Introduction to Computer Vision

Danna Gurari
University of Colorado Boulder
Fall 2023

https://home.cs.colorado.edu/~DrG/Courses/RecentAdvancesInComputerVision/AboutCourse.html

Today's Topics

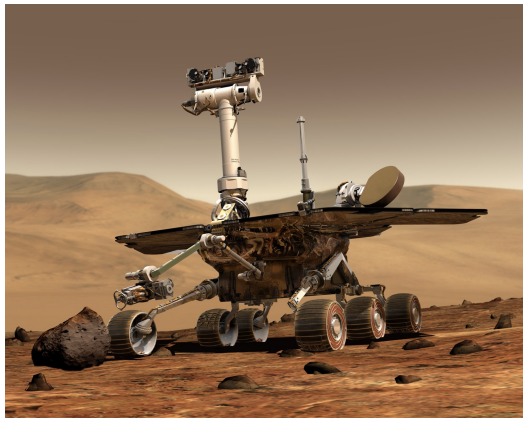
- Computer vision: origins
- What makes computer vision hard?
- Research in computer vision
- Course logistics

Today's Topics

- Computer vision: origins
- What makes computer vision hard?
- Research in computer vision
- Course logistics

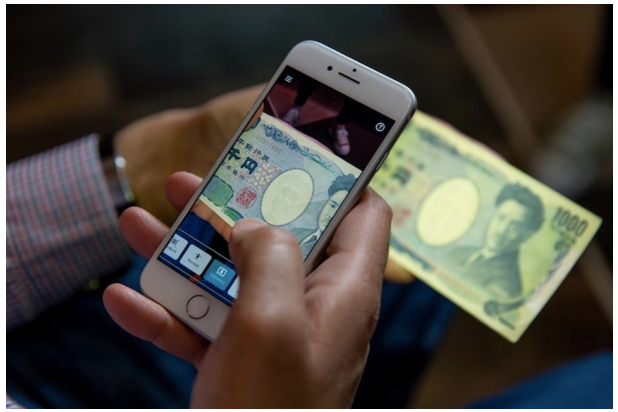
Computer Vision: Computers that “See”

Self-driving cars

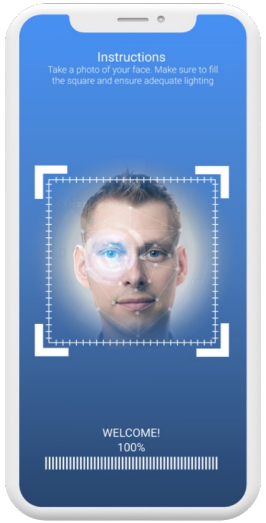


Exploration on Mars

Guided surgery



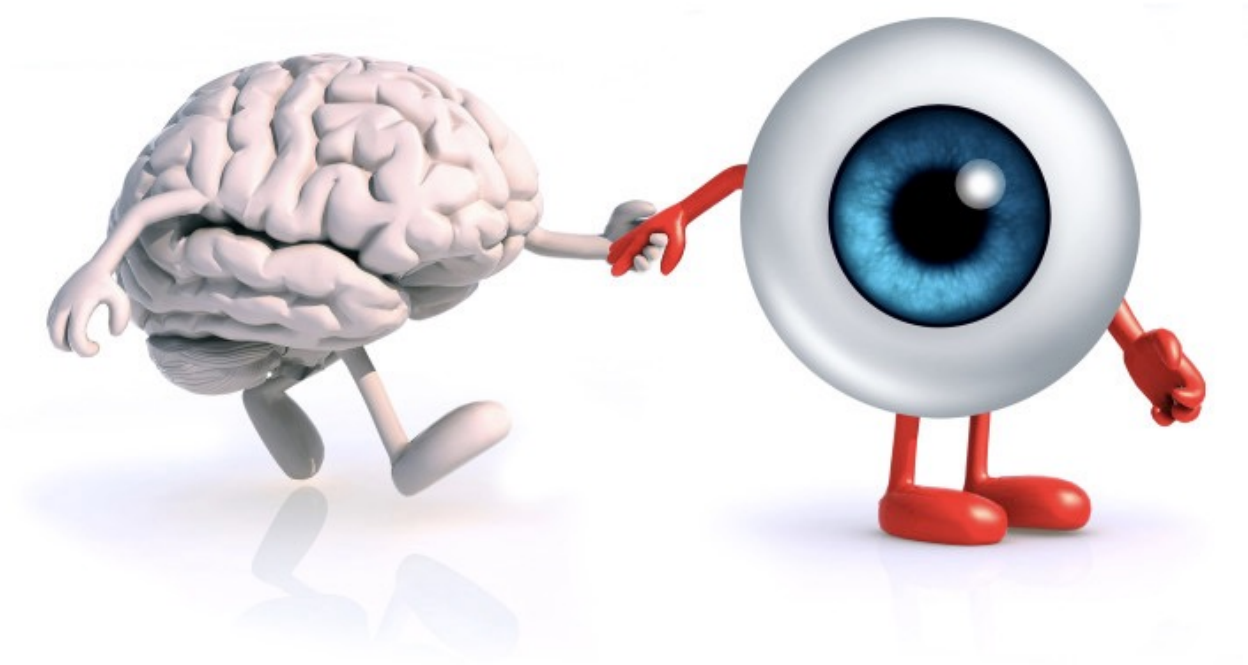
Visual assistance for people who are blind



Security

Origins of Computer Vision

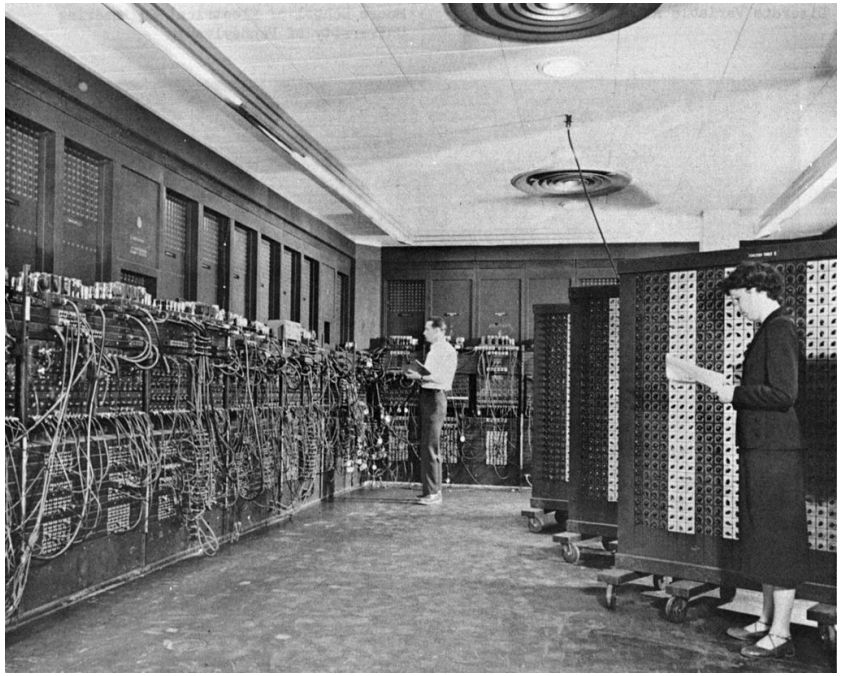
Emulating the basic ingredients of sight:



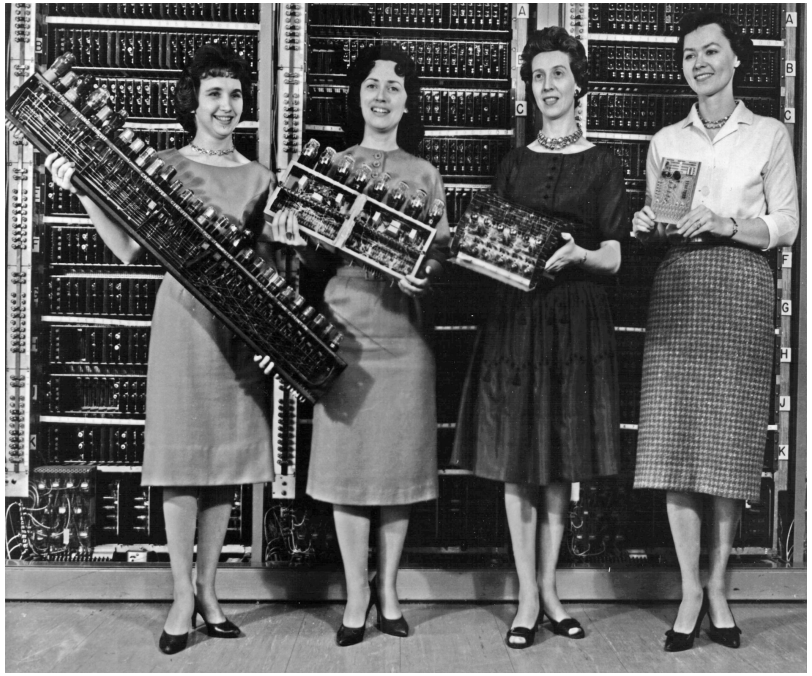
Origins of Computer Vision

1945

First programmable machine



ENIAC (Electronic Numerical Integrator and Computer) created during World War II
(could compute 5,000 additions in one second)



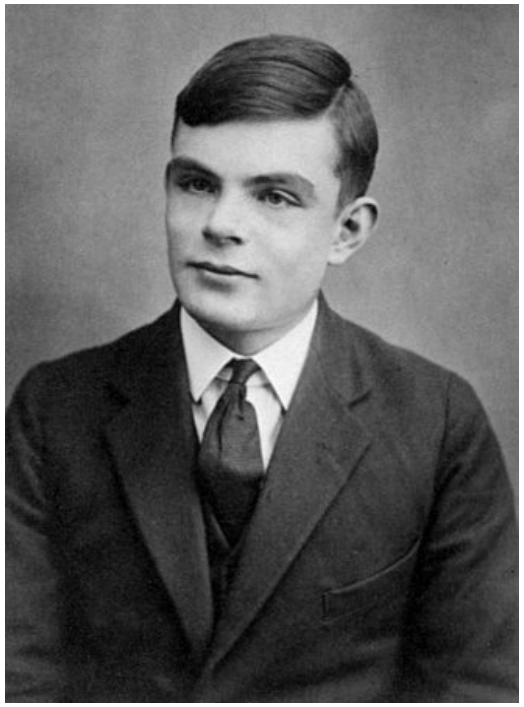
First programmers

Origins of Computer Vision

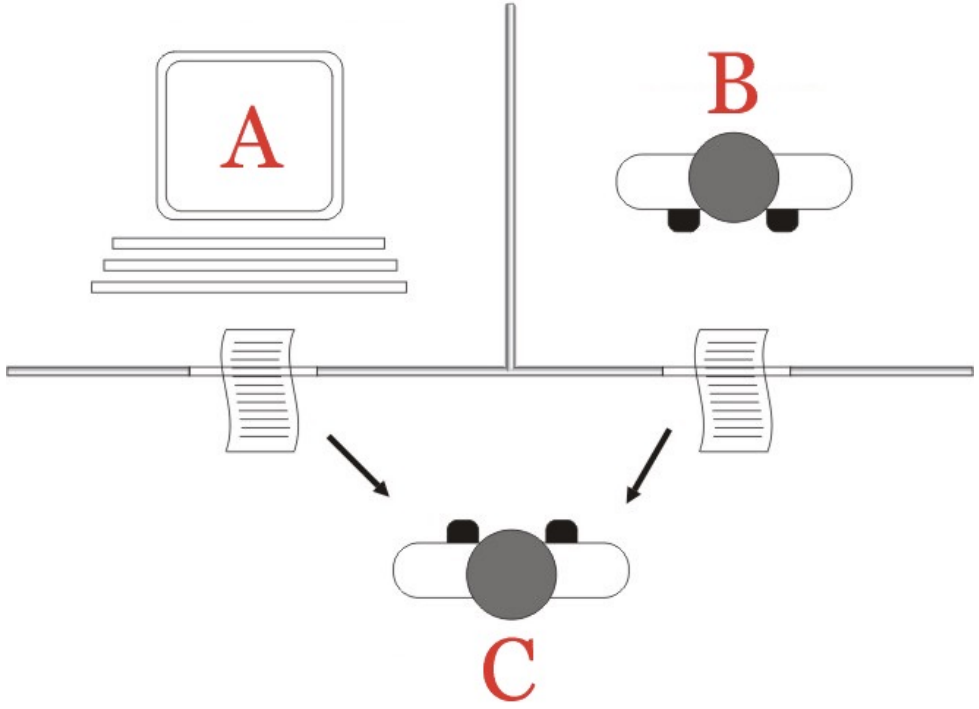
1945 1950

First programmable machine

Turing Test



Alan Turing
(1912-1954)



Turing Test: can ”C” decide whether text responses come from a machine or human

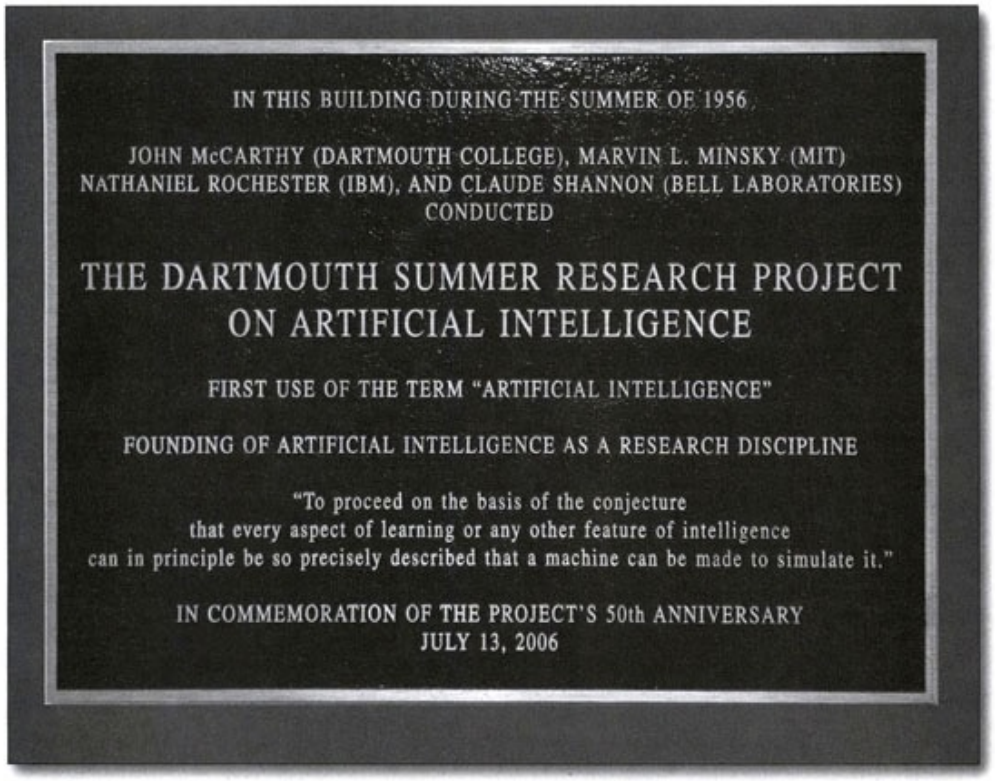
Origins of Computer Vision

1945 1950 1956

First programmable machine

Turing Test

AI birth



"Artificial intelligence" established as a field at a workshop

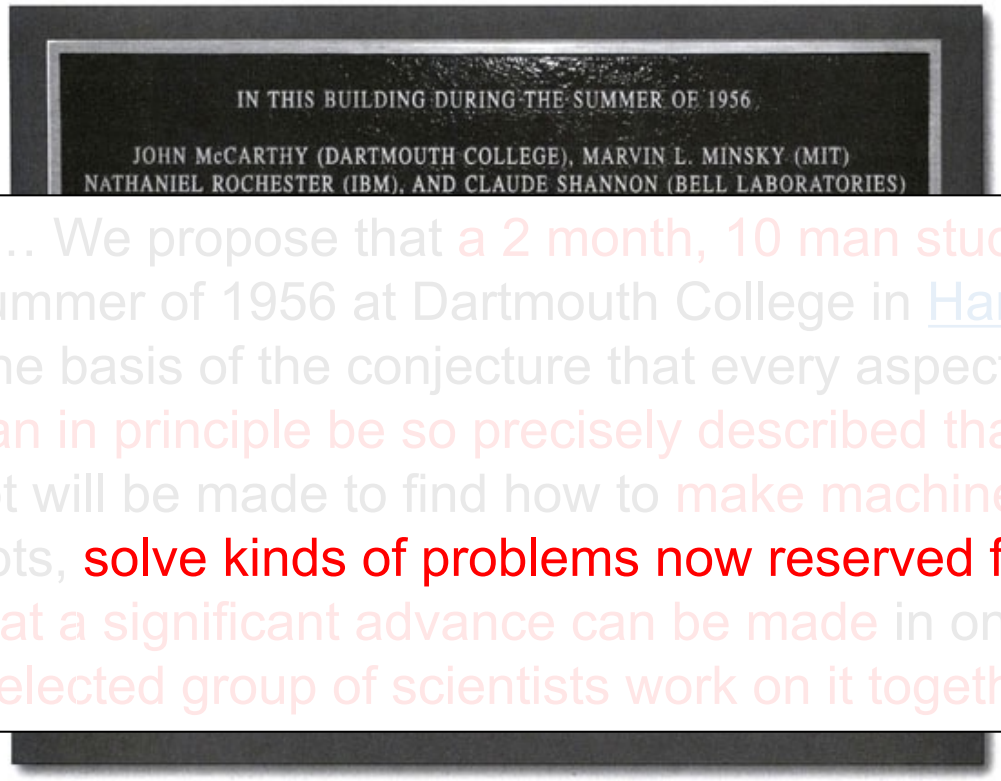
Origins of Computer Vision

1945 1950 1956

First programmable machine

Turing Test

AI birth



Workshop Proposal: "… We propose that a 2 month, 10 man study of artificial intelligence be carried out during the summer of 1956 at Dartmouth College in Hanover, New Hampshire. The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves. We think that a significant advance can be made in one or more of these problems if a carefully selected group of scientists work on it together for a summer…"

"Artificial intelligence" established as a field at a workshop

Origins of Computer Vision

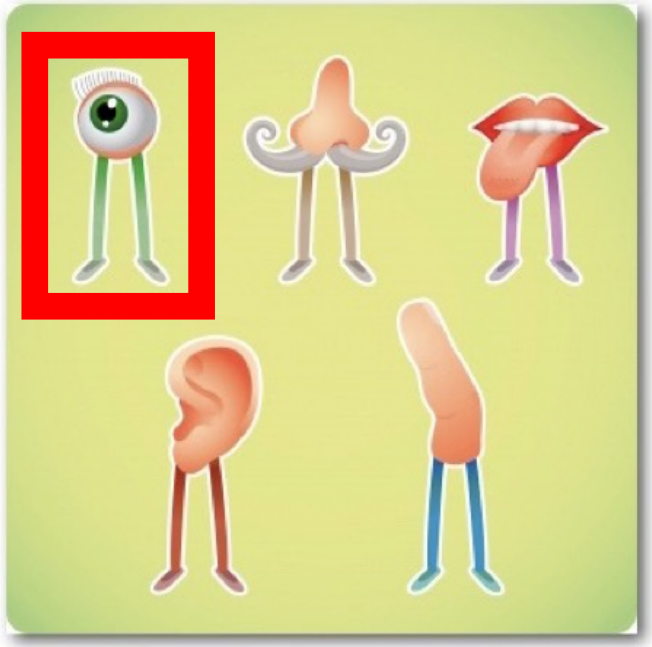
1945 1950 1956

First programmable machine

Turing Test

AI birth

What human intelligence might computers imitate?



Origins of Computer Vision

1945 1950 1957

First programmable machine

Turing Test

AI birth

First digital image

176 x 176 pixels

Origins of Computer Vision

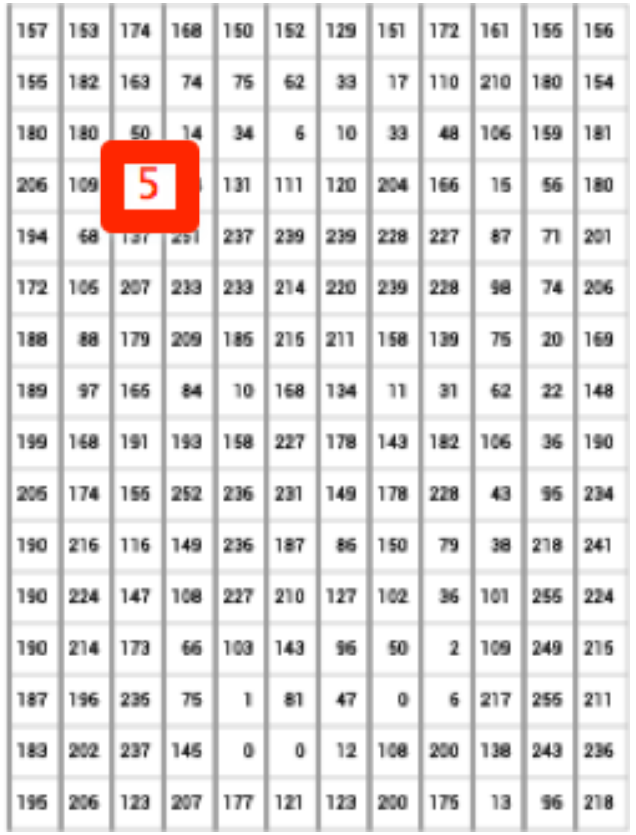
1945 1950 1957

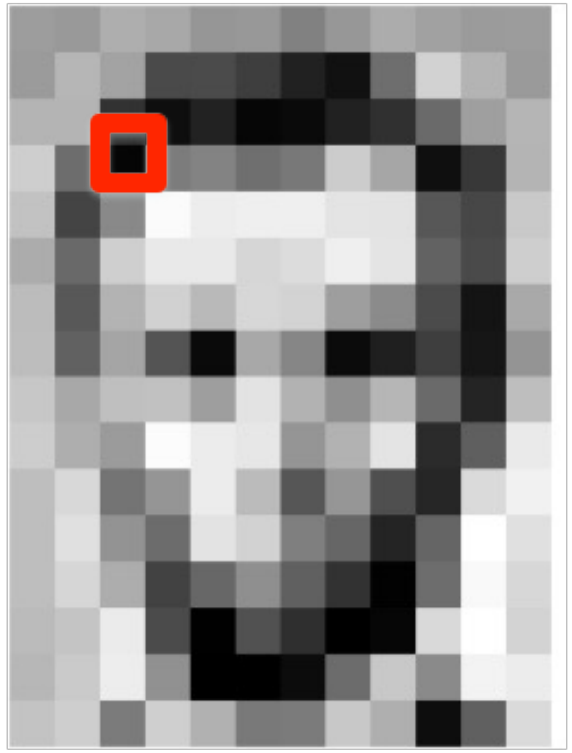
First programmable machine

Turing Test

AI birth

First digital image





What a Computer Sees:

Origins of Computer Vision

1945 1950 1957

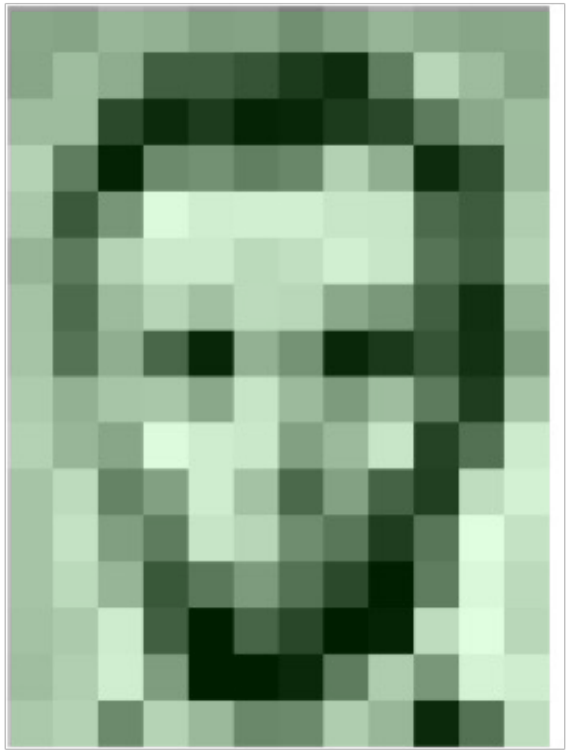
First programmable machine

Turing Test

AI birth

First digital image

157	153	174	168	150	152	129	151	172	161	155	156
155	182	163	74	75	62	33	17	110	210	180	154
180	180	50	14	34	6	10	33	48	106	159	181
206	109	5	124	131	111	120	204	166	15	56	180
194	68	137	251	237	239	239	228	227	87	71	201
172	105	207	233	233	214	220	239	228	98	74	206
188	88	179	209	185	215	211	158	139	75	20	169
189	97	165	84	10	168	134	11	31	62	22	148
199	168	191	193	158	227	178	143	182	106	36	190
205	174	155	252	236	231	149	178	228	43	95	234
190	216	116	149	236	187	86	150	79	38	218	241
190	224	147	108	227	210	127	102	36	101	255	224
190	214	173	66	103	143	96	50	2	109	249	215
187	196	235	75	1	81	47	0	6	217	255	211
183	202	237	145	0	0	12	108	200	138	243	236
195	206	123	207	177	121	123	200	175	13	96	218



What a Computer Sees:

Origins of Computer Vision

1945 1950 1957

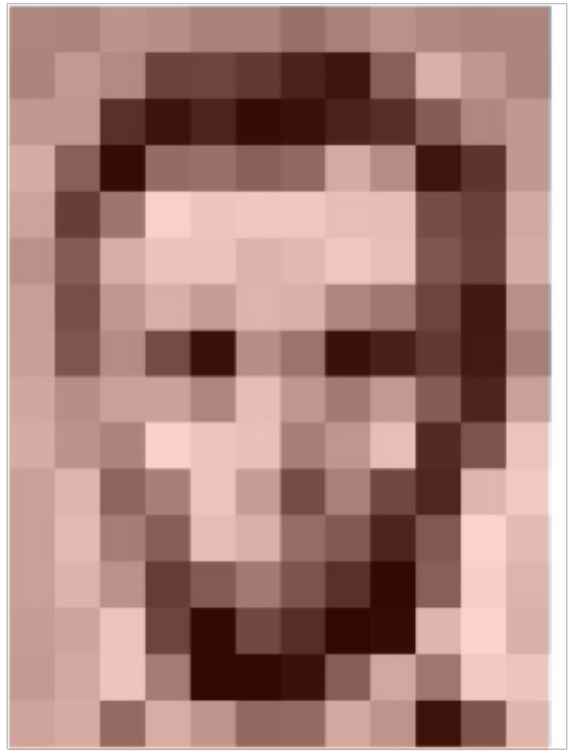
First programmable machine

Turing Test

AI birth

First digital image

157	153	174	168	150	152	129	151	172	161	155	156
155	182	163	74	75	62	33	17	110	210	180	154
180	180	50	14	34	6	10	33	48	106	159	181
206	109	5	124	131	111	120	204	166	15	56	180
194	68	137	251	237	239	239	228	227	87	71	201
172	105	207	233	233	214	220	239	228	98	74	206
188	88	179	209	185	215	211	158	139	75	20	169
189	97	165	84	10	168	134	11	31	62	22	148
199	168	191	193	158	227	178	143	182	106	36	190
205	174	155	252	236	231	149	178	228	43	95	234
190	216	116	149	236	187	86	150	79	38	218	241
190	224	147	108	227	210	127	102	36	101	255	224
190	214	173	66	103	143	96	50	2	109	249	215
187	196	235	75	1	81	47	0	6	217	255	211
183	202	237	145	0	0	12	108	200	138	243	236
195	206	123	207	177	121	123	200	175	13	96	218



What a Computer Sees:

Origins of Computer Vision

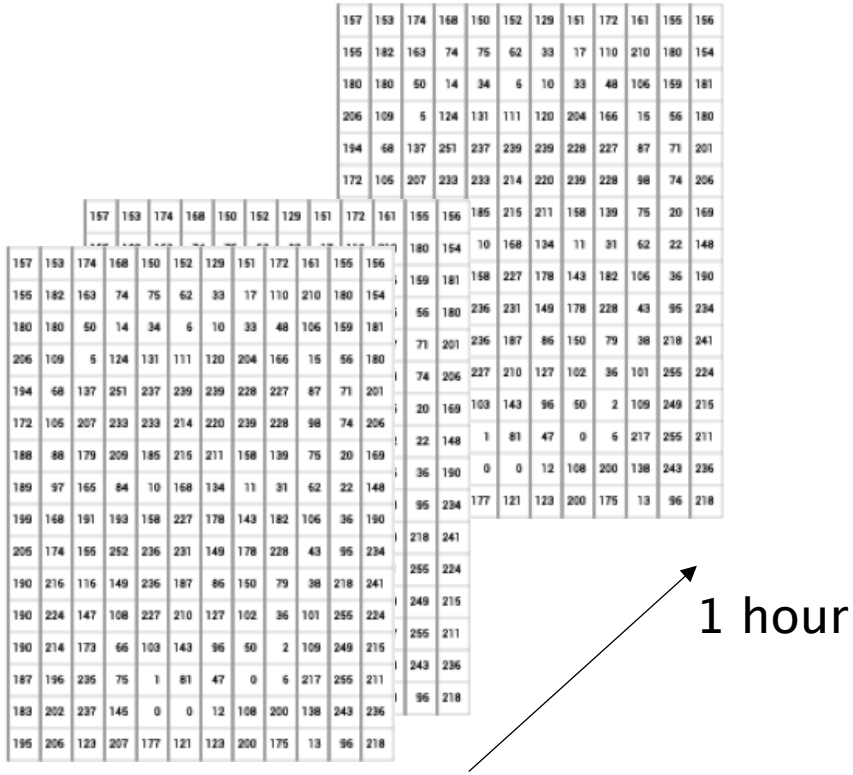
1945 1950 1957

First programmable machine

Turing Test

AI birth

First digital image

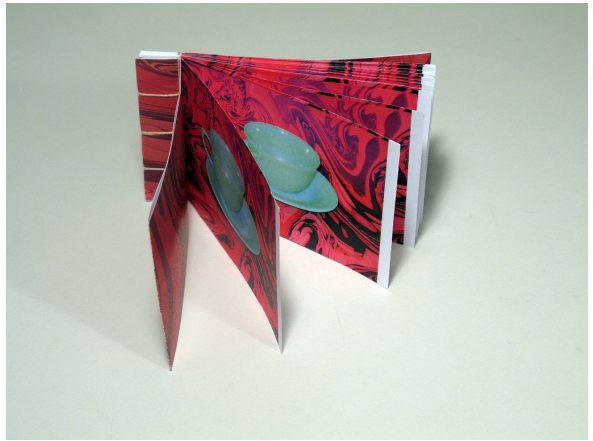


What a Computer Sees:

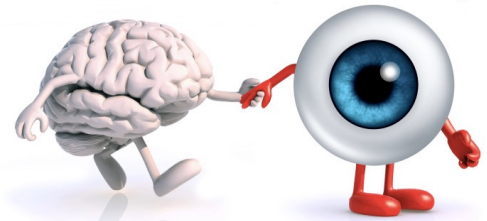
Time 1

1 hour

Analogous to (for video):



Origins of Computer Vision



1945 First programmable machine

1950 Turing Test

1957 AI birth

1957 First digital image

1966 Birth of Computer Vision

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

PROJECT MAC

Artificial Intelligence Group
Vision Memo. No. 100.

July 7, 1966

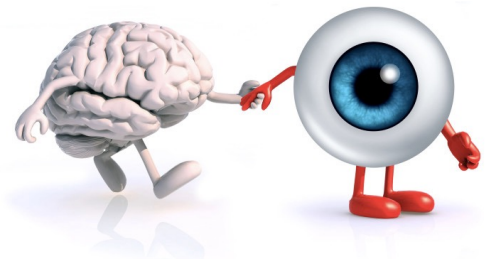
THE SUMMER VISION PROJECT

Seymour Papert

The summer vision project is an attempt to use our summer workers effectively in the construction of a significant part of a visual system.

sub-problems which will allow individuals to work independently and yet participate in the construction of a system complex enough to be a real landmark in the development of "pattern recognition".

Origins of Computer Vision



1945 First programmable machine

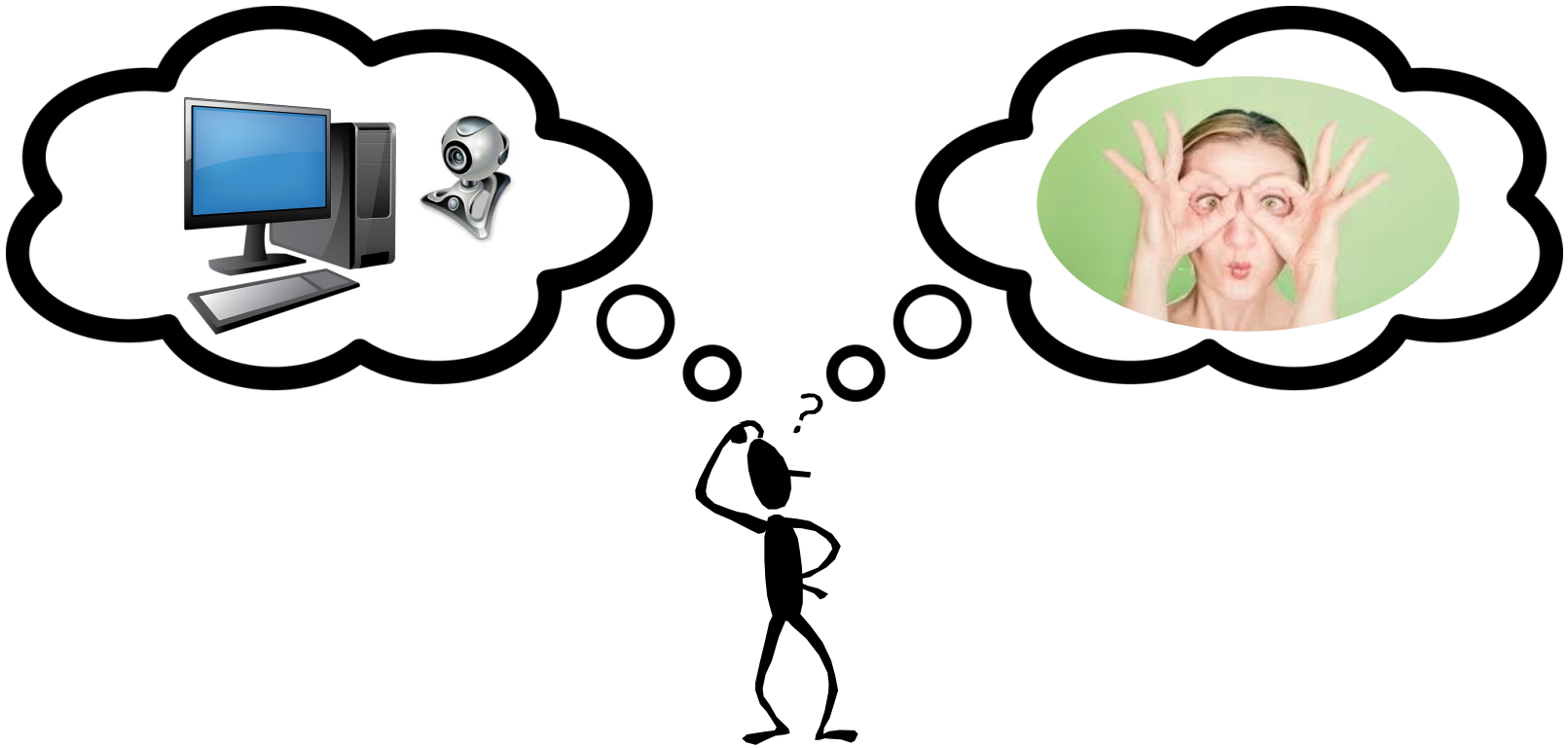
1950 Turing Test

1957 AI birth

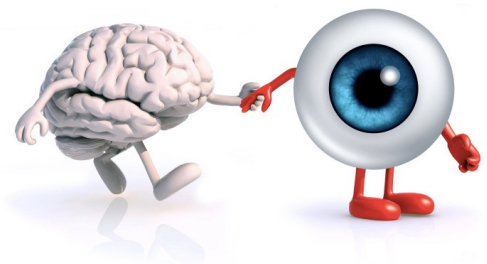
1957 First digital image

1966 Birth of Computer Vision

Turing Test: design “computer vision” that is indistinguishable from “human vision”



Origins of Computer Vision



1945 First programmable machine

1950 Turing Test

1957 AI birth

1957 First digital image

1966 Birth of Computer Vision

e.g.,

157	153	174	168	150	152	129	151	172	161	155	156
155	182	163	74	75	62	33	17	110	210	180	154
180	180	50	14	34	6	10	33	48	106	159	181
206	109	5	124	131	111	120	204	166	15	56	180
194	68	137	251	237	239	239	228	227	87	71	201
172	105	207	233	233	214	220	239	228	98	74	206
188	88	179	209	185	215	211	158	139	75	20	169
189	97	165	84	10	168	134	11	31	62	22	148
199	168	191	193	158	227	178	143	182	106	36	190
205	174	155	252	236	231	149	178	228	43	95	234
190	216	116	149	236	187	86	150	79	38	218	241
190	224	147	108	227	210	127	102	36	101	255	224
190	214	173	66	103	143	96	50	2	109	249	215
187	196	235	75	1	81	47	0	6	217	255	211
183	202	237	145	0	0	12	108	200	138	243	236
195	206	123	207	177	121	123	200	175	13	96	218

What is this?
- A picture of a person

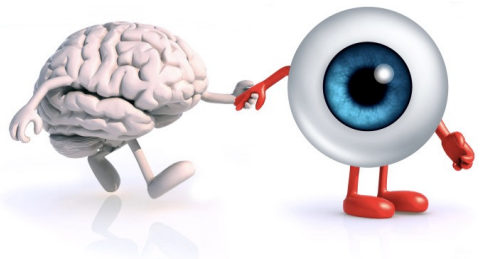
Could you describe this person?
- Long face
- Angular jaw
- Has a beard

Who is this person?
- Abraham Lincoln

Is this person happy?
- I am not sure.

Is this person attractive?
- ~70% of people would say “yes”

Origins of Computer Vision



1945 First programmable machine

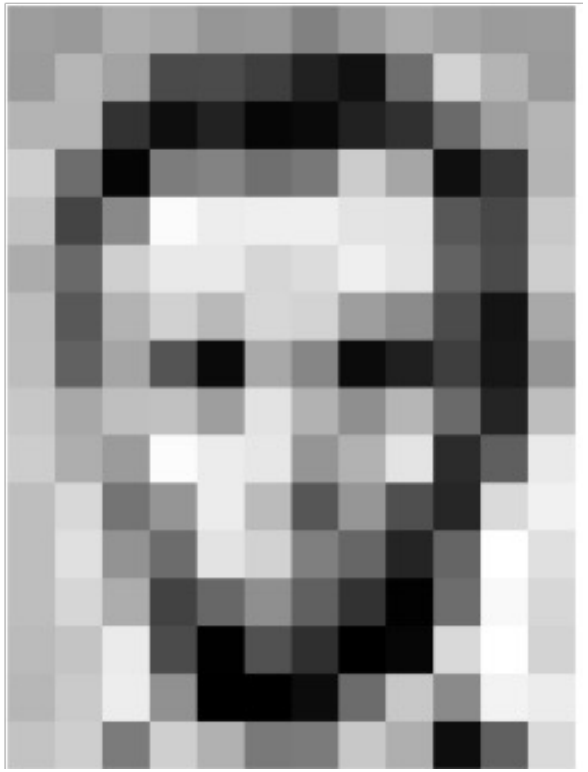
1950 Turing Test

1957 AI birth

1957 First digital image

1966 Birth of Computer Vision

e.g.,



What is this?
- A picture of a person

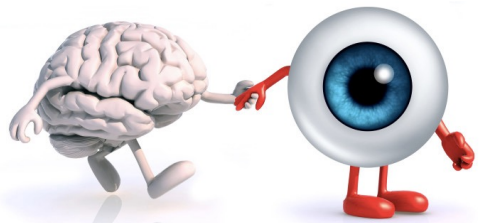
Could you describe this person?
- Long face
- Angular jaw
- Has a beard

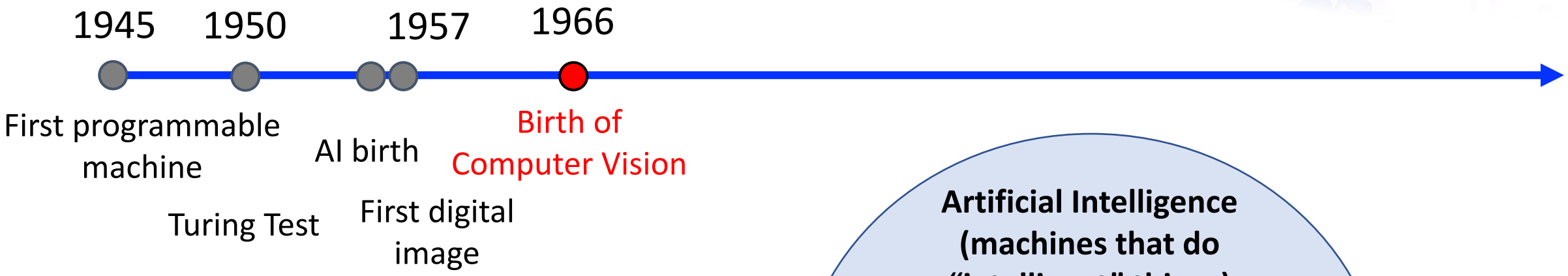
Who is this person?
- Abraham Lincoln

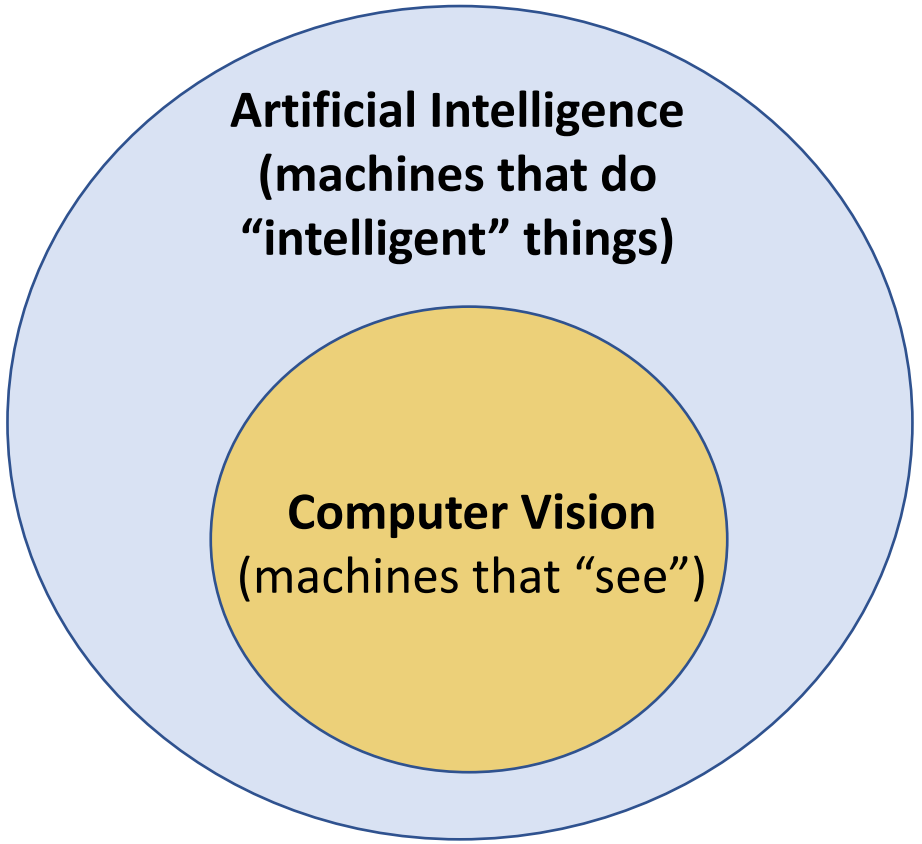
Is this person happy?
- I am not sure.

Is this person attractive?
- ~70% of people would say “yes”

Origins of Computer Vision



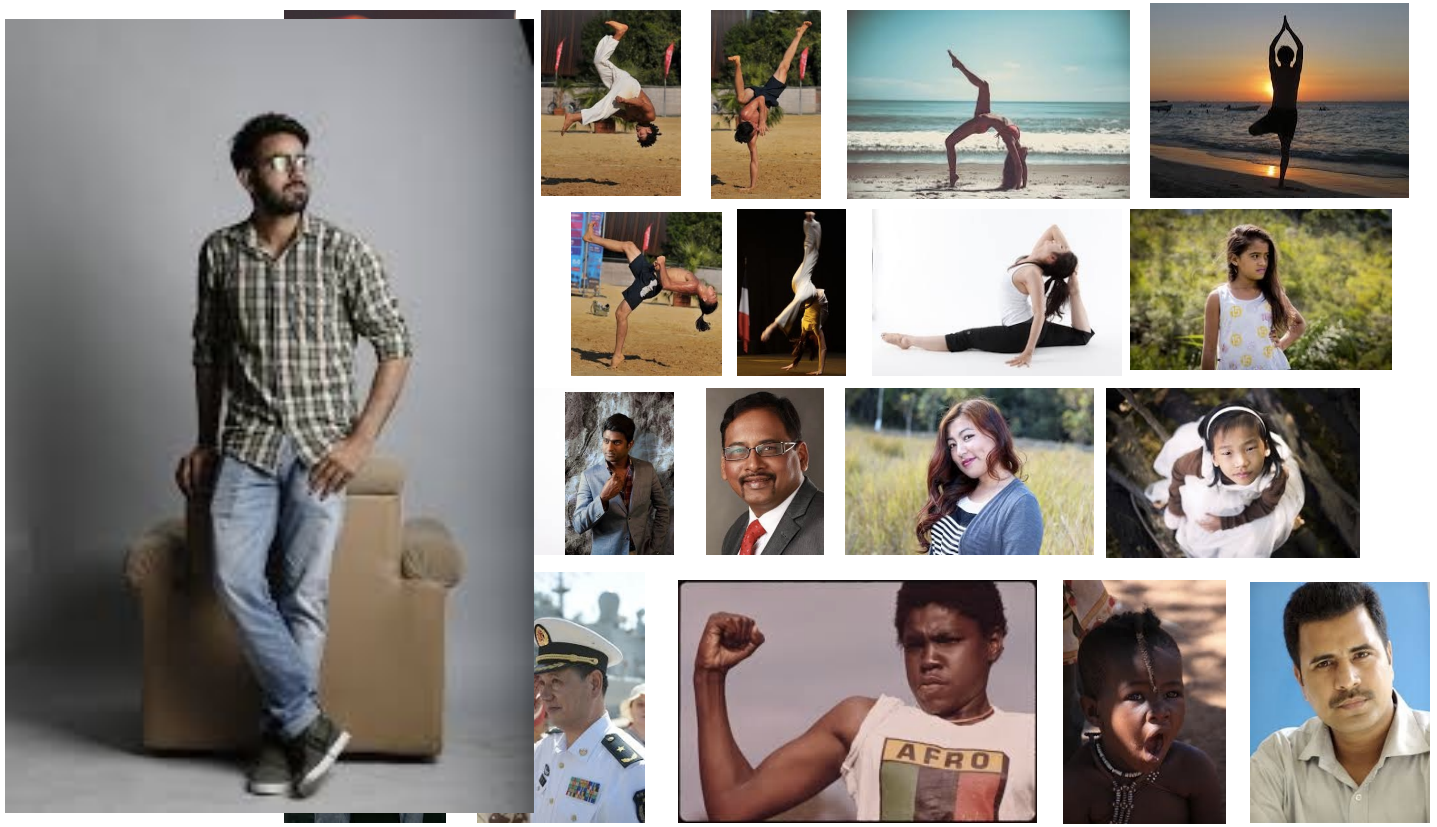




Today's Topics

- Computer vision: origins
- What makes computer vision hard?
- Research in computer vision
- Course logistics

Class Discussion: How Would You Program a Computer to Answer “Is a Person in the Image?”



Key Challenge: Replicate Human Vision for So Much Variation for **So Many Tasks**!

- Object recognition
- Scene classification
- Attribute labeling
- Object detection
- Segmentation
- Image Captioning
- Visual Question Answering
- Activity/Event Recognition
- Object Tracking
- Subjective Problems
- And more…

Key Challenge: Replicate Human Vision for So Much Variation for **So Many Tasks**!

- Object recognition
- Scene classification
- Attribute labeling
- Object detection
- Segmentation
- Image Captioning
- Visual Question Answering
- Activity/Event Recognition
- Object Tracking
- Subjective Problems
- And more…

e.g., take a picture of an object and find where to buy it

Key Challenge: Replicate Human Vision for So Much Variation for **So Many Tasks**!

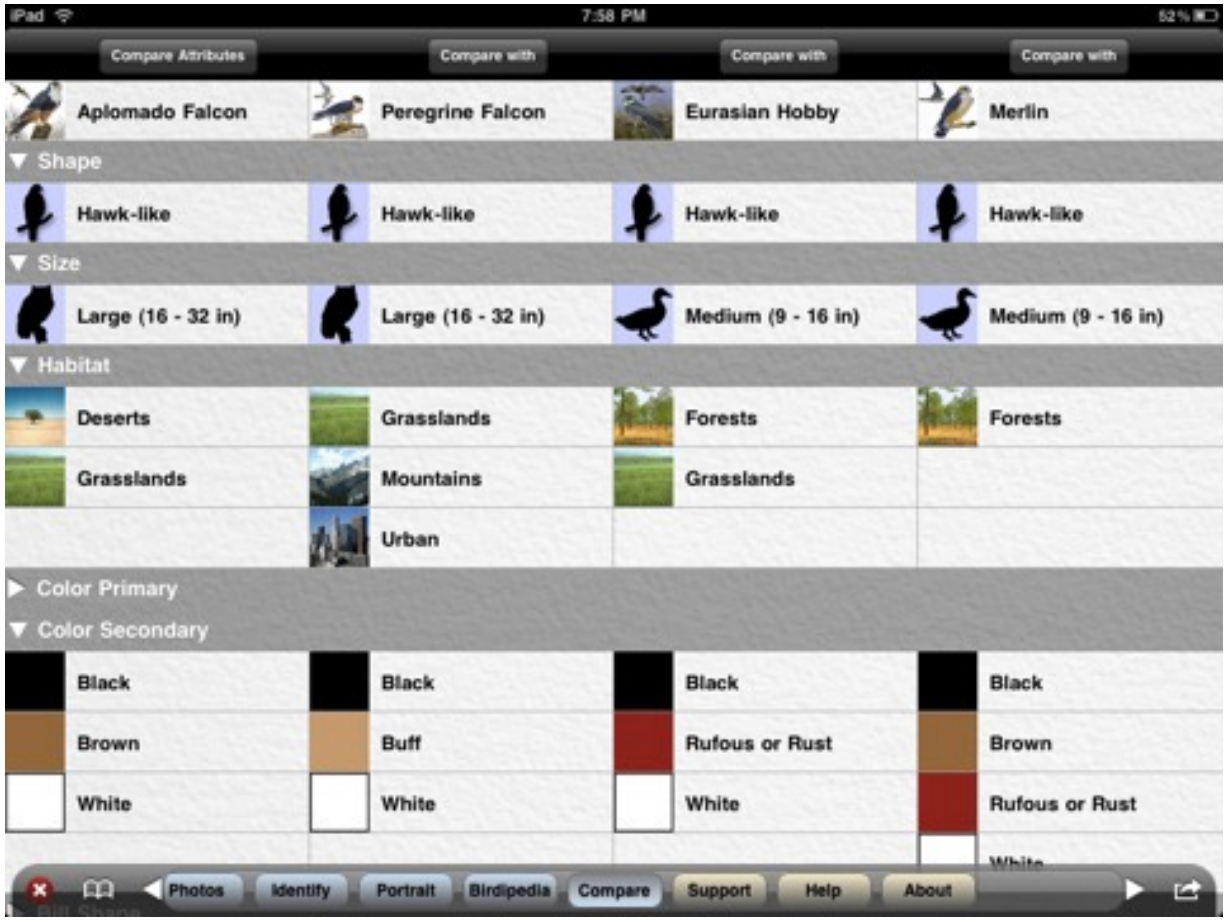
- Object recognition
- Scene classification
- Attribute labeling
- Object detection
- Segmentation
- Image Captioning
- Visual Question Answering
- Activity/Event Recognition
- Object Tracking
- Subjective Problems
- And more…

Kitchen

Store

Key Challenge: Replicate Human Vision for So Much Variation for **So Many Tasks**!

- Object recognition
- Scene classification
- Attribute labeling
- Object detection
- Segmentation
- Image Captioning
- Visual Question Answering
- Activity/Event Recognition
- Object Tracking
- Subjective Problems
- And more…

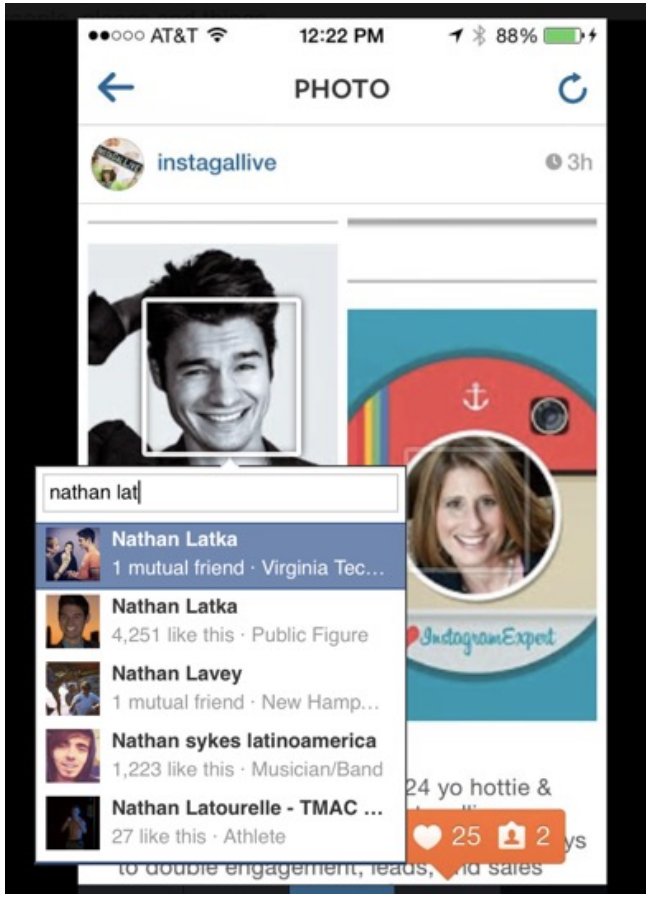


e.g., describe a bird to learn what type it is

Demo: https://www.youtube.com/watch?v=UPcz9Y17iCc

Key Challenge: Replicate Human Vision for So Much Variation for **So Many Tasks**!

- Object recognition
- Scene classification
- Attribute labeling
- Object detection
- Segmentation
- Image Captioning
- Visual Question Answering
- Activity/Event Recognition
- Object Tracking
- Subjective Problems
- And more…



e.g., detect faces to tag

Key Challenge: Replicate Human Vision for So Much Variation for **So Many Tasks**!

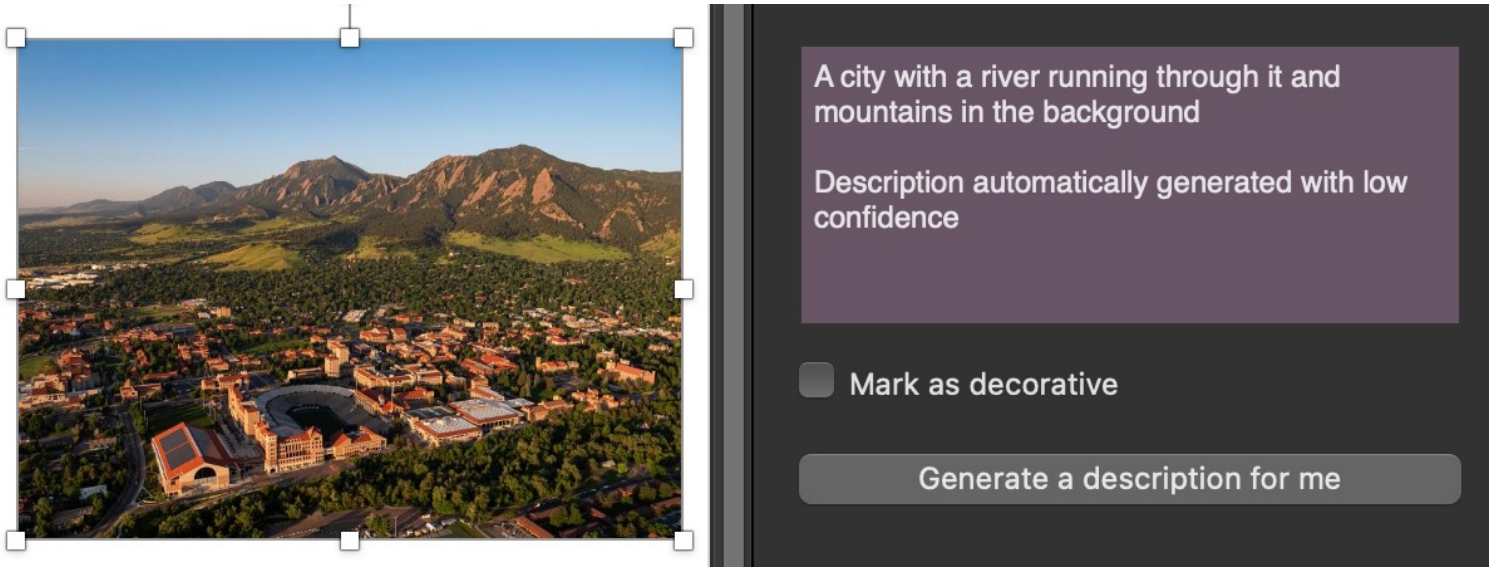
- Object recognition
- Scene classification
- Attribute labeling
- Object detection
- Segmentation
- Image Captioning
- Visual Question Answering
- Activity/Event Recognition
- Object Tracking
- Subjective Problems
- And more…

e.g., rotoscoping

https://www.starnow.co.uk/ahmedmohammed1/photos/4650871/before-and-after-rotoscopinggreen-screening

Key Challenge: Replicate Human Vision for So Much Variation for **So Many Tasks**!

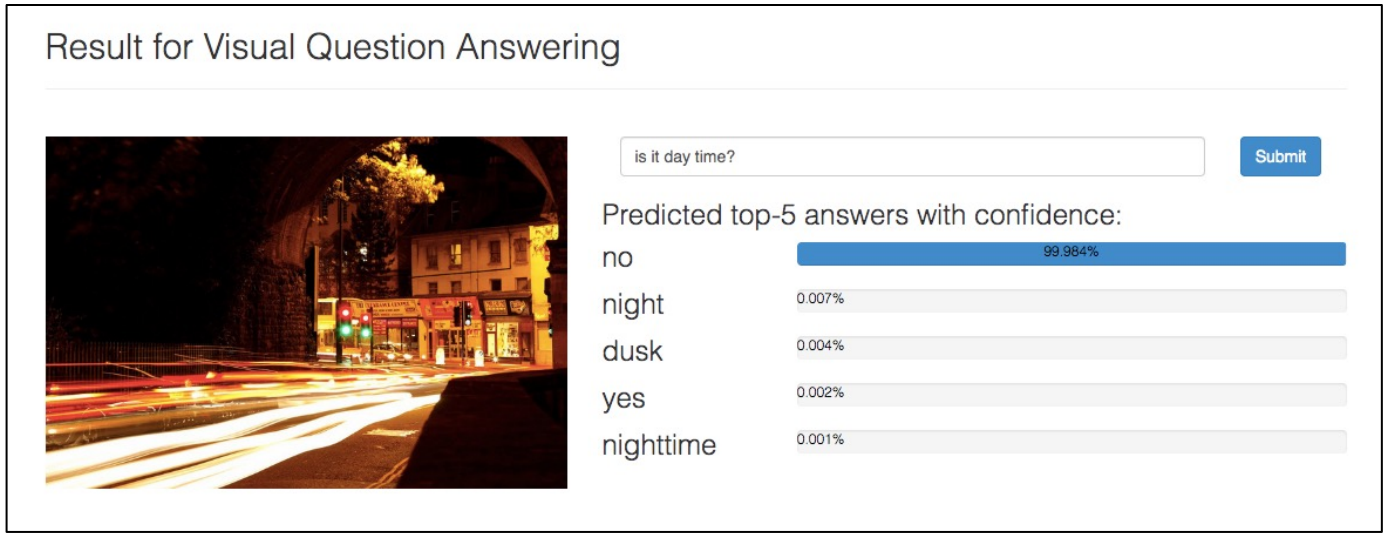
- Object recognition
- Scene classification
- Attribute labeling
- Object detection
- Segmentation
- Image Captioning
- Visual Question Answering
- Activity/Event Recognition
- Object Tracking
- Subjective Problems
- And more…



e.g., Microsoft Power Point

Key Challenge: Replicate Human Vision for So Much Variation for **So Many Tasks**!

- Object recognition
- Scene classification
- Attribute labeling
- Object detection
- Segmentation
- Image Captioning
- Visual Question Answering
- Activity/Event Recognition
- Object Tracking
- Subjective Problems
- And more…



Demo: 3

Key Challenge: Replicate Human Vision for So Much Variation for **So Many Tasks**!

- Object recognition
- Scene classification
- Attribute labeling
- Object detection
- Segmentation
- Image Captioning
- Visual Question Answering
- Activity/Event Recognition
- Object Tracking
- Subjective Problems
- And more…

e.g., shopping without a cashier

Key Challenge: Replicate Human Vision for So Much Variation for **So Many Tasks**!

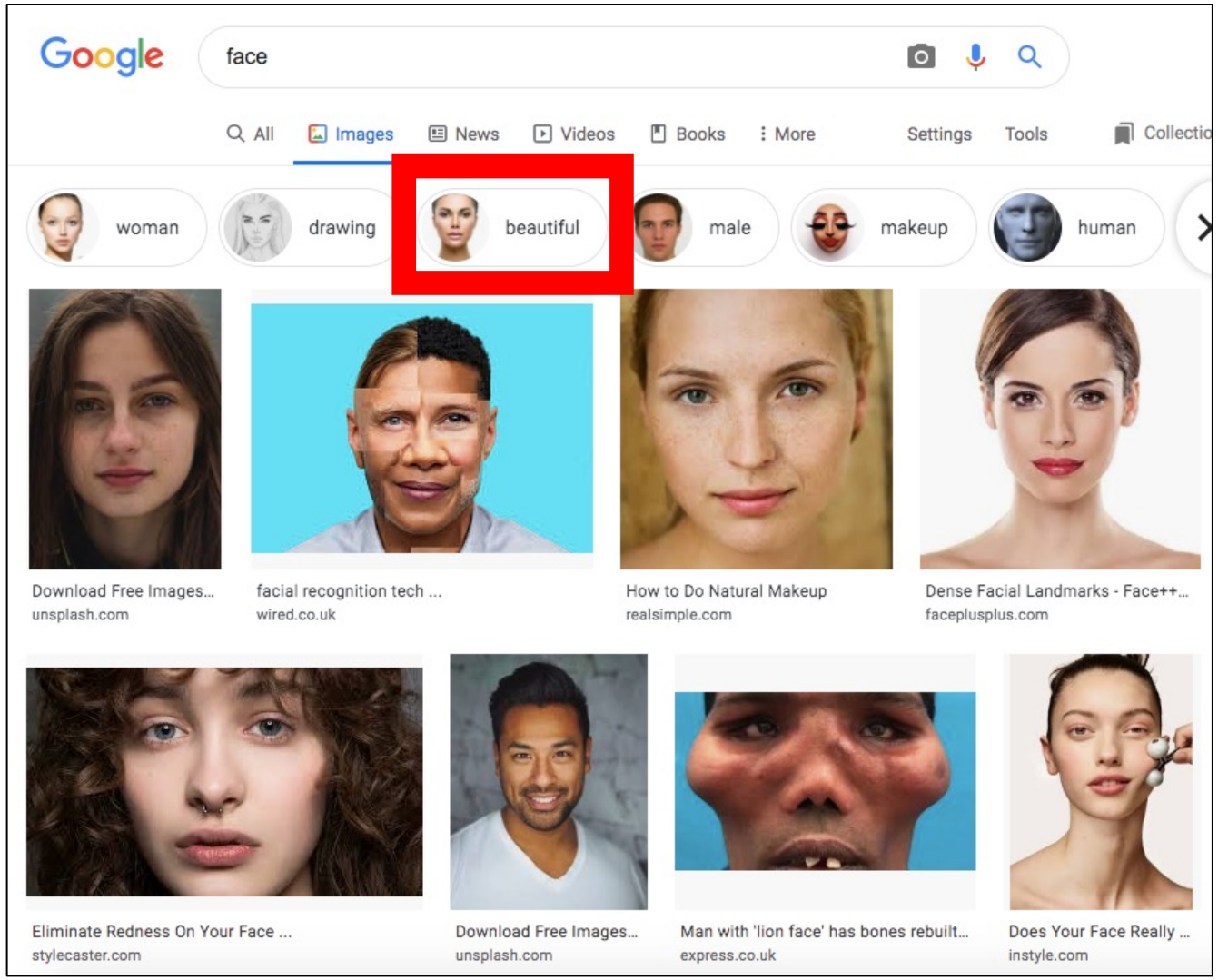
- Object recognition
- Scene classification
- Attribute labeling
- Object detection
- Segmentation
- Image Captioning
- Visual Question Answering
- Activity/Event Recognition
- Object Tracking
- Subjective Problems
- And more…

e.g., track bowling ball path

e.g., calculate bat speed

Key Challenge: Replicate Human Vision for So Much Variation for **So Many Tasks**!

- Object recognition
- Scene classification
- Attribute labeling
- Object detection
- Segmentation
- Image Captioning
- Visual Question Answering
- Activity/Event Recognition
- Object Tracking
- Subjective Problems
- And more…



Key Challenge: Replicate Human Vision for **So Much Variation** for So Many Tasks!

Illumination

Object pose

Clutter

Occlusions

Intra-class appearance

Viewpoint

Slide by Kristen Grauman

Key Challenge: Replicate Human Vision for **So Much Variation** for **So Many Tasks**!

images on hard drive: 10^5
(500 GB/2 MB = 250,000)

images seen during my first 10 years: 10^9
(24 images/sec * 60 sec * 60 min * 16 hr * 365 days * 10 yrs = 5,045,760,000)

images seen by all humanity: 10^{20}
(7.5 billion humans[1] * 24 images/sec * 60 * 60 * 16 * 365 * 60 yrs = $2.23 * 10^{20}$)

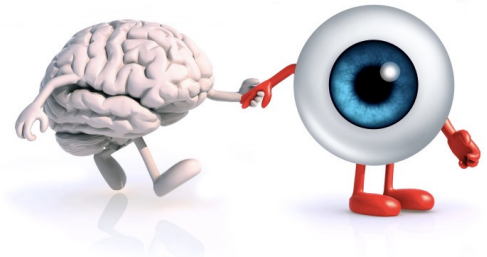
[1] http://www.worldometers.info/world-population/

Adapted from slides by Antonio Torralba

Today's Topics

- Computer vision: origins
- What makes computer vision hard?
- Research in computer vision
- Course logistics

Recall: Origins of Computer Vision



1945 First programmable machine

1950 Turing Test

1957 AI birth

1957 First digital image

1966 Birth of Computer Vision

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

PROJECT MAC

Artificial Intelligence Group
Vision Memo. No. 100.

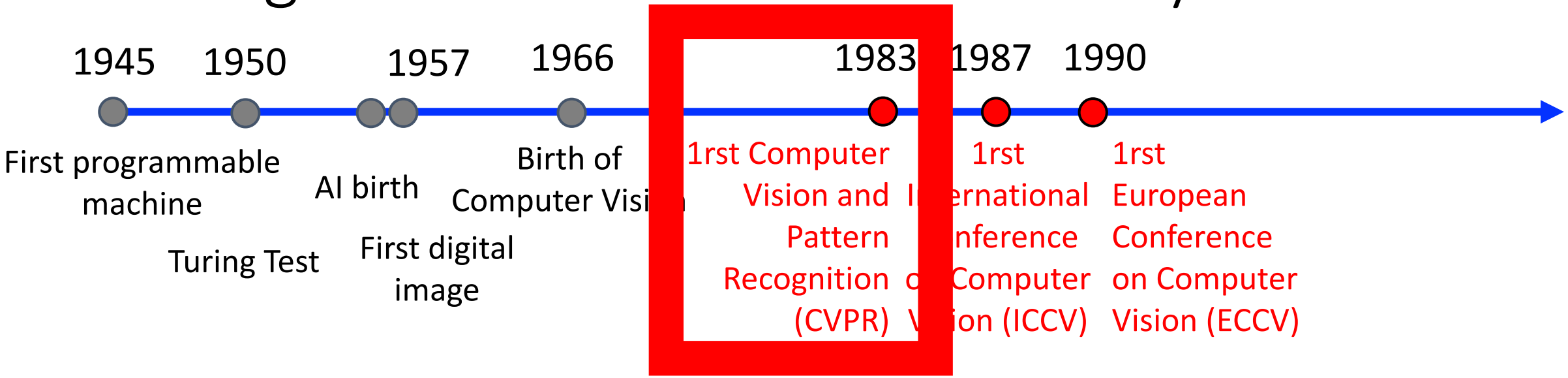
July 7, 1966

THE SUMMER VISION PROJECT

Seymour Papert

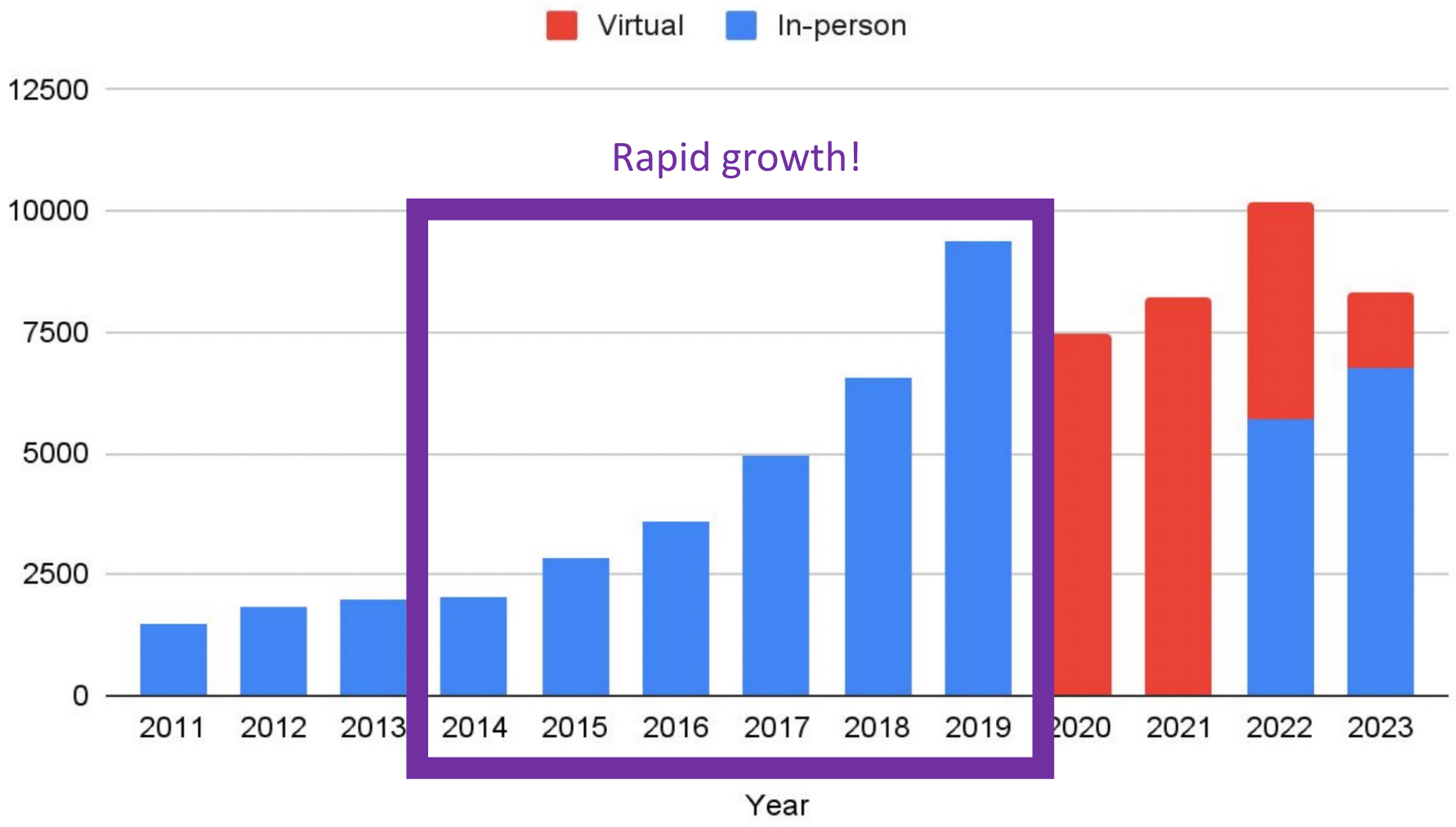
The summer vision project is an attempt to use our summer workers effectively in the construction of a significant part of a visual system. The particular task was chosen partly because it can be segmented into sub-problems which will allow individuals to work independently and yet participate in the construction of a system complex enough to be a real landmark in the development of "pattern recognition".

Emergence of Research Community



Research Community Size

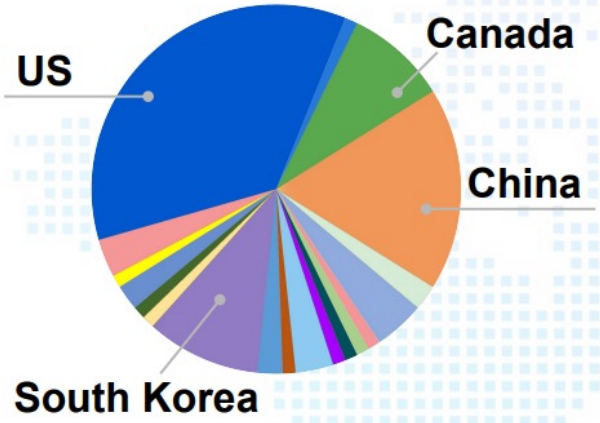
Number of CVPR 2023 attendees:



https://cvpr.thecvf.com/media/cvpr-2023/Slides/23313.pdf

Research Community Size

Attendees from 75 Countries/Regions!



US	**2779**
Canada	**672**
Mexico	25
Puerto Rico	1

Brazil	11
Colombia	9
Chile	5
Ecuador	3
Peru	3
Argentina	2

Germany	**372**
UK	**303**
Switzerland	**173**
France	**141**
Italy	85
Spain	52
Sweden	50
Netherlands	47
Poland	38
Denmark	30
Belgium	27
Czech Republic	26

Austria	16
Ireland	15
Russia	15
Turkey	13
Finland	10
Norway	8
Portugal	8
Romania	8
Serbia	7
Luxembourg	6
Hungary	5
Croatia	4

Greece	4
Estonia	3
Bulgaria	2
Slovenia	2
Ukraine	2
Iceland	1
Latvia	1
Malta	1
Slovakia	1

China	**1413**
South Korea	**815**
Japan	**304**
Singapore	**148**
Hong Kong	**119**
Israel	**118**
India	**101**
Taiwan	60
Saudi Arabia	41
UAE	34
Vietnam	16
Malaysia	8
Macau	5
Armenia	3
Qatar	3
Thailand	3
Iran	2
Oman	1

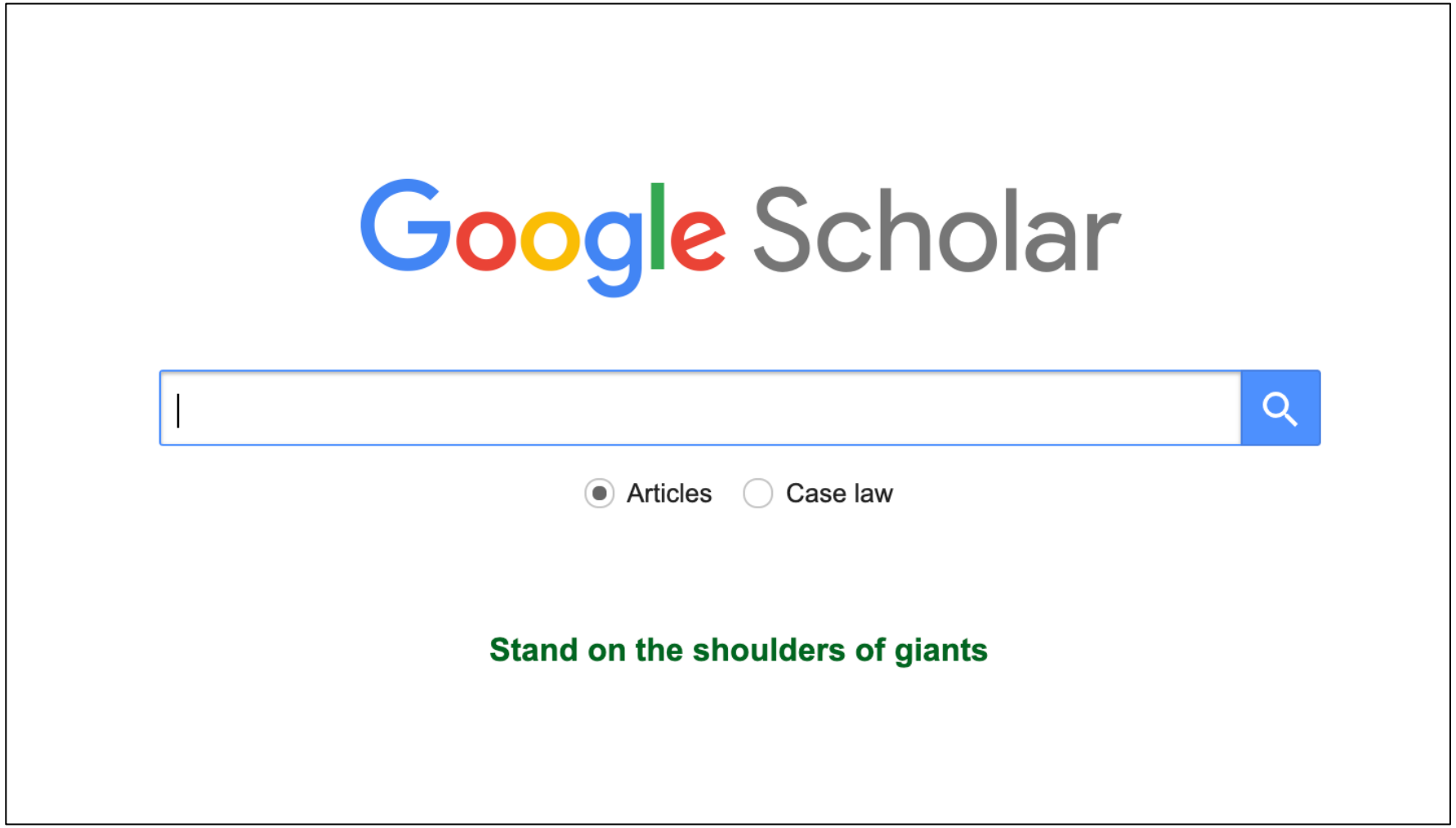
Ethiopia	16
Senegal	6
Nigeria	5
South Africa	2
Sudan	2
Algeria	1

Burundi	1
Cameroon	1
Egypt	1
Kenya	1
Libya	1
Tunisia	1

Australia	**115**
New Zealand	4

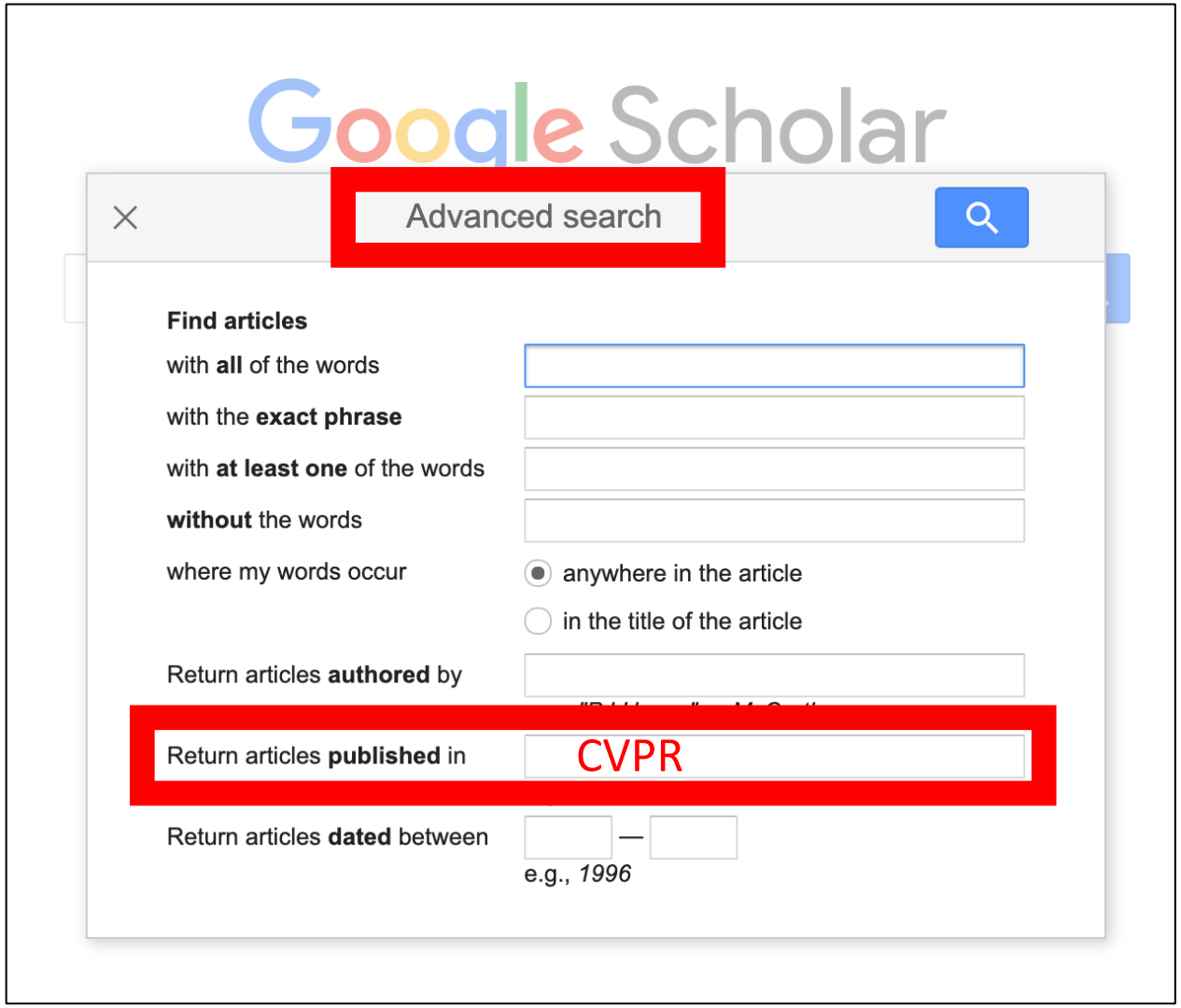
https://cvpr.thecvf.com/media/cvpr-2023/Slides/23313.pdf

Researchers' Success Metric: Publications



Tool to find publications

Researchers' Success Metric: Publications



Many Computer Vision Publication Venues Today

Conferences

- Recurring event, often annual
- Most prestigious publication venue
- Papers are length constrained

e.g., beyond CVPR, ICCV, and ECCV, also:

- Winter conference on Applications in Computer Vision (WACV)
- Asian Conference on Computer Vision (ACCV)
- British Machine Vision Conference (BMVC)
- Medical Image Computing and Computer-Assisted Intervention (MICCAI)
- Conference on Automatic Face and Gesture Recognition (IEEE FG)

Journals

- Periodical publication
- Next most prestigious publication venue
- Papers can be any length
- Often, longer review cycle than conferences

e.g.,

- International Journal of Computer Vision (IJCV)
- Transactions on Pattern Analysis and Machine Intelligence (PAMI)

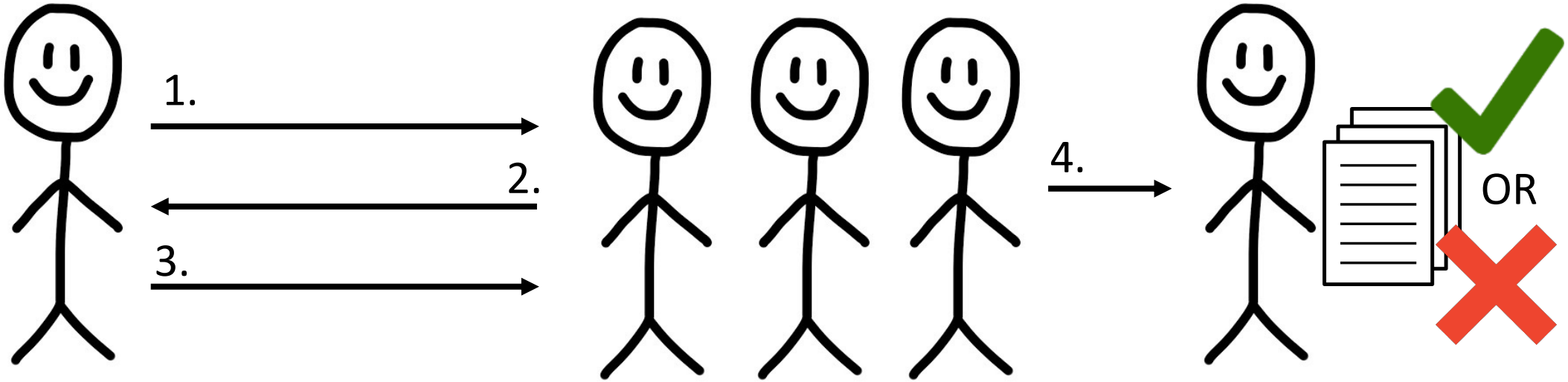
Workshops

- Typically associated with a conference, focusing on a specialized topic (some recur and even grow into conferences)
- Least prestigious publication venue
- Often, papers are length constrained
- Shorter review cycle than conferences

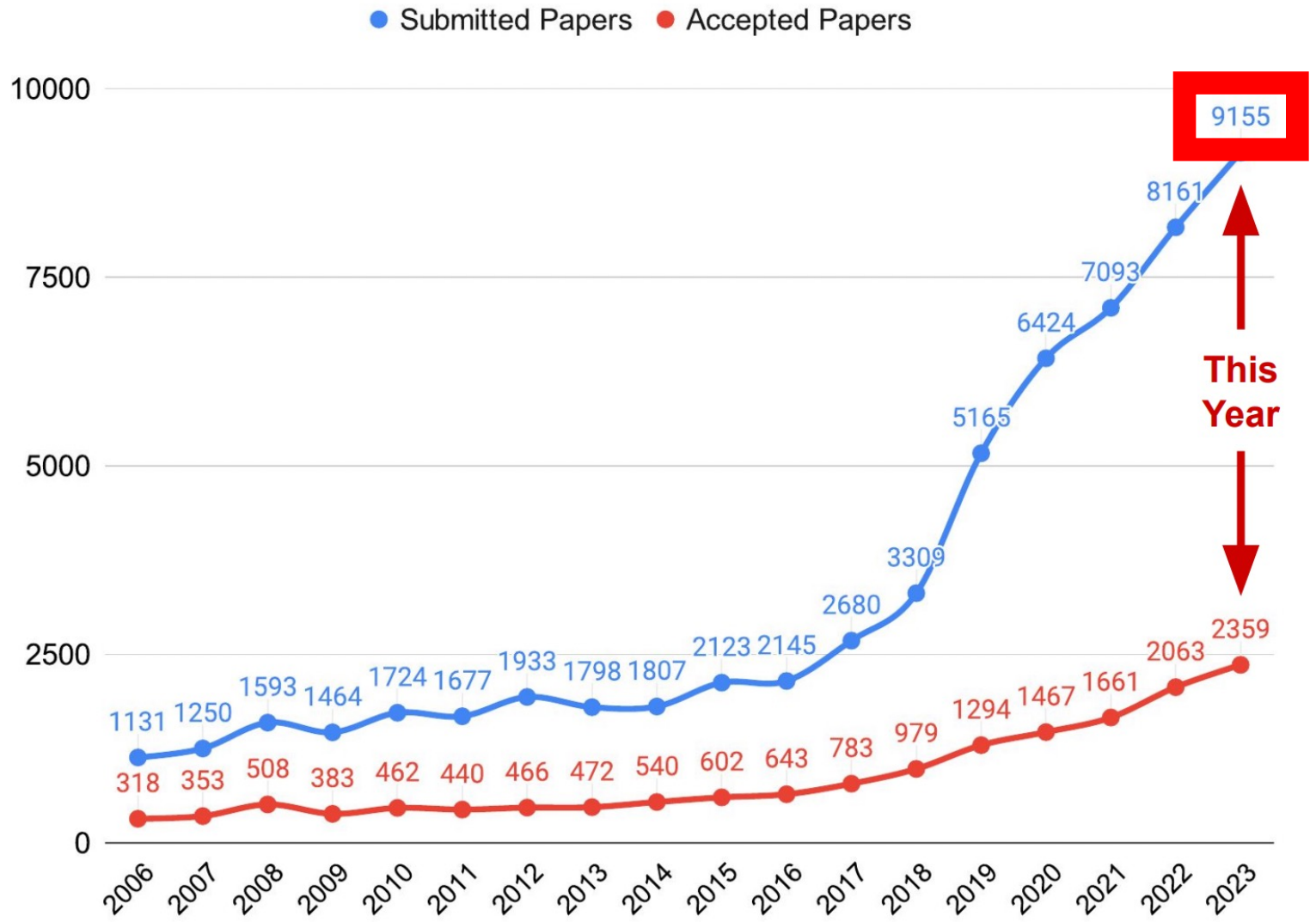
Typically, 10s associated with major conferences

Publication Process

- Typically, 1-3 years to publish in top venues (e.g., CVPR, ICCV, ECCV)
  1. Submit paper about research you conduct
  2. ~2 months later, receive reviews from at least 3 reviewers
  3. Write short response (rebuttal) to reviewers' feedback in ~1 week
  4. Meta-reviewer makes final decision ~1 month later: accept or reject
  5. If accepted, paper is published ~4 months later



Research Community Labor; e.g., CVPR



https://cvpr.thecvf.com/media/cvpr-2023/Slides/23313.pdf

Research Community Labor

Number of papers submitted to CVPR 2023: 9,155

Author labor estimate: ~$1.4 billion

- 9,155 submissions x 2 authors/submission x 1 year/author x $75,000/year = $1,373,250,000
    * median of 5 authors per paper

Reviewer labor estimate: ~$1.9 million

- 7508 researchers x 5 hr/reviewer x $50/hr = $1,877,000
    * 7,069 reviewers
    * 404 area chairs
    * 30 senior area chairs
    * 5 program chairs

https://cvpr.thecvf.com/media/cvpr-2023/Slides/23313.pdf
https://public.tableau.com/views/CVPR2023SubjectAreasbyTeamSize/Dashboard1?:showVizHome=no

Prestige of Computer Vision (Google Scholar)

Top publications

Top cited publications over the last five years Learn more

	Publication	h5-index	h5-median
1.	Nature	467	707
2.	The New England Journal of Medicine	439	876
3.	Science	424	665
4.	IEEE/CVF Conference on Computer Vision and Pattern Recognition	422	681

CVPR is 4th most impactful publication venue of all journals and conferences in all of science!

h5-index: largest number such that h articles published in 2018-2023 have at least h citations each

Prestige of Computer Vision (Google Scholar)

	Publication	h5-index	h5-median
1.	IEEE/CVF Conference on Computer Vision and Pattern Recognition	422	681
2.	Advanced Materials	326	415
3.	Neural Information Processing Systems	309	503
4.	International Conference on Learning Representations	303	563
5.	International Conference on Machine Learning	254	463
6.	Journal of Cleaner Production	246	321
7.	European Conference on Computer Vision	238	390
8.	Advanced Energy Materials	236	312
9.	IEEE Access	233	350
10.	Advanced Functional Materials	230	312
11.	IEEE/CVF International Conference on Computer Vision	228	366

In engineering and computer science, CV venues rank high for most impactful publication venues of all journals and conferences!

h5-index: largest number such that h articles published in 2018-2023 have at least h citations each

Course Focus: What is the Research Community Talking About and Where Is It Going?

Today's Topics

- Computer vision: origins
- What makes computer vision hard?
- Research in computer vision
- Course logistics

Course Objectives

- Understand core computer vision problems and their typical solutions:

    1. Recognize and define core computer vision problems

    2. Identify benchmarks used by the research community to track progress on each problem (i.e., data source, data annotation process, evaluation metrics)

    3. Identify types of algorithms commonly used to solve each problem alongside their general properties that make them well-suited for the problem

Course Objectives

- Analyze and present cutting-edge research:

    1. Identify in research papers the novelty claims, mechanisms used to validate the claims (e.g., theories and experiments), and why the papers' contributions matter to society

    2. Deliver oral presentations that explain research papers

    3. Discuss the merits and limitations of research papers

    4. Prepare and moderate discussions about research papers

Course Objectives

- Conduct research to enhance your expertise on a topic of your choice:

1. Design and execute a research project on a novel idea involving computer vision, such as analyzing an existing method, implementing/evaluating a new method, surveying the state-of-art for a specific problem, or outlining how to tackle a new computer vision problem

2. Deliver an oral presentation that explains the research

3. Review fellow students' presented research and provide constructive feedback

4. Communicate the research through a final report

Q&A: “Do I have the appropriate pre-requisites/background?”

- Familiarity with machine learning required, and ideally with neural networks

Note: It is common to feel an “imposter syndrome” in this course, as research papers are incredibly dense with many unexplained details. If you feel this, know that you are likely in good company!

Q&A: “What are required textbooks?”

- None. Links to required readings will be posted on the course website for each class meeting.

Q&A: “What are the assignments?”

- Reading assignments most weeks (first assignment due next week)
- Student-led presentations about computer vision research papers
- Final project: research on a topic of your choice
- Late policy
  - Penalized 1% of grade per hour for up to 2 hours
  - No credit if more than 2 hours late

Q&A: “How is my final grade determined?”

	% of Final Class Grade
Class Participation	10%
Reading Assignments	30%
Student-Led Lecture	30%
Final Project	30%

Student-Led Lectures

- Process (described in syllabus and on course website):
  - **3rd week**: assigned topic with partner
  - **2+ weeks prior to lecture** (10%): schedule meeting with me to review 4-6 recent publications at a premiere computer vision conference (e.g., CVPR, ICCV,ECCV) that you will share 48+ hours beforehand; we will assign 1 as required reading and 1 as optional reading
  - **At least 1 week prior to first lecture** (30%): schedule meeting with me to review a completed draft of the lecture slides that you will share 24+ hours before this meeting
  - **Lecture** (60%): ~50-minute presentation followed by ~25-minute discussion facilitated by instructor based on student-submitted discussion points
  - **Partner relationships**: you can divide and conquer or do everything together; up to you!

Course Topics: Tentative Schedule

Week	Topic(s)
Background	
1	Introduction
2	Object Recognition
3	Image Classification
4	Object Detection
5	Semantic Segmentation, Image Captioning
6	Foundation Models, Prompts

Week	Topic(s)
7-12: student-led presentations	
13	Image Synthesis and Video Analysis
14	Efficient and Responsible Computer Vision
15	Responsible Computer Vision

Course Resources

- Website:
  - https://home.cs.colorado.edu/~DrG/Courses/RecentAdvancesInComputerVision/AboutCourse.html
- Syllabus:
  - https://home.cs.colorado.edu/~DrG/Courses/RecentAdvancesInComputerVision/Syllabus.pdf

Instructor: Danna Gurari; aka, Dr. G

(preferred pronouns: she/her)

My Experience Relating to Computer Vision

2004-2005

Masters student designing system to record ultrasound images

2005-2007

Software engineer helping to record satellite images

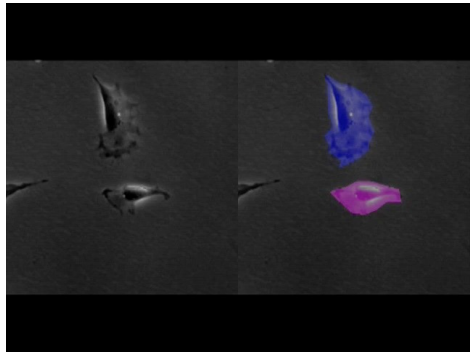
2007-2010

Software developer and project manager helping to record visible and infrared video

Source: Boulder Imaging

2010-2015

PhD student designing methods to segment and track cells in images and videos



2015-Present

Postdoctoral fellow and then assistant professor overseeing research projects related to many computer vision problems

e.g., image classification, object detection, semantic segmentation, object tracking, image captioning, visual question answering, style transfer, image inpainting, and image search

My Experience Relating to Computer Vision

2004-2005 | 2005-2007 | 2007-2010 | 2010-2015 | 2015-Present

Masters student designing system to record ultrasound images

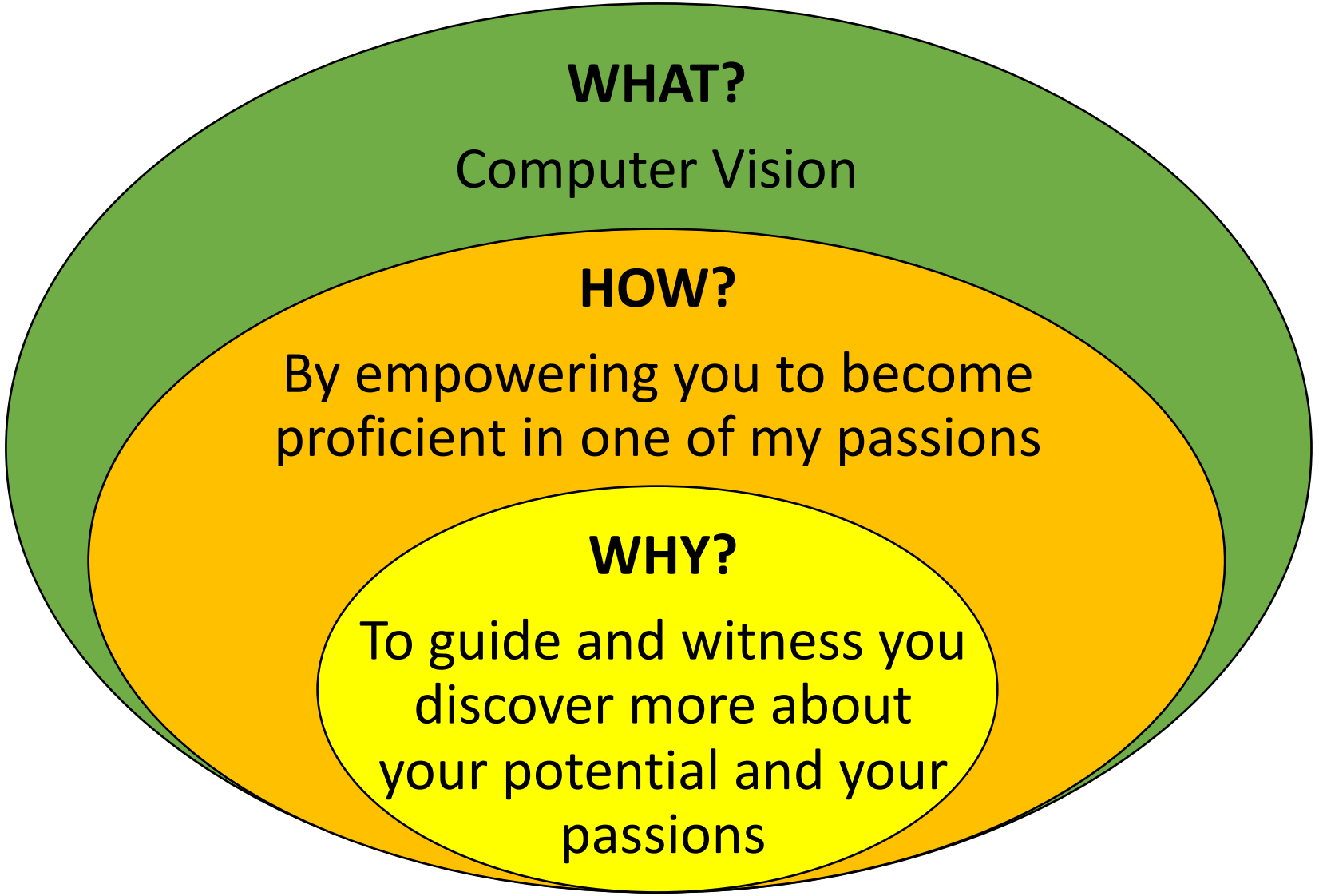
Software engineer helping to record satellite images

Software developer and project manager helping to record visib and infrared video

Source: Boulder Imaging

54 publications, most involving computer vision

What is My “Why” for Teaching You…



Today's Topics

- Computer vision: origins
- What makes computer vision hard?
- Research in computer vision
- Course logistics


The End