How research gets done part I

- This mini-series:
  - Aims to give you a feel for how research in deep learning gets done
  - Can guide your explorations
  - Aims to debunk and demystify

Step 1 of deep learning research:

Get a solid understanding of the fundamentals. This course is the perfect way to do so.
This means:

Aim to understand both the theoretical parts and slightly more importantly the practical parts.
Begin to read papers. It matters little (at first) which ones, just read what you find exciting.
While at first they might be hard to understand, soon you will understand more and more.

How research gets done part II

Step 2 of deep learning research

[Recap: step 1: understand fundamentals, read papers.]

How to read papers?

Quick advice: think in terms of "passes":

1st pass: Title -> abstract -> figures/tables -> conclusion -> Introduction

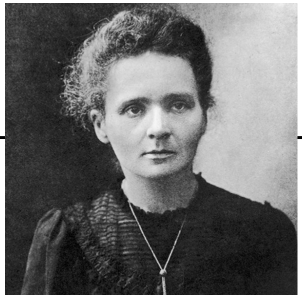
2nd pass: Intro->…->Conclusion, but skip details/don't try to understand maths

3rd pass: Try to recap what you didn't understand, reread those parts, be critical.

…. Dive into the code

After every pass you can drop out. Which is good. No need to detail-read *every* paper.

How research gets done part III



Step 3 of deep learning research

[Recap: step 1&2: understand fundamentals, read papers.]

Next, start reproducing results and playing around.

After reading an interesting paper, try to implement it from scratch.

Compare against the open-source version of the authors/other researchers.

Try to reproduce numbers and get familiar with tricks and hacks.

How research gets done part 4

Previous parts:
[fundamental understanding, read papers, how-to-read-papers, implement & tinker with code]

Today:

- Be curious.
- **“The most exciting phrase to hear in science, the one that heralds new discoveries, is not ‘Eureka!’ but ‘That’s funny…’” --- *Isaac Asimov***
- “That’s funny…” --> there’s a delta between what’s expect *vs.* what’s observed.
  - For this: need to be able to trust a) your experiment, b) understand broader context to form expectation
  - (see also incongruity theory re: why jokes are funny)
  - (see also Hubel & Wiesel from earlier)
- For this you need to be able to hold two opposing thoughts/ideas in your head (at least for a while)
- The advice: Don’t ignore this moment. Instead follow this up by *structured tinkering (next part)*

[kinda like this]

How research gets done part 5

Previous parts:
[fundamental understanding, read papers, how-to-read-papers, implement & tinker with code, realise and seek *funny* moments]

Today:

- We found something funny (or we have an inkling that there's something interesting here)
- How do we analyse this further?
  1. Establish benchmarks (we cannot know what we cannot measure)
  2. Establish baselines (is 58% good or bad?)
  3. Figure out what the *minimum viable proof of principle* is: if this works, it shows our idea is right
  4. Compare and contrast this to existing ideas:
     1. Why might it work?
     2. Why not?
     3. What are the *principles* at play here?
  5. Next times: when to give up and when not, how to design ablations

How research gets done part 6

"I was lucky…"

Previous parts:
[fundamental understanding/read papers, how-to-read-papers, implement & tinker with code, realise and seek *funny* moments, MVP/principles/benchmarks/baselines]

Today:

- When to give up and try something else?
  - Impossible to answer: different for everyone.
  - Possible factors to consider:
    - impact *vs.* work-required tradeoff
    - Amount of fun-while-working-on-it
    - Existence of small progresses
    - Opportunity costs (*what could you be doing instead)*
      - These can be big in a field like deep learning
- The more familiar you become in a topic, the more ideas you will get
- Having ideas is almost never the bottleneck after some point
- Evaluating the quality of ideas… requires intuition, which is developed with time

How research gets done part 8

Katalin Karikó

Previous parts:
[fundamental understanding/read papers, how-to-read-papers, implement & tinker with code, realise and seek *funny* moments, MVP/principles/benchmarks/baselines, when to (not) give up/impact-vs-work]

Today:

- Ok, so you didn't give up and you're on to something non-trivial.
- Next: How do you show/ analyze what's happening or why your method is better?
- Answer: Ablations
  - The key idea is to "only vary one thing at a time"
  - (Same principle behind when designing experiments in the investigation phase!)
  - Never change two things at the same time, you won't know if it was A or B that helped
  - Some examples:
- Show simple, easy to understand cases (sometimes toy examples)
- **One idea per paper!**

Table 3: **Ablation** of multi-modality, Modality Alignment and Gaussian marginals. Decorrelated Heads. Models are evaluated at 75 epochs on the VGG-Sound dataset.

Method	(image)	(audio video)	MA?	G.?	DH?	**Acc**	**ARI**	**NMI**
(a) SeLa	✓	✗	–	–	–	6.4	2.3	20.6
(b) Concat	✗	✓	–	✗	✗	7.6	3.2	24.7
(c) SeLaVi	✗	✓	✗	✗	✗	24.6	15.6	48.8
(d) SeLaVi	✗	✓	✗	✓	✓	26.6	18.5	50.9
(e) SeLaVi	✗	✓	✓	✗	✓	26.2	17.3	51.5
(f) SeLaVi	✗	✓	✓	✓	✗	23.9	14.7	49.9
(g) **SeLaVi**	✗	✓	✓	✓	✓	26.6	17.7	51.1

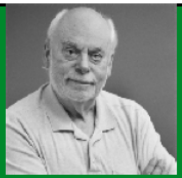
net-depth-features	AP	AP_{50}	AP_{75}
ResNet-50-C4	30.3	51.2	31.5
ResNet-101-C4	32.7	54.2	34.3
ResNet-50-FPN	33.6	55.2	35.3
ResNet-101-FPN	35.4	57.3	37.5
ResNeXt-101-FPN	**36.7**	**59.5**	**38.9**

(a) **Backbone Architecture**: Better backbones bring expected gains: deeper networks do better, FPN outperforms C4 features, and ResNeXt improves on ResNet.

	AP	AP_{50}	AP_{75}
softmax	24.8	44.1	25.1
sigmoid	**30.3**	**51.2**	**31.5**
	+5.5	*+7.1*	*+6.4*

(b) **Multinomial *vs.* Independent Masks** (ResNet-50-C4): *Decoupling* via per-class binary masks (sigmoid) gives large gains over multinomial masks (softmax).

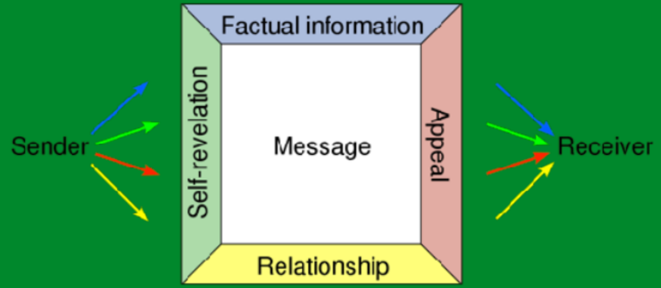
How research gets done: part 9



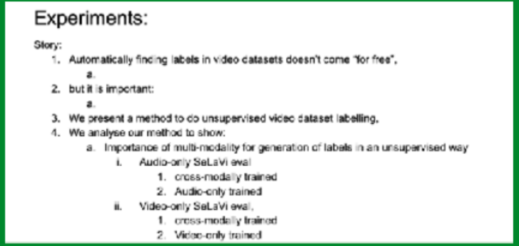
Fraser Stoddart:
"You've got to break the rules"

Previous parts:
[fundamental understanding/read papers, how-to-read-papers, implement & tinker with code, realise and seek *funny* moments, MVP/principles/benchmarks/baselines, when to (not) give up/impact-vs-work, importance of Ablations]

- Ideally before or latest when all previous steps are (more or less) completed we develop the *storyline*
- Why a story? Aren't we writing a hard, cold, scientific paper?
  - Yes, but: (science) communication not as easy:
  - So we need to put in a lot of work
  - What's the *rode draad*/overarching motif?

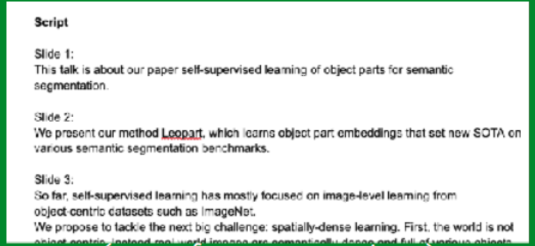


- Use google docs, don't make it super nice, just re-iterate from scratch multiple times.



Experiments:

Story:
1. Automatically finding labels in video datasets doesn't come "for free",
   a.
2. but it is important:
   a.
3. We present a method to do unsupervised video dataset labelling.
4. We analyse our method to show:
   a. Importance of multi-modality for generation of labels in an unsupervised way
      i. Audio-only SeLaVi eval
         1. cross-modally trained
         2. Audio-only trained
      ii. Video-only SeLaVi eval.
         1. cross-modally trained
         2. Video-only trained



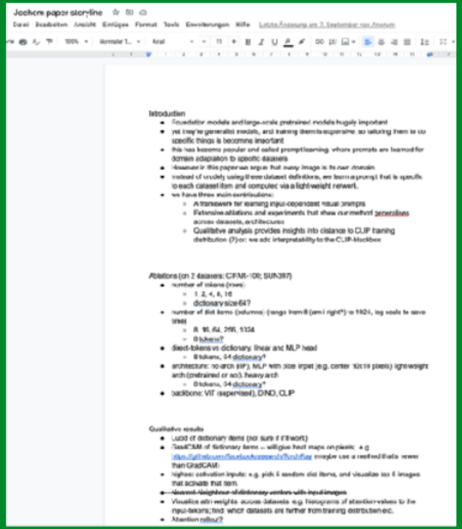
Script

Slide 1:
This talk is about our paper self-supervised learning of object parts for semantic segmentation.

Slide 2:
We present our method Leopart, which learns object part embeddings that set new SOTA on various semantic segmentation benchmarks.

Slide 3:
So far, self-supervised learning has mostly focused on image-level learning from object-centric datasets such as ImageNet.
We propose to tackle the next big challenge: spatially-dense learning. First, the world is not

(same for presentations)



- Best thing: you'll discover important missing experiments & have the introduction part of paper almost done

How research gets done: part 10

You

Previous parts:
[fundamental understanding/read papers, how-to-read-papers, implement & tinker with code, realise and seek *funny* moments, MVP/principles/benchmarks/baselines, when to (not) give up/impact-vs-work, importance of Ablations, storyline]

- So you have made great research and submitted your paper to some venue
- Congratulations. Please celebrate this.
  - Acceptance of a paper is sometimes stochastic, but finishing a piece of work needs to be celebrated on its own
- Take a break (it will allow you to see things from a new perspective)
- Alongside an ML paper, we
  - Publish the code on github, please follow reproducibility guidelines: https://github.com/paperswithcode/releasing-research-code
  - Make a website for a paper:
    examples: https://richzhang.github.io/colorization/ https://single-image-distill.github.io/ https://www.di.ens.fr/willow/research/mil-nce/ https://www.matthewtancik.com/nerf
  - Sometimes make a twitter thread about it, or write a layperson-blogpost about it
- All in order to increase the accessibility and reach of our research, because
  - The field moves so quickly it's hard to keep track but,
  - Research should be accessible, available and understandable