this task, you will have to carry out the following steps:

1. Split the dataset into training, test and validation dataset containing 70%, 20% and 10% of the examples respectively. Make sure while splitting the dataset that the images of a car model (specified uniquely by model_id and model_year) should not be split across train, test and validation splits.

2. Implement data loaders in PyTorch for the training, validation and test dataset. The images in our automotive dataset are of different dimensions than ImageNet images the network is pre-trained on. Therefore, they should be resized to a fixed standard resolution (`224x224`). Feel free to use the `RandomResizedCrop` or other random transformations provided in `torchvision.transforms` for the training dataset. For the validation and test dataset you are not allowed to use random transformations. As shown in tutorial notebook 9, do not forget to normalise your datasets using the mean and standard deviation computed for ImageNet dataset.

3. You will employ transfer learning with a ResNet-18 network pre-trained on ImageNet dataset for classifying car images. Review the section on Transfer Learning from lecture notes by Andrej Karpathy, and pick one of the following two approaches for transfer learning which based on what we have discussed in the lecture may be more suitable for this task.
   A. **ResNet-18 as fixed feature extractor** - Take the ResNet-18 pre-trained on ImageNet. Therefore, we can treat the ResNet-18 as a fixed feature extractor for the automotive image dataset. The last layer of ResNet-18 outputs are the 1000 class scores for ImageNet dataset. However, in our task we have a different number of categories. Therefore, you should replace the last fully connected layer in ResNet-18 with a new one with random weights and train only the parameters of this new layer. Select the hyper-parameters to maximise the performance on the validation dataset.
   B. **Fine-tuning the ResNet-18** - Another way to exploit transfer learning is to not only replace the last layer, but to also fine-tune the weights of the pretrained network. Therefore, in this setting, you should replace the last fully connected layer in ResNet-18 with a new one with random weights and train the parameters of this new layer as well as the pre-

trained part. Next, you should again for this setting, select the hyper-parameters that maximise the performance on the validation dataset.

4. Evaluate and report the performance of the best performing model from step 3 on the test set.

5. Randomly select images for 4 car models each from the test set for the years 2000, 2006, 2010, 2014, 2017. Visualise the images along with their design modernity score.

## Automotive Design Typicality

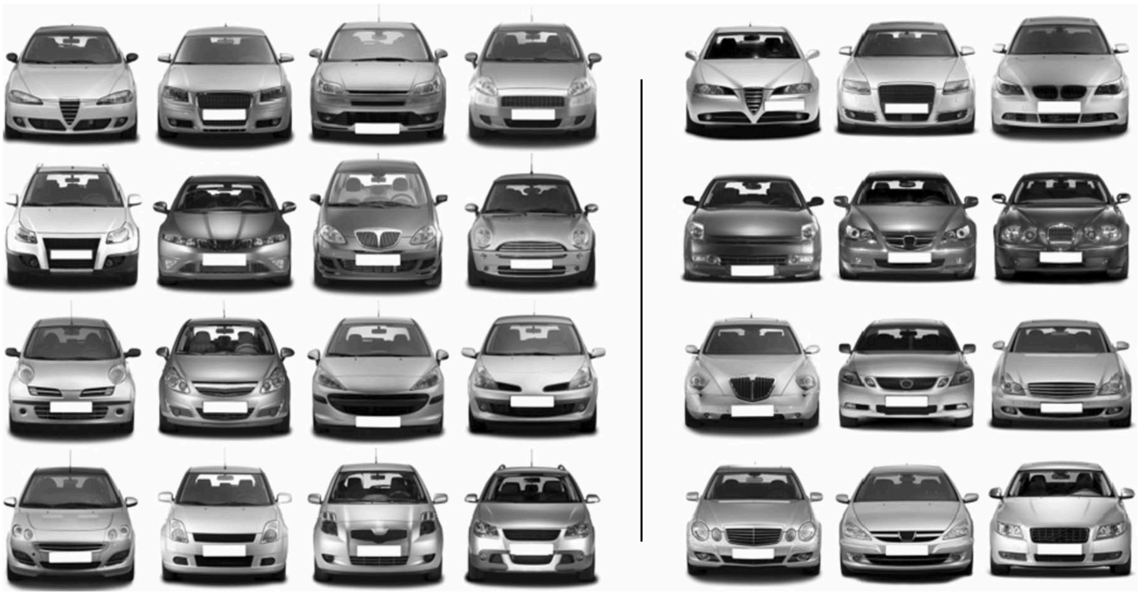


Image Source: Landwehr 20122. Gut Liking for the Ordinary: Incorporating Design Fluency Improves Automobile Sales Forecasts

The second visual characteristic we will learn to infer from images is Automotive Design Typicality. Design typicality describes how representative a certain object is of a group (Blijlevens et al., 2012). Previous research has shown that consumers prefer typical over atypical designs (Hekkert et al., 2003).

We can leverage neural networks to automatically compute typicality ratings in contrast to manual approaches (e.g., Landwehr et al., 2011). Specifically, we compute the design typicality for a car model as the cosine similarity between the visual product features and features of the morph (Lake et al., 2015). Here, the morph of a group is a representative product of that group computed by averaging the feature vectors across all products in the group. To complete this task, you will have to carry out the following steps:

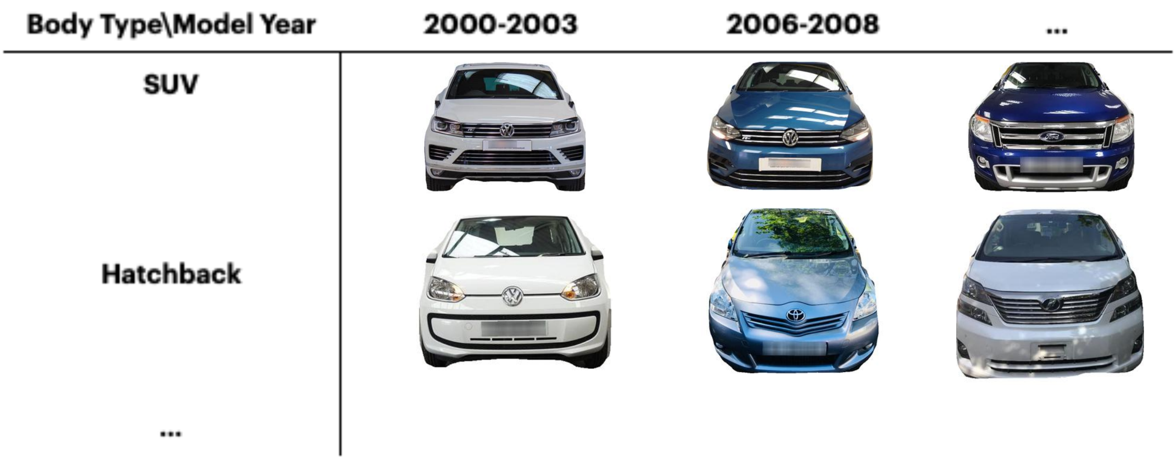
1. To compute the product features, we extract the features from a second ResNet18 CNN classifier. The second ResNet18 network is also first pre-trained on ImageNet. The pre-trained network is next transfer learned (refer to the last section for instructions) to the task of classifying car images into body-type categories: 'Hatchback', 'SUV', 'MPV', 'Saloon' and 'Convertible'. The labels for body-type are available in the the autos.csv in assignment_3 directory on ILIAS.

2. After transfer learning on the task of body-type classification, the features output from the network layer prior to the last custom linear layer give us the product features. For ResNet18, this corresponds to the output of the last `avgpool` layer.

3. Next, we can compute the morph for each group, where a group is defined by a combination of body-type and model-year category. For example, SUVs for model year 2012-2014 can be considered one group. For each group, you must compute the morph by computing the average of the product features extracted from the ResNet after transfer learning for all products in that group.

4. For each group, compute the cosine similarity between the morph of that group and all other auto products images in that group. This similarity measure gives us the design typicality for each automotive.

5. For each group, visualise the automotive image with the highest typicality score as shown below.




High typicality score means that the product is closest to the morph of that group and thus can be considered a representative design for the group. Therefore, this visualisation will allow us to see the evolution of product designs over the years for different automotive segments.

6. Finally select five groups of your choice. For each group visualise an automative product image with high, low and medium typicality.

<u>References:</u>


Blijlevens, J., Carbon, C.-C., Mugge, R., & Schoormans, J. P. L. (2012). Aesthetic appraisal of product designs: Independent effects of typicality and arousal.

Hekkert, P., Snelders, D., & Wieringen, P. C. W. V. (2003). 'Most advanced, yet acceptable': Typicality and novelty as joint predictors of aesthetic preference in industrial design.

Landwehr, J. R., Labroo, A. A., & Herrmann, A. (2011). Gut liking for the ordinary: Incorporating design fluency improves automobile sales forecasts.

Lake, B. M., Zaremba, W., Fergus, R., & Gureckis, T. M. (2015). Deep neural networks predict category typicality ratings for images.

Keep in mind the computational resource constraints on Google Colab or the Uni-Cluster and setup your experiments judiciously.  Starting with the assignment early on is also a good strategy in order to maximise GPU time.

Along with the code, please provide a short analysis of the results in the notebook using text boxes. **Run all cells, and do not clear out the outputs, before submitting.** Download notebook to your computer and upload it as submission to ILIAS.
This assignment is worth **20 points** and is due on **Friday, July 15, at 17:00 AM CET**. In order to receive credit for the assignment, please be prepared to present your solutions to randomly selected sections from the assignment or similar exercises in the lecture on **Wednesday, July 20, 2022**.