### Introduction to Visual AI for Cultural Heritage

Les Futurs Fantastiques 2021

2/12/2021

Dr. Giles Bergel

Visual Geometry Group, Department of Engineering Science, University of Oxford

## Agenda

1. Presentation 40 mins

2. Independent tryouts 60 mins

3. Discussion 20 mins

4. After the workshop, feel free to get in touch!

### About me

POWERS OF THE PRESS



THE HOMEPAGE OF GILES BERGEL

ABOUT RESEARCH TEACHING CV NOTES CONTACT

#### About the author

I am an academic book historian and a digital humanist. I am equally interested in how texts came into existence in books, and in how digital culture is changing the nature of the book (and much else besides). My interests include ballads, broadsides, chapbooks, typefaces, printers' woodblocks, diagrams, copyright and book trade history. This website contains links to a number of my research projects, which are mostly digital in nature.

http://www.printing-machine.org

## My organisation



http://www.robots.ox.ac.uk/~vgg/

#### ♥Visual AI Home Themes News Publications Datasets Software People ✓ Events ✓



AN OPEN WORLD INTERPRETABLE VISUAL TRANSFORMER

A large research project based in the Department of Engineering Science, University of Oxford, in partnership with The University of Edinburgh and University of Bristol. It is funded by the EPSRC (Engineering and Physical Sciences Research Council), and will run from 2020 - 2025.



#### Objectives

To carry out fundamental research to develop next generation computer vision methods that are able to input audio-visual data (such as a video stream, or large scale datasets) and transform it into usable semantic and geometric information, with human-like capabilities for visual analysis, flexibility and interpretability

To transfer these methods to industry and to other academic disciplines (such as Archaeology, Art, Geology, Medicine, Plant sciences and Zoology)

### My role

"To carry out research applying visual AI in the humanities... To disseminate the research outputs to appropriate communities in the humanities and cultural heritage fields... and to communicate back research questions, ideas for new development, feature requests etc..."

## Why cultural heritage?

- Cultural heritage (or LAM) organisations have large datasets and good-quality metadata.
- Many humanities research questions have been difficult to address computationally, and are intrinsically interesting
- An academic research collaboration can be mutually beneficial and demonstrate impact. Impact may include new research findings, better metadata, improved collections management, refinements to methods and software... or entirely new research questions.

## AI is not just algorithms

- Algorithms (i.e. software)
- Computing resources (storage, hardware)
- Datasets (for training and testing)
- Curation (of training data and outputs)

### Evolution of VGG tools: from research software..



#### Descriptor Learning Using Convex Optimisation

Code and learnt models for feature descriptor learning, computation, and evaluation [Simonyan et al., PAMI 2014].



#### Detecting Upper Body Configurations

MATLAB code for detecting upper body configurations [Hoai and Zisserman, CVPR 2014]



#### Efficient Additive Kernels via the Homogeneous Kernel Map

Code for large scale training of non-linear SVMs based on additive kernels [Vedaldi and Zisserman, CPVR 2010]



#### Encoding Methods Evaluation Toolkit MATLAB code for evaluation of different bag of visual words encoding schemes over standard image classification

test datasets [Chatfield et al.,

BMVC 2011]



#### Efficient Structured Output SVM Ranking

The reference implementation of the structured output ranking algorithm with linear time constraint generation method [Mittal et al., ECCV 2012]



<u>Fast Semantic</u> <u>Segmentation via Soft</u> <u>Segments</u>

An implementation of the method described in: "Visual vocabulary with a semantic twist" [Arandjelovic and Zisserman, ACCV 2014]

### ... to releases for a wider public



#### **VGG Text Search**

A web-based engine for running onthe-fly text string searches over userdefined datasets.



#### Image Comparator

A browser-based image comparison tool to compare two images in near real-time. It has a range of options to visualize the comparison results.



#### VGG Face Finder

A web-based engine for running onthe-fly face searches over user-defined datasets.



# swan



An standalone image annotator application packaged as a single HTML file (< 200 KB) that runs on most modern web browsers.



#### VGG Image Search Engine

Standalone application able to make a large collection of images searchable by using image regions as query.

http://www.robots.ox.ac.uk/~vgg/software/

## Common computer vision tasks

- Comparison (spotting the differences between images)
- Matching (finding out what images have in common)
- Classification (describing what is in an image its content or subject)
- Facial recognition (combines some elements from the above)

## Common computer vision tasks

- <u>Comparison (spotting the differences between images)</u>
- Matching (finding out what images have in common)
- Classification (describing what is in an image its content or subject)
- Facial recognition (combines some elements from the above)

### Spot the difference...?

Enter Tamora, Andronicus, and Lucius.

Enter Tamora, Andronicses, and Lucius,

https://collation.folger.edu/2013/04/first-folios-online/

### Spot the difference (2)



Shakespeare First Folio Works (1623): Bodleian and Boston Public Library copies

### Spot the difference (2)



Shakespeare First Folio Works (1623): Bodleian and Boston Public Library copies

## Visualising variance with Traherne

Stop, ftop, ao helpe? Enter Glofter, and Sermants with Torches. Glo. Now Edward, where's the villaine? Baft. Here flood he in the dark, his fharpe Sword out, Mumbling of wicked charmes, conjuring the Moone To ftand auspicious Miffris. Glo. But where is he? Baft. Looke Sir, I bleed. Glo. Where is the villaine, Edmand? Esft. Fled this way Sir, when by no meanes he could. Glo. Pursue him, horgo after. By no meanes, what? Baft. Perswade me to the murther of your Lordship, Gafin

### Visual Geometry Group

Department of Engineering Science, University of Oxford

### **Traherne Digital Collator**

Abhishek Dutta, Joon Son Chung and Andrew Zisserman

#### Overview

Traherne Digital Collator is an image comparison tool developed for the <u>The Oxford Traherne</u> project. While the user interface is designed specifically to address the needs of Traherne project users, this tool should be useful for any standard digital collation task. It is an <u>open</u> <u>source project</u> developed at the <u>Visual Geometry Group</u> and released under the BSD-2 clause <u>license</u>.

<u>Version 1.3</u> (and older releases) of the original collator application was developed by <u>J. S. Chung</u> and more details about these releases are available <u>here</u>. We recommend using the newer version (i.e. 2.x.y) of this tool whose user interface has been redesigned and codebase rewritten to create a standalone and lightweight application.

#### **Downloads**

Windows	MacOS	Unix		
<ul> <li><u>Traherne-2.0.5-Win32</u> (for Windows 7 and 8.1)</li> <li><u>Traherne-2.0.5-Win64</u> (for Windows 10)</li> </ul>	• Traherne-2.0.5-MacOS.dmg	Compile from source		

Software - http://www.robots.ox.ac.uk/~vgg/software/traherne/

### Visual Geometry Group Department of Engineering Science, University of Oxford



#### Image Comparator (IMCOMP)

#### Shrinivasan Sankar, Abhishek Dutta, and Andrew Zisserman

#### Overview

Image Comparator (or, IMCOMP) is a web application to automatically compare a pair of images using geometric and photometric transformations. It is an open source project maintained by the Visual Geometry Group.

#### Here are some features of IMCOMP:

- · available as an online tool that can be accessed from any modern web browser
- · a large number of visualizations are available to help users spot the difference between two images
- · Supports photometric transformation to compensate for colour differences between two images
- · Supports different types of geometric transformations (e.g. similarity, affine, thin-plate spline, etc.) to enable comparison of images containing many types of deformations.
- · Results can be saved as an image.

#### **Online Application**

#### Click here to load the IMCOMP online tool

#### Screenshots



http://www.robots.ox.ac.uk/~vgg/software/imcomp/

## Maps and music





## Engraved illustrations



### **Comparing Curved Pages**



d on the base image, drag mouse cursor to select a region.

☑ Comp. < [1] Marston - What You Will, Austin TX, sig. c4r.jpg ♥ > ↔ ↑ ↑ ↑ Comp. (full image) ○ Comp. (cropped) ○ Overlap ○ Base (cropped) WHATTOV WIL. Enter a Schole-mailler Aravs the curtains behind with, Battus Nows, Slip, Nathaniell and Holifernes Pippo, Schole-boyer, fitting with booker in their head.

#### WHAT TOP WIL.

Sener a Schole-mailler, de aver the curtains behind with, Battus Nows, Slip, Nathaniell and Holifernes Pippo, febole-bører, fitting with bioket in their kunde.

All, Salue Mazifler. Ped Saluere pueri effote falui, vos faluere exopio vobis falurem, Batte my fili, fili sei Batte. Bat, Quid vie.

Ped Stand forth repeat your leffon with out booke. Eat. A nowne is the name of a thing that may be feene felt heard or vnderftood,

Pra,Good boy,on on, Bat. Of nownes fome bee fubftantiues and fome beefub-

fantiucs. Ped.Adiectines. Bat. Adjectimes, a nowne fubftantiue ether is propper to the thing that it betokneth.

Ped Well to numbers.

Barran, In Nownes bee two numbers, the Singuler and the Plurall, the Singuler number speaketh of one as Lapira Stone, the Plurall speaketh of more then one, as Lapides

Ped, Good childe , now thou art past Lupider Stones, proteed to the cafes Nons , fay you next Nows , wher's your leffon None.

Now, I am in a verbe forfooth.

- Pol Say on for footh fay fay, News, A verbe is a part of Ipeach declined with mood and innee and betokneth doing as Amo Houe, Ped How many kind of verbes at there?

Ness, 2 Perfonall and imperfonall, Ped Of verbs perfonalls, how many kinds.

New, Fine, Actine Paffine Neuter Deponent and Common, A Verbe Active endeth in O and beetokneth to doe as

Must love and by putting to R it may bee a pathice as Amer I

## Common computer vision tasks

- Comparison (spotting the differences between images)
- <u>Matching (finding out what images have in common)</u>
- Classification (describing what is in an image its content or subject)
- Facial recognition (combines some elements from the above)

### **Image Matching**

Are these images of the





instance

same object?

### **Image Classification**

What object is this?





category

### Image matching

DOWNLOAD DEMO NEWS COMMUNITY SUPPORT ABOUT

#### VGG Image Search Engine (VISE)

Abhishek Dutta, Relja Arandjelović, Andrew Zisserman

VGG Image Search Engine (VISE) is a free and open source software for visual search of a large number of images using an image as a search query.



Draw search query region and click "Search"



 Windows 10
 GNU/Linux
 macOS
 Source Archive

 Image: Construction of the state of

robots.ox.ac.uk/~vgg/software/vise

Image-matching with SIFT features within a 'bag of visual words'



## What are SIFT features?

- Geometric features, defined by differences in texture
- These features are scale invariant (size doesn't matter)
- These features survive rotation and skewing, but not flipping or warping [without modification], so features within images to be matched must be on the same plane (e.g. details of the façade of a building)

## What is a 'bag of visual words'?

### Counterpart to statistical models of words in texts



Keywords in documents

but

### Matching logos in BBC video



https://www.robots.ox.ac.uk/~vgg/research/on-the-fly/

### Finding printed illustrations





#inliers= 79

Putative Boxes Lines Regions Draw again



High resolution full image Search on full image

http://ballads.bodleian.ox.ac.uk

#### http://15cbooktrade.ox.ac.uk/illustration/

## What else can VISE match?

• Tiled surfaces, mosaics, coins, stamps, seals, printing surfaces (i.e. blocks and plates) and their offprints..

• Duplicate image files where there are small differences (changes in file format, rotation, size, filename...)

## Matching of woodblocks – preprocessing



#inliers = 40

## Matching of woodblocks – postprocessing



#inliers = 273



James Basire, Monument of Eleanor, Queen of England, James Basire, Bodleian Library Gough Copperplates d.12

### Matching the plate to an offprint



### Deduplication – Ashmolean Museum case study



Offline use-case: deduplication of Ashmolean Museum studio images



## Common computer vision tasks

- Comparison (spotting the differences between images)
- Matching (finding out what images have in common)
- <u>Classification (describing what is in an image its</u> <u>content or subject)</u>
- Facial recognition (combines some elements from the above)

### **Image Matching**

Are these images of the same object?



All Souls college



instance

### **Image Classification**

4

What object is this?





category

### Deep learning with a Convolutional Neural Network (CNN)



https://www.aldec.com/en/company/blog/177--how-to-develop-high-performance-deep-neural-network-object-detectionrecognition-applications-for-fpga-based-edge-devices

### Training by example: the VGG ArtUK search system



http://www.robots.ox.ac.uk/~vgg/research/art\_search/

### Case study: ARTUK (formerly Your Paintings)



#### Elliot J. Crowley, Ernesto Coto and Andrew Zisserman

#### Overview

The objective of this research is to find objects in paintings by learning classifiers from photographs on the internet. We provide a demo that allows a user to search for an object of their choosing in a dataset of over 200,000 paintings in a matter of seconds.



#### Live Demo

You can try out the demo system using one of the following examples or by clicking the demo button:

abstract	arch	baby	beard	bird	boat	button	cap	chair	church	cow
crowd	dog	earrings	fire	flower	forest	fruit	garden	geometric	horse	jug
lace	man	medal	mountain	moustache	road	sand	seascape	sheep	shoe	snow

http://www.robots.ox.ac.uk/~vgg/research/art\_search/





Home with Catch



Amsterdam, Harbour...



Shipping in the Tham...



Dutch Canal Scene



Wiew of Berwick-upon...



Next >

Cockenzie Harbour



Sunderland Harbour,...

The Market Boat





#### SS 'Moor' off Coast...



#### Fishing off Chioggia...



#### St Ives Harbour



Hull Pier from the...



Deep Sea Liners



Shipping in the Mers...







Newport Docks



Boats in the Inner...





Architectural Capric...



Rome: The Interior...



Interior of a Church



Interior of a Cathed...



Courtyard Scene



St Peter Delivered...



Interior of St Peter...



Imaginary Cathedral...



Classical Ruins



Procession in a Cath...



Wiew in St Bartholom ...



Cathedral Interior



The Interior of St...



Alexander Slaying...



The Church of St...



Next >

The Interior of a...



Interior of a Classi...



The Interior of Alkm...





III John Lloyd (1771-182...



The Meet of a Hunt...



Point-to-Point



A Couple of Foxhound...



Lord Anson's Hunt...



John Press and The...



Hunting Scene of...



The Burton Hunt



Quorn Hounds, 'Cruis...



Six of the Duke of...



Two Couple of Hounds...



The Dogs' Home



A Warwickshire Hunt



A Meet of the Fife...



III The Llanharan Hunt



Serpentine, Hyde ...



A Riding School in...





Stamford Bridge



Carabou 3



Water Painting



III Peziza



Celebration



Moonwalker



Ship in Harbour



Abstract No.17\*



Circus (Big Top Perf...



The Road to Issel



III Hop Garden in Kent







Next >

Untitled



Spiral Construction



III PS 'Waterwitch'

### Misclassification – not dogs, but horses



### Misclassification – a two-part explanation

a. Dogs and horses are visually somewhat similar

b. Dogs and horses *co-occur* in the training data: it is *insufficiently granular* for the classifier to learn features that are only specific to dogs

### **Object Detection**

Object detection – a method in computer vision in which the goal is to predict a bounding box containing a specific class of object.

I.E. – localised image classification



Detecting people, buses and umbrellas with the EfficientDet object detector trained on the COCO dataset

# Retraining EfficientDet on NLS Chapbooks dataset with VGG List Annotator (LISA) software





https://gitlab.com/vgg/lisa

### Positive results



### Two false positives and one false negative



Question: how can we explain what we are looking for in terms that computers can understand?

false detection caused by semitransparent blank page

missed detection

Hup.

Current research – detection and classification of regions within images – training using VGG Image Annotator (VIA)



Envisioning Dante: collaboration with Dr. Guyda Armstrong, University of Manchester

### Image Segmentation – comparison of architectures



Image courtesy Suny Shtedritski (VGG)

## Common computer vision tasks

- Comparison (spotting the differences between images)
- Matching (finding out what images have in common)
- Classification (describing what is in an image its content or subject)
- Facial recognition

### British Library Broadcast News collaboration



https://www.bl.uk/collection-guides/television-and-radio-news

### Automatic detection and identification of people in British Library archive videos

Andrew Brown Ernesto Coto Andrew Zisserman

Visual Geometry Group, University of Oxford



http://www.robots.ox.ac.uk/~vgg/research/person\_id\_in\_video/

## Face/painting recognition



Figure 4: Successful retrievals using a CNN representation. In each case, the five query photos are shown beside the top retrieved painting. (a) Alexandra Wentworth, (b) Amy Sedaris, (c) Bradley James, (d) Christopher Ecclestone, (e) David Cross, (f) Deborah Harry, (g) James Brown, (h) Mary Tyler Moore, (i) Richard Harris, (j) Stevie Nicks.

## Face/painting 'nearest neighbor' search



Figure 6: Photos of famous people and their closest matching portrait from 'Your Paintings'. (a) Jennifer Lawrence, (b) Simon Cowell, (c) David Cameron, (d) Madonna, (e) Benedict Cumberbatch, (f) Natalie Dormer.

P	ortrait Matcher	
	Drop a photo here Drop a photo here Drop a photo here Submit	
	Click on one of the images below to try it !	
9 3		

http://www.robots.ox.ac.uk/~vgg/research/face\_paint/

## **Portrait Matcher**

These are your top 3 matches. Click on your favourite one.



Who is this?

Who is this?

Who is this?

http://www.robots.ox.ac.uk/~vgg/research/face\_paint/

### Racial and gender bias in facial recognition



Joy Buolamwini, 'How I'm fighting bias in algorithms', TedX talk (2017), www.poetofcode.com

## An (inadequate) excuse

Several technologists we spoke to mentioned that photos of people with darker skin tend to have less color contrast, making it harder to extract the features that [some] algorithms use to compare faces.

Brendan F. Klare et al., Face Recognition Performance: Role of Demographic Information, 7 IEEE Transactions on Information Forensics and Security 1789, 1797 (2012)

## Skin tone and media bias



Figure 1 Polaroid Shirley card (Printed with permission of Polaroid)

Figure 2 Traditional Kodak Shirley card (Printed with permission of Kodak)

Roth, Lorna. Looking at Shirley, the Ultimate Norm: Colour Balance, Image Technologies, and Cognitive Equity. *Canadian Journal of Communication*, v. 34, n. 1, mar. 2009.

## Addressing bias





#### Publication

Turning a Blind Eye: Explicit Removal of Biases and Variation from Deep Neural Network Embeddings



#### Presentation

Bias Estimation in Face Datasets, ECCV 2018 Workshop Presentation



#### **Download Dataset\***

List of URLs with ancestral origin annotations

## Ethnicity classification



#### Western

#### Eastern

Fu, Siyao, Haibo He, and Zeng-Guang Hou. "Learning race from face: A survey." IEEE transactions on pattern analysis and machine intelligence 36.12 (2014): 2483-2509

## Ethnicity classification (2)

### **Coming soon**

PATH TRACKING ETHNICITY RECOGNITION

Recognizes a person's ethnicity



https://ntechlab.com

## Ethnicity classification (3)



https://www.kairos.com/blog/ethnicity-diversity-how-we-detect-it-and-why-it-matters

## Test and training data for critical AI studies





Joy Buolamwini and Timnit Gebru. "Gender shades: Intersectional accuracy disparities in commercial gender classification." *Conference on Fairness, Accountability and Transparency*. 2018

### Questions?

### Slides, worksheet and demo data

## tinyurl.com/FF21VISUAL-AI

### Contacts

VGG - <u>http://www.robots.ox.ac.uk/~vgg/</u>

Me – <u>giles.bergel@eng.ox.ac.uk</u> Twitter- @ChapBookPro