Object recognition and computer vision 2023
Reconnaissance d’objets et vision artificielle (RecVis)

Class logistics

Lecturers:

Gül Varol  Jean Ponce  Armand Joulin  Josef Sivic  Ivan Laptev  Cordelia Schmid  Mathieu Aubry

http://imagine.enpc.fr/~varolg/teaching/recvis23/
Object recognition and computer vision 2023

Lecturers:

Gül Varol  Jean Ponce  Armand Joulin  Josef Sivic  Ivan Laptev  Cordelia Schmid  Mathieu Aubry

TAs:

Ricardo Garcia  Guillaume Le Moing  Charles Raude

http://imagine.enpc.fr/~varolg/teaching/recvis23/
## Schedule

**Tuesdays 16h - 19h**  
**Location: Salle Dussane**

<table>
<thead>
<tr>
<th>#</th>
<th>Date</th>
<th>Lecturer</th>
<th>Topic and reading materials</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Oct 3</td>
<td>Gül Varol, Jean Ponce</td>
<td>Class logistics; assignments, final projects, grading (G. Varol); Introduction to visual recognition; Camera geometry; Image processing (J. Ponce)</td>
</tr>
<tr>
<td>2</td>
<td>Oct 10</td>
<td>Gül Varol</td>
<td>Instance-level recognition: local invariant features, correspondence, image matching materials Assignment 1 out.</td>
</tr>
<tr>
<td>3</td>
<td>Oct 17</td>
<td>TAs</td>
<td>Python/Pytorch tutorial. Attendance is optional.</td>
</tr>
<tr>
<td>4</td>
<td>Oct 24</td>
<td>Armand Jaulin</td>
<td>Supervised learning and deep learning; Optimization and regularization for neural networks; Introduction to sequence models Assignment 1 due. Assignment 2 out.</td>
</tr>
<tr>
<td>5</td>
<td>Oct 31</td>
<td>Gül Varol</td>
<td>Neural networks for visual recognition: CNNs and image classification</td>
</tr>
<tr>
<td>6</td>
<td>Nov 7</td>
<td>Gül Varol</td>
<td>Beyond CNNs: Transformers; Beyond classification: Object detection; Pose estimation; Segmentaion</td>
</tr>
<tr>
<td>7</td>
<td>Nov 14</td>
<td>Josef Sivic</td>
<td>Large-scale image and video search Assignment 2 due. Assignment 3 out.</td>
</tr>
<tr>
<td>8</td>
<td>Nov 21</td>
<td>Gül Varol</td>
<td>Generative models: VAEs, GANs, diffusion; Vision &amp; language</td>
</tr>
<tr>
<td>9</td>
<td>Nov 28</td>
<td>Ivan Laptev</td>
<td>Weakly-supervised learning; Self-supervised learning; Vision for robotics Assignment 3 due. Final project topics are out.</td>
</tr>
<tr>
<td>10</td>
<td>Dec 5</td>
<td>Cordelia Schmid</td>
<td>Human action recognition in videos Final project proposal due.</td>
</tr>
<tr>
<td>11</td>
<td>Dec 12</td>
<td>Mathieu Aubry</td>
<td>3D computer vision</td>
</tr>
<tr>
<td>12</td>
<td>Jan 8 - Jan 9</td>
<td></td>
<td>Final project presentations The presentations may be virtual. Instructions will be provided. Final project reports due on 15/01.</td>
</tr>
</tbody>
</table>

---

Follow updates & exceptions on class webpage !!!!
Practical information: Participation

Class webpage: http://imagine.enpc.fr/~varolg/teaching/recvis23/
Google Classroom: Register with the code wbj5g7w to receive announcements.

Time: 16h00-19h00, Tuesdays, starting Oct 3
Location: Salle Dussane, ENS Ulm, 45 rue d’Ulm 75005, Paris
Format: In-person lectures. Slides provided after each lecture.

For externals: You are welcome to attend the course (either for auditing or validation) provided there are enough free places in the lecture hall. If your school requires a proof of attendance, you need to get signatures from teachers after every lecture.
3 programming assignments (50%)

• A1: Instance-level recognition
• A2: Neural networks
• A3: Image classification competition

Some experience with Python (numpy, pytorch) will be useful, but we will provide an optional crash-lecture on Python/Pytorch for computer vision

Final project (50%)

• More independent work, resulting in a report and a class presentation.
• We will provide Google Cloud credits for each student.

Policy

Assignments are strictly individual

Copy-paste of the code, results, parts of the report → 0p.

FPs can be done in groups of max 2 people
Assignment I: Instance-level recognition

- Part I: Sparse features for matching specific objects in images
  - Feature detector and descriptor
  - Robust match filtering techniques
  - Augmented reality
- Part II: Compact descriptors for image retrieval
Assignment II: Neural networks

- Part 1: Neural Network’s theory:
  - Forward pass, Backward pass
  - Parameter update
- Part 2: Building blocks of convolutional neural networks
- Part 3: Training a CNN on CIFAR-10 dataset with PyTorch
Assignment III: Image classification competition

- Class Kaggle competition
- Example task: Bird image classification - the assignment will cover a similar task
Final project

• Can be done individually or as a **group of max 2 people**
• The proposed project topics are from the recent top-conference publications in computer vision, see example topics from 2022 here: [https://www.di.ens.fr/willow/teaching/recvis22/](https://www.di.ens.fr/willow/teaching/recvis22/)
• Student-defined projects are welcome.
• Final project can be joint with another MVA course.
• We arranged $100 Google Cloud credits for the project.
  ◦ This will be announced through Google Classroom before projects start

• Select the topic + write project proposal
• Present the work in the class
• Write project report
Practical: Python tutorial

Fill-in the Python tutorial participation form linked from the class webpage by Mon Oct 16.

The tutorial will be on Tue Oct 17, starting at 16h00 (until 18h00) at:
INRIA/Willow, 2 rue Simone IFF, 75012 Paris.

Note there will be no lecture on that day.

Who should participate?
- Students with no or limited experience with Python. Attendance is optional.

Topics covered:
- Installing Anaconda.
- Brief introduction to Python.
- Introduction to Numpy, PyTorch for computer vision.
- Using Jupyter notebooks.
Course outline

1. Instance-level recognition
   • Camera geometry
   • Image processing
   • Image correspondence

2. Category-level recognition
   • Supervised learning
   • Neural networks for visual recognition
   • Object recognition, detection, and segmentation

3. Advanced topics
   • Large-scale image and video search
   • Generative models
   • Vision for robotics
   • Human action recognition in videos
   • 3D computer vision

Vibrant portrait painting of Salvador Dali with a robotic half face
Course outline

1. Instance-level recognition
   - Camera geometry
   - Image processing
   - Image correspondence

2. Category-level recognition
   - Supervised learning
   - Neural networks for visual recognition
   - Object recognition, detection, and segmentation

3. Advanced topics
   - Large-scale image and video search
   - Generative models
   - Vision for robotics
   - Human action recognition in videos
   - 3D computer vision
Course outline

1. Instance-level recognition
   • Camera geometry
   • Image processing
   • Image correspondence

2. Category-level recognition
   • Supervised learning
   • Neural networks for visual recognition
   • Object recognition, detection, and segmentation

3. Advanced topics
   • Large-scale image and video search
   • Generative models
   • Vision for robotics
   • Human action recognition in videos
   • 3D computer vision

Vibrant portrait painting of Salvador Dali with a robotic half face

Boiling eggs
3. Advanced topics

- Large-scale image and video search
- Generative models
- Vision for robotics
- Human action recognition in videos
- 3D computer vision
Recap:

1. Register on the Google Classroom
   • Assignment submissions, discussions and announcements will be done on Google Classroom.
   • Assignment 1 – Instance-level recognition (due Oct 24 2023)

2. Fill-in Python tutorial participation form (by Oct 16)
IMAGINE and WILLOW teams are active in computer vision.

http://imagine.enpc.fr

http://www.di.ens.fr/willow/

There will be master internships available. Talk to us if you are interested!