

9th December 2020 **#VDLM**

virtual edition



The Organizers:





Alex Schindler AIT & TU Wien



Jan Schlüter JKU Linz



René Donner contextflow

Agenda for Today



Welcome & Introduction

Announcements (Jobs, Events)

Practical experiences in accurate video segmentation

by Antonis Makropoulos, Head of Machine Learning at contextflow

Latest News & Hot Topics in Deep Learning

by René Donner and Michael Pieler

Networking in Breakout-Rooms

Announcements

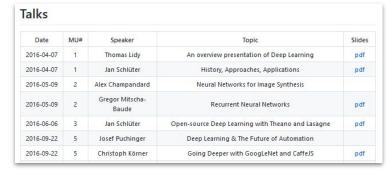


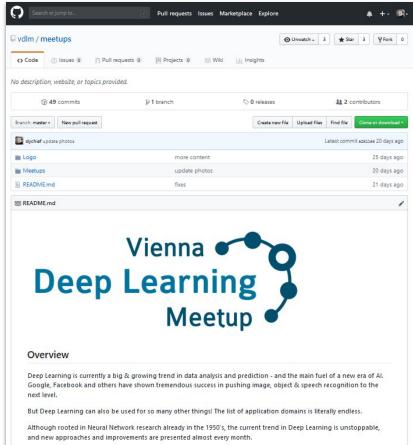
VDLM on Github

https://github.com/vdlm/meetups

- all talks
- slides
- photos
- videos
- Wiki

#	Date	Place	Topic	Link	Video	Meetup.con
1	2016-04-07	Sector 5	intro	more		link
2	2016-05-09	Sector 5		more		link
3	2016-06-06	Sector 5		more		link
4	2016-07-07	TU Wien		more		link
5	2016-09-22	Automic Software GmbH		more		link
6	2016-10-12	Sector 5		more		link
7	2016-12-01	Agentur Virtual Identity		more		link
8	2017-01-17	TU Wien Informatik		more		link





Meetup

Pre- & PostDoc Position: Machine Learning & Data Science (40h)



At the **Institute of Creative\Media/Technologies** (IC\M/T)

- 40+ researchers (pre/postdoc level)
- Our focus: Interdisciplinary research on human-centered interactive technologies and time-based media (http://icmt.fhstp.ac.at)
- Basic and applied research (e.g. FWF, FFG, EU Projects)
- Long-term perspective (no contract limitation as in other universities)
- Just 30 minutes from Vienna
- We support your efforts towards PhD



contact: m.zeppelzauer@fhstp.ac.at

Pre- & PostDoc Position: Machine Learning & Data Science (40h)



You will work on:

- Machine/deep learning, multimodal data analysis, computer vision, pattern recognition
- Interdisciplinary / human-centered topics
- Interactive ML on different types of data
- Write **publications** / hold presentations
- Acquire research **projects** (e.g. FWF / FFG)
- Possibility for teaching

Your profile:

- MSc / PhD in computer science (or comparable)
- Advanced skills in focus areas: ML/DL,
 computer vision / pattern recognition
- Python programming, DL frameworks
- Successful publication and project acquisition
- Excellent English skills
- Open-minded personality

contact: <u>m.zeppelzauer@fhstp.ac.at</u>

Pre- & PostDoc Position: Machine Learning & Data Science (40h)



We offer:

- Challenging **interdisciplinary** research projects
- Open-minded internationally-oriented environment
- **Long-term perspective:** unlimited contract (40h, part-time possible)
- **Career model** for personal development
- **Flexible** working hours, home office, family-friendly
- Comprehensive program for further education
- Opportunity to shape your **own research agenda**
- **Salary** at least: 37,100€ (PreDoc), 46,900€ (PostDoc)

We explicitly encourage all genders to apply.

We explicitly encourage people with disabilities to apply.

Apply here:

https://tinyurl.com/ICMT-PreDochttps://tinyurl.com/ICMT-PostDo

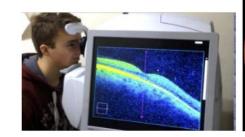
contact: m.zeppelzauer@fhstp.ac.at

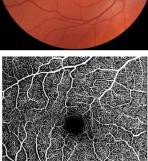


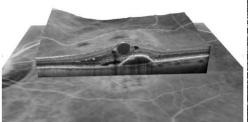


Retina Choroid Sclera Macula

- Retinal Image Analysis with Deep Learning
 - Modern medical imaging modalities: OCT, OCT-Angiography
- Interdisciplinary and International Research Lab
 - 60% PhD, 40% MD
- Availability of Large Curated Longitudinal Datasets
 - Multimodal imaging (0.5M OCT scans) and EHR
 - Well structured and annotated
- Industry Collaborations
 - · Big Pharma: Novartis, Genentech, Bayer, ...
 - · OCT device companies
- Academic Collaborations with JKU and ICL
 - FWF project with S. Hochreiter and G. Klambauer
 - Wellcome Trust project with D. Rueckert





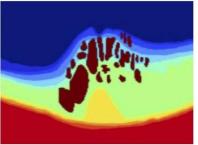




General research topics

Semantic segmentation





Disease progression modeling

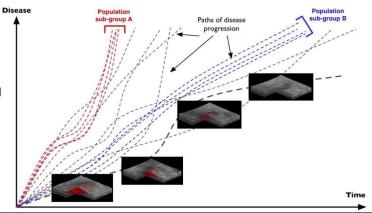
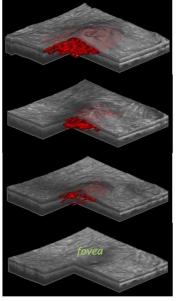
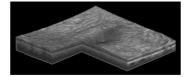


Image registration



0





Open positions (full-time)

- Software/Research Engineer (~ €38k)
 - Supporting clinical researchers and company partners
 - Maintaining our software and hardware infrastructure
 - Good programming skills (Python, C/C++, ...)
 - Some experience with deep learning (PyTorch, TensorFlow, ...)
- PhD student (~ €42k, MSc required)
 - · Industry applications: Collaboration with companies on ML for clinical decision support
 - Basic research: Part of the FWF-Research Group with JKU
 - Co-supervision of MSc students
- **PostDoc** (~ €56k, PhD required)
 - · Co-supervision of PhD students
 - Strong background (publications) in machine learning, medical image analysis or computer vision



Musimap is a "Music Emotion AI company"

We build Deep Learning based music search, tagging & recommendation products for music & advertising industry







Job offer

Data Scientist with a focus on (Music) Recommendation

- MSc or PhD in Computer Science or related
- experience in at least one of audio signal analysis, deep learning, similarity computation or recommender systems
- ideal: experience with big data analysis and/or highly scalable retrieval systems
 good knowledge of Python and Data Science frameworks
- good programming and documentation skills
- a goal driven and product-oriented mindset
- good team-worker plus ability to work also independently
- proficiency in English

Musimap is an equal opportunity workplace, female applicants are encouraged.

Full-time job ... 75% remote, 5 days (1 week) per month at Brussels HQ.

Starting date: January 2021 Deadline: Dec. 2nd -> 11th 2020 (last chance!)

http://www.musimap.com/jobs/data-scientist

Direct Contact: tom @ musimap.com

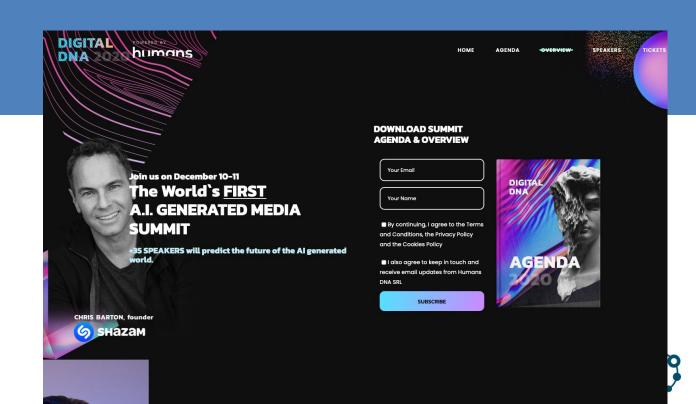




Conference Ticket Give-Away!

digitaldnasummit.joinhumans.com - Dec 10th + 11th

Send email now :-)
rene @ contextflow.com



Agenda for Today



Welcome & Introduction

Announcements (Jobs, Events)

Practical experiences in accurate video segmentation

by Antonis Makropoulos, Head of Machine Learning at contextflow

Latest News & Hot Topics in Deep Learning

by René Donner and Michael Pieler

Networking in Breakout-Rooms

Hot Topics & Latest News

a short block at every meetup to briefly present recent papers and news in Deep Learning

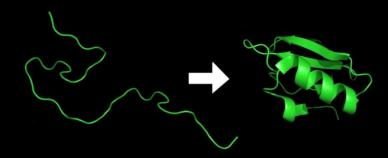
Send us contributions (tom.lidy@gmail.com) or propose to present some hot topics yourself!



DeepMind AlphaFold 2

Proteins and Protein Folding

- Amino acids are basic building blocks of life (21 of them)
- Proteins are chains of amino acids and are workhorses of living organisms (structure providers, movers, reaction catalysts, etc)



- Protein folding: amino acid sequence --> 3D structure
 - Uniqueness: The sequence usually maps 1-to-1 to a 3D structure
 - Function: 3D structure determines its function. Misfold --> disease.
 - Problem: 10^143 way to fold Levinthal's paradox
 - Disease: Misfold --> disease
 - Dataset: 200 million proteins. 170,000 protein 3D structures.
 - Cost: X-ray crystallography costs \$120,000 and takes 1 year [23]

DeepMind AlphaFold 2

AlphaFold1:

CNNs

AlphaFold2:

Attention mechanisms (transformers)

What happened?

- DeepMind's AlphaFold 2 solves protein folding (50 yeard old grand challenge)
 - "Solves" = Achieves 87+ GDT on CASP competition.



How big is this accomplishment? *

- Biggest advancement in structural biology of the past 20+ years.
- Biggest advancement in artificial intelligence of the past 20+ years.
 - ImageNet moment (aka AlexNet moment)
 - Competition: AlexNet (deep learning), AlphaZero (RL, self-play), GPT-3 (transformers), Tesla Autopilot, Waymo, smart speakers, Boston Dynamics, etc.
- Prediction: First Nobel Prize for machine learning model
- * This whole section is subjective and opiniated for **fun**. I can make an argument for any of the other breakthroughs listed (and some not listed)

NeurIPS 2020

Virtual (obviously)

100 USD

Until Dec 12th

Prerecorded videos, live chat during presentation

January 13th - VDLM Meetup:

NeurIPS Review



NeurIPS 2020 – Keynotes

Invited Talk:

You Can't Escape Hyperparameters and Latent Variables: Machine Learning as a Software Engineering Enterprise

Charles Isbell

Tue, Dec 8th, 2020 @ 02:00 - 04:00 CET

Invited Talk:
A Future of Work for the Invisible Workers in A.I.
Saiph Savage

Thu, Dec 10th, 2020 @ 02:00 - 04:00 CET



NeurIPS 2020 – Tutorials

Tutorial: (Track1) Federated Learning and Analytics: Industry Meets Academia

Brendan McMahan, Virginia Smith, Peter Kairouz

Tutorial and Q&A: Mon, Dec 7th @ 22:30 CET - Tue, Dec 8th @ 01:00 CET

Extra Q&A session: Thu, Dec 10th, 2020 @ 21:00 - 21:50 CET

Tutorial: (Track2) Explaining Machine Learning Predictions: State-of-the-art, Challenges, and Opportunities

Himabindu Lakkaraju, Julius Adebayo, Sameer Singh

Tutorial and Q&A: Mon, Dec 7th @ 22:30 CET - Tue, Dec 8th @ 01:00 CET

Extra Q&A session: Wed, Dec 9th, 2020 @ 12:00 - 12:50 CET

Tutorial: (Track2) Deeper Conversational AI

Pascale N Fung, Yun-Nung (Vivian) Chen, Zhaojiang Lin, Andrea Madotto

Tutorial and Q&A: Mon, Dec 7th, 2020 @ 09:00 - 11:30 CET

Extra Q&A session: Thu, Dec 10th, 2020 @ 09:00 - 09:50 CET



Exploring Simple Siamese Representation Learning

Xinlei Chen, Kaiming He (FAIR)

"SimSiam"



But before we look into SimSiam a small recap of:

SimCLR, BYOL, and SwAV



SimCLR (simple framework for contrastive learning)



BYOL (bootstrap your own latent)

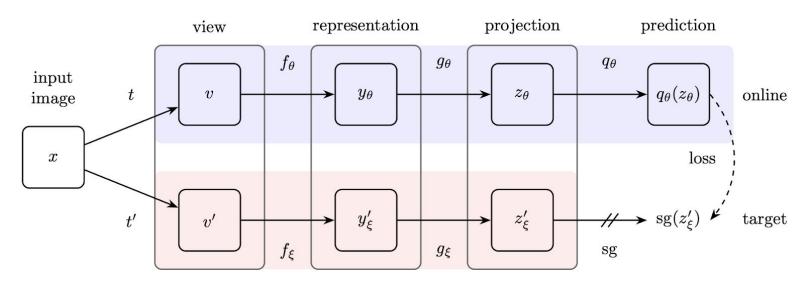


Figure 2: BYOL's architecture. BYOL minimizes a similarity loss between $q_{\theta}(z_{\theta})$ and $\operatorname{sg}(z'_{\xi})$, where θ are the trained weights, ξ are an exponential moving average of θ and sg means stop-gradient. At the end of training, everything but f_{θ} is discarded, and y_{θ} is used as the image representation.



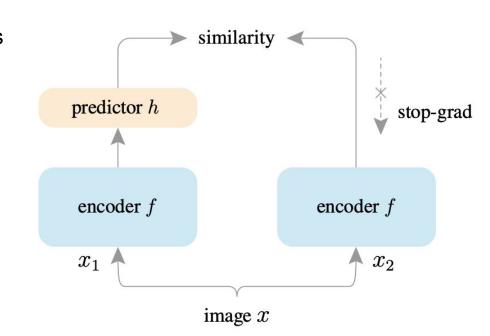
SwAV (swapping assignments between views)



SimSiam (simple siamese network)

Surprising empirical results show that simple Siamese networks can learn meaningful representations even when using none of the following:

- negative sample pairs
- large batches
- momentum encoders



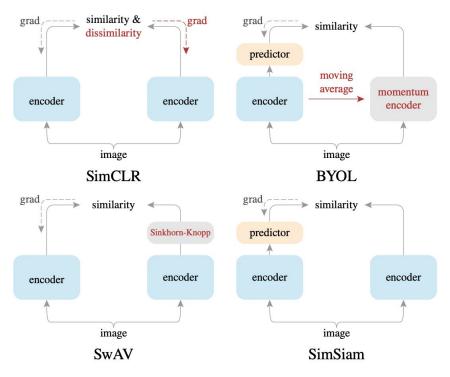


Algorithm 1 SimSiam Pseudocode, PyTorch-like

```
# f: backbone + projection mlp
# h: prediction mlp
for x in loader: # load a minibatch x with n samples
   x1, x2 = aug(x), aug(x) # random augmentation
   z1, z2 = f(x1), f(x2) # projections, n-by-d
  p1, p2 = h(z1), h(z2) # predictions, n-by-d
  L = D(p1, z2)/2 + D(p2, z1)/2 \# loss
  L.backward() # back-propagate
  update(f, h) # SGD update
def D(p, z): # negative cosine similarity
   z = z.detach() # stop gradient
  p = normalize(p, dim=1) # 12-normalize
   z = normalize(z, dim=1) # 12-normalize
  return -(p*z).sum(dim=1).mean()
```



SimSiam vs. others



The dash lines indicate the gradient propagation flow. In BYOL, SwAV, and SimSiam, the lack of a dash line implies stop-gradient, and their symmetrization is not illustrated for simplicity. The components in red are those missing in SimSiam. https://arxiv.org/abs/2011.10566

Deep Learning

Meetup

SimSiam results

method	batch size	negative pairs	momentum encoder	100 ep	200 ep	400 ep	800 ep
SimCLR (repro.+)	4096	✓		66.5	68.3	69.8	70.4
MoCo v2 (repro.+)	256	✓	✓	67.4	69.9	71.0	72.2
BYOL (repro.)	4096		✓	66.5	70.6	73.2	74.3
SwAV (repro.+)	4096			66.5	69.1	70.7	71.8
SimSiam	256			68.1	70.0	70.8	71.3

Table 4. Comparisons on ImageNet linear classification. All are based on ResNet-50 pre-trained with two 224×224 views. Evaluation is on a single crop. All competitors are from our reproduction, and "+" denotes *improved* reproduction vs. original papers (see supplement).

Ĩ	VOC 07 detection			VOC 07+12 detection			COCO detection			COCO instance seg.		
pre-train	AP_{50}	AP	AP_{75}	AP ₅₀	AP	AP ₇₅	AP ₅₀	AP	AP_{75}	AP ₅₀ ^{mask}	AP ^{mask}	AP ₇₅ mask
scratch	35.9	16.8	13.0	60.2	33.8	33.1	44.0	26.4	27.8	46.9	29.3	30.8
ImageNet supervised	74.4	42.4	42.7	81.3	53.5	58.8	58.2	38.2	41.2	54.7	33.3	35.2
SimCLR (repro.+)	75.9	46.8	50.1	81.8	55.5	61.4	57.7	37.9	40.9	54.6	33.3	35.3
MoCo v2 (repro.+)	77.1	48.5	52.5	82.3	57.0	63.3	58.8	39.2	42.5	55.5	34.3	36.6
BYOL (repro.)	77.1	47.0	49.9	81.4	55.3	61.1	57.8	37.9	40.9	54.3	33.2	35.0
SwAV (repro.+)	75.5	46.5	49.6	81.5	55.4	61.4	57.6	37.6	40.3	54.2	33.1	35.1
SimSiam, base	75.5	47.0	50.2	82.0	56.4	62.8	57.5	37.9	40.9	54.2	33.2	35.2
SimSiam, optimal	77.3	48.5	52.5	82.4	57.0	63.7	59.3	39.2	42.1	56.0	34.4	36.7

Table 5. **Transfer Learning**. All unsupervised methods are based on 200-epoch pre-training in ImageNet. *VOC 07 detection*: Faster R-CNN [32] fine-tuned in VOC 2007 trainval, evaluated in VOC 2007 test; *VOC 07+12 detection*: Faster R-CNN fine-tuned in VOC 2007 trainval + 2012 train, evaluated in VOC 2007 test; *COCO detection* and *COCO instance segmentation*: Mask R-CNN [18] (1× schedule) fine-tuned in COCO 2017 train, evaluated in COCO 2017 val. All Faster/Mask R-CNN models are with the C4-backbone [13]. All VOC results are the average over 5 trials. **Bold entries** are within 0.5 below the best.

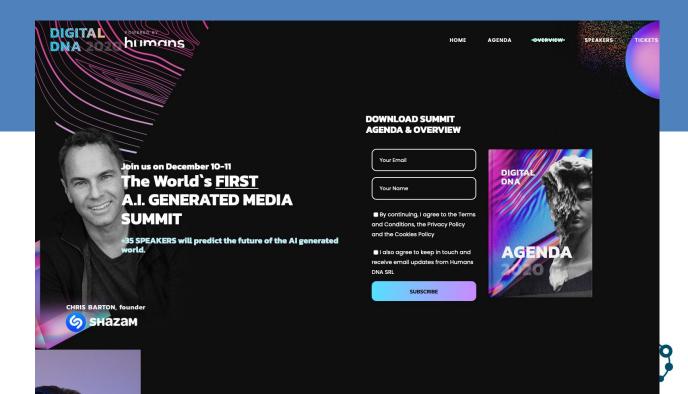


Conference Ticket Give-Away!

digitaldnasummit.joinhumans.com - Dec 10th + 11th

And the winner is ...:

I. H





Next Meetup: January 13th, 2021

NeurIPS Review

www.meetup.com/Vienna-Deep-Learning-Meetup

