EECS 598-012: Unsupervised Visual Learning

Instructor: Andrew Owens Winter 2021

Tentative schedule

Note: The schedule and syllabus are likely to change significantly over the coming week!

Lecture	Date	Topic	Materials
Lec. 1	Wed, Jan 20	**Introduction** About the course Unsupervised Learning	• Smith & Gasser: Six Lessons from Babies (optional)
		Generative models	
Lec. 2	Mon, Jan 25	**Energy-based models**	• Song & Kingma: How to Train Your Energy-Based Models • Grathwohl et al: Your classifier is secretly an energy based model and you should treat it like one
Lec. 3	Wed, Jan 27	**Variational autoencoders**	• van den Oord et al.: Neural Discrete Representation Learning • Vahdat & Kautz, NVAE: A Deep Hierarchical Variational Autoencoder
Lec. 4	Mon, Feb 1	**Normalizing flows**	
Lec. 5	Wed, Feb 3	**GANs**	• Large Scale Adversarial Representation Learning
Lec. 6	Mon, Feb 8	**Autoregressive models**	• Pixel-CNN • Generative pretraining from pixels
		Discriminative methods	
Lec. 7	Wed, Feb 10	**Pretext tasks**	• Unsupervised Visual Representation Learning by Context Prediction • Unsupervised Representation Learning By Predicting Image Rotations • Colorful Image Colorization
Lec. 8	Mon, Feb 15	**Contrastive learning**	• Representation Learning with Contrastive Predictive Coding • He et al.: Momentum Contrast for Unsupervised Visual Representation Learning • Contrastive Multiview Coding
		Learning from non-visual signals	
Lec. 9	Wed, Feb 17	**Language & vision**	• VirTex: Learning Visual Representations from Textual Annotations • Radford et al: Learning Transferable Visual Models From Natural Language Supervision
Lec. 10	Mon, Feb 22	**Sound & vision**	• Afouras et al.: Self-Supervised Learning Of Audio-Visual Objects From Video • Asano, Patrick et al.: Labelling unlabelled videos from scratch with multi-modal self-supervision
		Video	
Lec. 11	Mon, Mar 1	**Motion estimation**	• Jabri et al.: Space-Time Correspondence as a Contrastive Random Walk • Jonschkowski et al.: What Matters in Unsupervised Optical Flow
Lec. 12	Wed, Mar 3	**Forecasting**	• High Fidelity Video Prediction with Large Stochastic Recurrent Neural Networks • Oops! Predicting Unintentional Action in Video
Lec. 13	Mon, Mar 8	**3D reconstruction**	• Unsupervised Learning of Probably Symmetric Deformable 3D Objects from Images in the Wild • Competitive Collaboration: Joint Unsupervised Learning of Depth, Camera Motion, Optical Flow and Motion Segmentation
		Advances in deep learning	
Lec. 14	Wed, Mar 10	**Attention**	• An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale • End-to-End Object Detection with Transformers
Lec. 15	Mon, Mar 15	**Optimization**	• Martens & Grosse: K-Fac
Lec. 16	Mon, Mar 22	**Theory of self-supervision**	
		Learning with less supervision	
Lec. 17	Wed, Mar 24	**Gradient-based meta-learning**	• Finn et al.: Model-Agnostic Meta-Learning for Fast Adaptation of Deep Networks • Wu, Ren et al.: Understanding Short-Horizon Bias In Stochastic Meta-Optimization • Maclaurin et al: Gradient-based hyperparameter optimization through reversible learning (optional)
Lec. 18	Mon, Mar 29	**More meta-learning**	• DARTS: Differentiable Architecture Search • Lorraine et al: Optimizing Millions of Hyperparameters by Implicit Differentiation

Lecture	Date	Topic	Materials
Lec. 19	Wed, Mar 31	**Architecture search**	
Lec. 20	Mon, Apr 5	**Semi- and weakly-supervised learning**	• Self-training with Noisy Student improves ImageNet classification (optional) • Meta Pseudo Labels • Evaluating Weakly Supervised Object Localization Methods Right
		Embodied vision	
Lec. 21	Wed, Apr 7	**Representation learning for action**	
Lec. 22	Mon, Apr 12	**Novelty**	
		Advances in deep learning	
Lec. 23	Wed, Apr 14	**Human vision**	
Lec. 24	Mon, Apr 19	**Project presentations**	
Lec. 25	Wed, Apr 21	**Project presentations**	

Staff & Office Hours

Andrew Owens
Instructor

Xixi Hu
GSI

Name	Office hours time
Andrew Owens	Monday 4:45 - 5:30pm (starting Jan 25)
Xixi Hu	TBD

Office hours will take place over video chat, using the same Zoom link as lecture.

Course information

Today's computer vision systems largely rely on supervision from humans, such as object labels, to learn about the world. This course will discuss recent efforts to create methods that avoid the need for this supervision by learning from unlabeled sensory data. Topics will include: deep generative models, self-supervised learning, multimodal models, learning from video, and semi-supervised learning. We will also cover recent advances in deep learning, such as meta-learning and network architectures, that support the goal of unsupervised learning. This is a seminar-based graduate-level class covering very recent advances in unsupervised learning in computer vision. The main focus of the class will be on *reading and critiquing recent research papers*. In each lecture, students will present and critique several recent research papers. You will also explore these ideas via a self-directed project. While the class is *not* intended to be an introduction to unsupervised learning or deep learning, we will give a single problem set reviewing some of the core concepts, which will be due approximately 2/3 of the way through the semester.

Lectures: Lectures will take place over Zoom on Monday and Wednesday, 3:00 - 4:30pm. Since this is a discussion-based class, *your attendance is required*. Missing more than two classes without an excuse will negatively affect your grade. Recordings will only be provided to students enrolled in the class.

Prerequisites: This is an advanced vision course. Students are expected to have taken an introductory vision course before enrolling (EECS 442, 504, or equivalent), so that they will be prepared to read and discuss recent research.

Paper reviews: You'll be required to submit one short paper review each week, beginning the week of Lecture 2. Your review should be based on the paper itself, rather than the discussion. It is therefore due *before* the paper is presented in class (i.e. at 3pm on Monday or Wednesday).

- Summarize the paper. For most papers, this means explaining technical contributions, such as key mathematical insights, algorithms, and architectures.
- Briefly explain how the paper relates to previous work, and why its contributions might be (or might not be) important.

- Summarize the key experiments.
- Discuss the paper's shortcomings: e.g. limitations to the methods, unconvincing aspects of experiments, presentation issues.

Reviews will be graded as: ✓+, ✓, ✓−, **0**. We will not accept late submissions without a valid excuse. However, we will *drop your 2 lowest review scores*.

Q&A: This course has a Piazza forum, where you can ask public questions. We also appreciate it when you respond to questions from other students! If you have an important question that you would prefer to discuss over email, you may email the course staff (eecs442-fa20-staff@umich.edu), or you can contact the instructor by email directly.

Textbooks: In this class, we'll mostly be reading research papers, rather than textbooks. The following might be useful as reference, though:

- Goodfellow, Bengio, Courville. *Deep Learning*. (available for free online)
- Szeliski. *Computer Vision: Algorithms and Applications, 2nd edition draft* (available for free online)

If you have feedback for the author of the Szeliski book draft, please submit it here, and we'll pass it along!

Grading: Final grades will be computed as follows:

Final project	45%
Class presentation	25%
Problem set	20%
Participation & reviews	10%

Academic integrity: While you are encouraged to discuss homework assignment with other students, *your programming work must be completed individually*. You must also write up your solution on your own. You may not search for solutions online, or use existing implementations of the algorithms. Please see the Michigan engineering honor code for more information.

Support: The counseling and psychological services center (CAPS) provides support for a variety of issues, including mental health and stress.

Presentation guidelines

You will be in charge of teaching one class, as part of a group of 3 people (starting Lec. 5). Each class will be organized around a topic of ongoing research. We'll send up a sign-up sheet after the first class, where you will rank

Organization: We suggest organizing most classes as follows:

1. Background (25 mins)
2. Paper 1 (25 mins)
3. Paper 2 (25 mins)
4. Discussion (5 mins)

The *background* section is usually the most important part of the class. It should resemble a mini-lecture, covering the "basics" that students will need to understand the paper presentations. For example, if the class is covering papers about variational autoencoders (VAEs), this section should review what a VAE is, and it should touch on any relevant findings that are necessary to understand the papers. Often, this will involve also describing prior attempts to solve the problems that the (much more recent) papers address.

Each *paper* section should be a *critical* presentation the work in the paper. You should explain what problem the researchers were addressing, their motivation for what their solution was, and how well they succeeded at that goal. Unlike introductory courses, where methods are largely well-understood and have passed the "test of time", the papers in this class will often have important limitations. We therefore encourage you to take a critical approach to reading the papers, and to describe possible shortcomings. We also encourage you to discuss things in the paper that you do not think were well-justified, and choices by the authors that you did not understand.

Finally, for the (optional) *discussion*, you will lead a brief interactive session, where students can debate the issues at stake in the papers. For example, you might run a Q&A session where you ask: *should we really consider language-based supervision to be "unsupervised"*, or *do we need to interact with the world to learn good representations?*. If you'd like the section to be particularly interactive, you can also do this using Zoom breakout sessions.

Slides: You are allowed to use existing slides and figures, but please clearly credit the authors. Please submit your slides to us in PDF form. By default, we will post your slides only on Canvas, so that they are only visible to those enrolled in the class. Howeve, we'd also be happy to post them publicly if you'd like.

Signing up: We'll assign people to presentation timeslots in two phases (i.e. the first and second halves of the class). You'll fill out a questionaire indicating which classes you'd like to participate in. If you happen to have a group of 3 in mind already, please indicate this on the form, and we will try to assign you to a single topic (we unfortunately cannot accommodate groups with other sizes).

Project guidelines

You'll do a self-directed group project, due at the very end of the course. Groups should be at most 4 students, unless you are given permission from the instructor. Deliverables include:

- Project proposal (due halfway through the semester).
- Report (4 pages in CVPR format)
- Presentation (a 5-min talk)

We'll provide more details as the semester progresses.