

# Data Pipeline for Images (Jenkins Automation) Computer Vision

Extract

Transform

Data Gathering

Raw Data

Python Script  
to transform image

Processed Data

Store raw data on Data Instance

- Grayscale
- Resize (shape and size)
- Clean

Python script for automated  
checking and testing

- Delete corrupted images
- Rename
- Delete duplicates
- Delete any images that are  
not in the specified shape  
or size

Load

Store

Feed to model

Pipeline ends here,  
Image augmentation  
in model training