# Introduction to Visual AI for Cultural Heritage

[PRESENTATION SLIDES](https://drive.google.com/file/d/1DCMw4hrA7Kg8DnEE7Qc-QyMTBCOpRMus/view?usp=sharing)

This worksheet introduces a number of tools addressing common tasks in computer vision. The tools have been developed by the University of Oxford’s [Visual Geometry Group](http://www.robots.ox.ac.uk/~vgg/) in collaboration with digital humanities and other researchers.

**How to follow this workshop**

Exercises may be followed through **online demos** or, optionally, through **local installation** of the tools on your own laptop. The online demos are **recommended** in the first instance. **Demo data** is needed for some of the exercises, and is provided below.

**Requirements for all**

A laptop (not a tablet).

All the tools require a modern web browser such as Firefox, Chrome, Edge or Safari. Internet Explorer is not supported.

**Requirements for local installation - s**ee [this guide](https://docs.google.com/document/d/1QrRIFnPWFNIZ0-fpHG10K-VY0FwDOxLKUc_c02LcIw0/edit?usp=sharing)

# Task 1. Matching images with VISE

**Overview** VGG Image Search Engine (VISE) provides *instance-based* image-matching. It is designed to match identical images, and can also often match the same regions within images.

### Online demos

Try at least **one** of the following three demos

1. Bodleian Ballads [ImageMatch](http://imagematch.bodleian.ox.ac.uk:8000/) (there is an alternate server [here](http://zeus.robots.ox.ac.uk/ballads/) if the former is slow)

This implementation of VISE indexes around 900 seventeenth-century English printed broadside ballads held by the Bodleian Library.

**How to:** Select a ballad from one of the thumbnails in the page at the link above. Click and drag a bounding-box around a region of interest, and then press ‘SEARCH.’

**Try:** Selecting ‘Detailed matches’ to see a match and its context. From here, you can tick the boxes ‘Regions’ and ‘Lines’ to see the local features within the image that the tool is matching.

**Try:** Selecting ‘Image Comparison’ to see a view of the two matching woodcuts (move your trackpad pointer across the images to flip between them)

**Try:** Finding the point where VISE returns an irrelevant result. Select [See list view](http://imagematch.bodleian.ox.ac.uk:8000/dosearch?docID=774&xl=545.00&xu=688.00&yl=73.00&yu=216.00&numberToReturn=20&startFrom=1&tile=false&) and observe how the score drops as you go down. What is the difference in scores between a relevant and an irrelevant result? What factors might affect the scores that VISE generates from a given set of images?

**Try**: Uploading one of the images from the folder ‘Bodleian Ballads upload queries’ in the workshop folder (you’ll need to download the images first). Tick the box marked ‘whole sheet’ for the Bodleian ballad.

Further information

· See [Bodleian Ballads Online](http://ballads.bodleian.ox.ac.uk/) for a database incorporating VISE within a conventional digital library resource.

· A catalogue of the woodcuts divided into specific woodblocks, with their impressions clustered by similarity and sequenced over time, is described [here](http://balladsblog.bodleian.ox.ac.uk/blog/1069).

2. [Oxford Buildings](http://www.robots.ox.ac.uk/~vgg/research/oxbuildings/index.html)

 100,000 mostly random photographs, but including a number of shots of Oxford buildings viewed from various angles and distances and with various things going on next to them.

**Try** searching for features of different sizes in Oxford buildings (e.g. the lamp-post in [this image](http://zeus.robots.ox.ac.uk/oxfordbuildings/search/?id=oxc1_magdalen_000545)), as well as the large building [here](http://zeus.robots.ox.ac.uk/oxfordbuildings/search/?id=oxc1_radcliffe_camera_000012)).

**Try** thinking about how to create a training dataset such as Oxford Buildings (which is described [here](https://www.robots.ox.ac.uk/~vgg/data/oxbuildings/)). How would you go about obtaining such a dataset of public images, and how should it be documented? What kinds of legal or issues might be involved? See [here](https://arxiv.org/abs/1803.09010), [here](https://arxiv.org/abs/1810.03993) and [here](https://megapixels.cc/) for some discussion of these issues (nb the Oxford Town Centre dataset described in the last link is not one of ours..)

### VISE local demo

To set up VISE on your own laptop, follow the Installation Instructions [here](http://www.robots.ox.ac.uk/~vgg/software/vise).

To set up the VISE environment and create your first index, see the [User Guide.](https://www.robots.ox.ac.uk/~vgg/software/vise/_old_2016-2020/UserGuide.html)

You can download sample VISE local data (images of parts of books - printed illustrations, bindings and bookplates) [here](https://drive.google.com/file/d/1nIGqt8RxxZILbXMi5PAHItqsfylzjSrO/view?usp=sharing).

# Task 2. Comparing images with Image Comparator/Traherne

This tool was originally built in order to detect resettings of type in early printed books, and was named after the Oxford edition of the works of Thomas Traherne.

As it can also be used to help find small variations in pairs of otherwise identical images, it has since been developed as Image Comparator - access the demo [here](http://meru.robots.ox.ac.uk/imcomp/index.html). An alternate server is [here](http://zeus.robots.ox.ac.uk/imcomp/index.html).

**To Use:** Either drag and drop the sample pairs of images at the bottom of the screen, or download and reupload images from samples provided [here](https://drive.google.com/file/d/1CAsuICFMhgKqquVQJLodC1xqHYffv_JC/view?usp=sharing) (I recommend you download the whole folder). Press Compare, then select a visualisation.

**Try:** If you have skewed or warped images (for example, curved book pages) ImageComparator can apply a geometric transformation to align them. Several transformations are provided: try them and see what they do.

**Also Try:** Select ‘Photometric Transform’ to equalise the lighting of each image.

### Installable demo

Traherne may be downloaded for Mac and Windows from [here](http://www.robots.ox.ac.uk/~vgg/software/traherne/) (N.B. you may need to bypass Windows Defender or Mac Gatekeeper to install the software). Demo data is [here](https://drive.google.com/file/d/19a7vOII_wvTdNNQ9xwaw27XYsDUw5gJD/view?usp=sharing). To use Traherne, click Load Base and select the first of a pair of files, then do the same for Load Comp (more than one file can be selected in each operation). Click Compare Base & Comp: Traherne will then register the two images: visualisations can be selected, as with ImageComparator.

# Task 3. Classifying images with VIC

**Overview:** VIC provides the ability to detect and classify objects in images: it includes a neural network model trained on the large ImageNet dataset.

NB: Initial VIC results are often surprising. You can refine the search by selecting relevant matches and searching again, but results depend on the consistency of the training data and its labelling, as well as the methods and the purposes to which they are put.

Online demos

### 1.Villanova

### Please visit [this spreadsheet](https://docs.google.com/spreadsheets/d/1EO5BMjn2MbRRGnAIw6gG3bzqr1-GW_U2wtC-hJtW1rs/edit?usp=sharing) and add your initials to reserve a demo (each instance of the demo can only support a few users at once) and then click the link. All demos provide the same materials, from [**Villanova University Library Image’s Collection.**](https://digital.library.villanova.edu/Item/vudl%3A98245)

The demo provides a number of search modes.

**Try** Browsingvia the words ‘map,’ ‘portrait’, and ‘dog’ – these are examples of so-called ‘curated queries’ where VIC has been pre-trained on specific, labelled examples and is then retrieving similar results (you can create new curated queries if you install VIC locally).

Then, try the search interface:



From left to right, you can:

1. Enter a keyword for retrieving Google-tagged images. VIC will then attempt to find similar images in the Villanova dataset.
2. Upload an image from your laptop, or select the URL of an image on the internet, which VIC will then ingest.
3. Pull down a list of image metadata terms (selected by the Villanova librarians) which act as browse points for onward visual searches. If you select \* you will see the contents of the collection as a whole. **Try** clicking on ‘similar’ to see images that VIC believes contain some matching features (NB: note the green tick (positive) or yellow triangle (negative) badge on the results). You can also refine the results by adding several images – click on the plus sign of each.

**TRY** clicking on a result to see the full metadata for an image, including a link back to the item in Villanova’s catalogue.

### 2. [ArtUK](http://zeus.robots.ox.ac.uk/artsearch/)

A large collection of paintings in public ownership in the UK. It uses the same technology as VIC but with a slightly different interface.

**Try** searching by colour and texture.

### Installable demo

To install VIC on your own machine, follow the Installation Instructions [here](http://www.robots.ox.ac.uk/~vgg/software/vic/). VIC can ingest data either from a folder on your laptop, or via a IIIF manifest or collection. In either case you will need to create a local set of folders (which we provide in our data downloads) and map them to the relevant Docker folders, as shown in the [installation instructions](http://www.robots.ox.ac.uk/~vgg/software/vic/).

To use data from a IIIF repository, [follow these instructions](http://www.robots.ox.ac.uk/~vgg/software/vic/downloads/docker/using_your_own_IIIF_images.pdf). N.B: it’s advisable to download one of our sample datasets (such as Paintings, linked to from [here](http://www.robots.ox.ac.uk/~vgg/software/vic/)) in order to set up the folder structure that VIC expects/

The following IIIF collections have been tested (copy and paste the URL):

[Villanova University Digital Images Collection](https://digital.library.villanova.edu/Collection/vudl%3A98245/IIIF) (the same dataset as the online demo)

[Cambridge University Library Curious Objects](https://cudl.lib.cam.ac.uk/iiif/collection/curiousobjects) (57 images of book-like or related objects)

[Bodleian Entertainment Ephemera](https://iiif.bodleian.ox.ac.uk/iiif/collection/entertainment-ephemera) (659 images - this will take 10-15 minutes to ingest)

[Durham University’s Manifest browser](http://iiif.durham.ac.uk/jalava/universe.html#http%3A%2F%2Fryanfb.github.io%2Fiiif-universe%2Fiiif-universe.json/http%3A%2F%2Fdigital.library.villanova.edu%2FCollection%2Fvudl%3A3%2FIIIF/https%3A%2F%2Fdigital.library.villanova.edu%2FCollection%2Fvudl%3A2001%2FIIIF) is a good source of other manifests. **(N.B. manifests or collections can link to many thousands of images - VIC will attempt to ingest them all in one go. It’s therefore recommended that you verify the numbers of images in a IIIF repository before ingesting).**

Task 4: Recognising faces with VFF

**Overview**: Facial recognition, like image classification, is now generally done with deep neural networks. VFF includes a model trained on millions of faces representing thousands of identities, with which classifiers for matching individual faces to identities can be built on the fly.

Online demos

1. VGGFace using VFF software

### Please visit [this spreadsheet](https://docs.google.com/spreadsheets/d/1ckwATfy-Rh-7oIl74VoTq91e3qGUl1bOMLFwNZBXUjU/edit?usp=sharing) and add your initials to book an online demo (each demo can only support a few users at once). All demos provide the same materials, from the [VGG Face](https://www.robots.ox.ac.uk/~vgg/data/vgg_face/) dataset of public-domain images of celebrities.

As with VIC, VFF provides a number of search modes. You can submit a photo; enter a name for retrieving Google-tagged images; or you can pull down a list of names of pre-tagged faces, which you can then submit as queries to try to retrieve untagged faces. See ‘Getting Started’ at the bottom right of the screen for more detailed instructions.

**TIP:** To see the collection as a whole, pull down the metadata search terms, click in the box, select the \*, click Submit, then click the Search icon.

**Try**: Selecting a name for an initial search point, or Subset 1 or Subset 2 for a random selection of identities.

**NB**: facial recognition is extremely controversial, in particular for policing and security applications. Bias is a particular concern: see the [work](https://scholar.google.com/citations?user=lemnAcwAAAAJ&hl=en) of [these](https://scholar.google.com/citations?user=lemnAcwAAAAJ&hl=en) and other [researchers](http://proceedings.mlr.press/v81/buolamwini18a.html?mod=article_inline) for critical commentary. [VGG-Face](https://www.robots.ox.ac.uk/~vgg/data/vgg_face/) includes this caution: ‘we note that the distribution of identities in the VGG-Face dataset may not be representative of the global human population. Please be careful of unintended societal, gender, racial and other biases when training or deploying models trained on this data.’

1. VGG also provide a demo for [searching paintings for your nearest “neighbour”](http://www.robots.ox.ac.uk/~vgg/research/face_paint/)

Local demo

### VFF can be installed from [here](http://www.robots.ox.ac.uk/~vgg/software/vff/), as can a set of images of celebrities matching the online demo.

### If you would like to try VFF using your own images, see [instructions here](http://www.robots.ox.ac.uk/~vgg/software/vff/docker_deployment_mac.html#own_images).