Sergey Tulyakov

Snap Inc.

stulyakov@snap.com
sergey.tulyakov@unitn.it

Web
GitHub
Google Scholar

Summary

Sergey Tulyakov is a Principal Research Scientist heading the Creative Vision team at Snap Research. His work focuses on creating methods for manipulating the world via computer vision and machine learning. This includes human and object understanding in 2D and 3D, photorealistic manipulation and animation, video synthesis, prediction and retargeting. Sergey co-invented the unsupervised image animation domain with MonkeyNet and First Order Motion Model that sparkled a number of startups in the domain. His work on Interactive Video Stylization received the Best in Show Award at SIGGRAPH Real-Time Live! 2020. He has published 30+ top conference papers, journals and patents resulting in multiple innovative products, including Snapchat Pet Tracking, OurBaby Snappable, Real-time Neural Lenses (gender swap, baby face, aging lens, face animation) and many others. Before joining Snap Inc., he built the face tracking and reconstruction engine behind http://facify.me app.

Technical skills: C++, Python, C#, Qt, Boost, CMake, Docker, Kubernetes, GCloud
Scientific computing: PyTorch, python scientific stack, large scale machine learning, Eigen, dlib, OpenCV, PointCloudLibrary(PCL), matlab

Work experience

03/2021 – present	**Principal research scientist, Research Lead** at SNAP Inc. (Santa Monica, CA, US)
	I am leading the creative vision team focusing on creative technologies using machine learning and computer vision techniques.
11/2018 – 03/2021	**Lead research scientist, Research Lead** at SNAP Inc. (Santa Monica, CA, US)
07/2017 – 11/2018	**Senior research scientist** at SNAP Inc. (Santa Monica, CA, US)
	Designed and built a system for distributed, reproducible, and accelerated ML training, bringing 60x speed-up in ML training job creation and deployment. Currently is the default system for running training of deep models, serving 50+ active users across the company, with more than 10K training jobs run to date.
	Submitted 5 machine learning and computer vision patents.
	Contributed to multiple neural style lenses including OurBaby snappable, young, old face style transforms. Key contributor to the generative neural network distillation technology. Contributed to cat and dog tracking lenses.
01/2017 – 04/2017	**Research intern** at NVIDIA (Santa Clara, CA, US)
	Worked on motion-content decomposed video generation using Generative Adversarial Networks with Ming-Yu Liu and Jan Kautz.
08/2016 – 11/2016	**Research intern** at Microsoft Research (Cambridge, UK)
	Together with Sebastian Nowozin and Andrew Fitzgibbon worked on building efficient hybrids of deep generative and discriminative models, that benefit from using unlabeled data.
07/2010 – 09/2012	**Senior software engineer**, project lead at HiQo Solutions, Inc (Minsk, Belarus)
06/2006 – 06/2010	**Software engineer** at Todes, Ltd. (Minsk, Belarus)

Education

Oct 2012 – Apr 2017	University of Trento (Trento, Italy) **PhD in Computer Science.** Thesis: A Computer Vision Perspective on Face Analysis: Registration, Tracking, Synthesis
Nov 2014 – Feb 2015	Carnegie Mellon University (Pittsburgh, USA) **Research intern** at Robotics Institute
Sept 2009 – July 2010	Belarusian State University of Informatics and Radioelectronics (Minsk, Belarus) Master in Computer Science. **Final grade: 9 out of 10**
Sept 2004 – July 2009	Belarusian State University of Informatics and Radioelectronics (Minsk, Belarus) Bachelor in Computer Science. **Diploma with distinction. GPA 8.9. Final grade: 10 out of 10**

Selected personal projects

- **Facify** is a face tracking and 3D face reconstruction technology, enabling real-time face analysis on low-power mobile devices. The technology features a tiny hard drive footprint (20mb) and impressive tracking speed (200 frames per second on iPhone 6)
- **FaceCept - Face perCept** is a technology that allows real-time analysis of people's faces. Key features: gender, age, facial expression, new/returning, attention time recognition. The technology is cross-platform: it runs even in a browser.
- **FaceCept3D** is a flexible open-source technology for 3D face analysis and recognition, available on GitHub. Key features: head pose, facial expression and action units recognition in real time. FaceCept3D handles head pose ranges much wider than other systems.

Selected scientific projects

- The **ACANTO** project aims at increasing the number of older adults who engage in a regular and sustained physical activity. Developed a real-time method for instantaneous heart rate recognition from face videos.
- The **DALi - Devices for Assisted Living** project aimed at extending autonomous live of elderly people beyond home. Developed a real-time system for head pose analysis, tracking and facial expression recognition under a wide range of head poses.
- The purpose of the **PerTe - Persuasive Technology** project was to aid groups of people in a brainstorming environment. Worked on user monitoring part: track faces, measure attention given/received and analyze speech activity of the subjects.

Professional activities

Events:	• Organising Committee of the 10th Workshop on Human Behaviour Understanding organized in conjuction with ICCV 2019. • Chair of the first Workshop on 3D Face Alignment in the Wild (3DFAW) & Challenge organized in conjunction with ECCV 2016.
Teaching:	• Graduate course on Teaching Computers to Imagine with Deep Generative Models, University of Trento, Italy. Nov 19-26, 2019 • CVPR Tutorial on Unlocking Creativity with Computer Vision: Representations for Animation, Stylization and Manipulation, CVPR Virtual 2021

Reviewer:	I serve as a reviewer at all major CV, ML and Graphics conferences, including CVPR, ECCV, ICCV, IJCV, NeurIPS, SIGGRAPH, BMVC
	International Journal on Computer Vision
	IEEE Transactions on Affective Computing
	IEEE Transactions on Multimedia
	ACM Transactions on Intelligent Systems and Technology
	Elsevier Image and Vision Computing

Talks:

- **Representations for Content Creation, Manipulation and Animation**, BAIR Research group, UC Berkeley, Feb 17, 2021
- **Representations for Content Creation, Manipulation and Animation**, Oxford, Torr Vision Group, Jan 19, 2021
- **Learning representations for content creation, manipulation and animation**, Stanford University, Apr 2020
- **Learning representations for video generation, prediction and retargeting**, The Fourth Christmas Colloquium on Computer Vision, Skoltech, Moscow, Dec 2018
- **Vision-based human understanding**, University of Southern California, Los Angeles, CA, Nov 2018
- **Vision-based human understanding at Snap Research**, Keynote at ICME Workshop on Faces in Multimedia, San Diego, CA, July 2018
- **Computer Vision Perspective on Face Analysis: Registration, Tracking, Synthesis**, BAIR at UC Berkeley, CA, Mar 2018
- **Towards Learning to Imagine Videos with Controlled Content**, GTC 2018, San Jose, CA, Apr 2018
- **Real-time 3D face tracking using 2D and 3D sensors**, Stanford Neuroscience Institute, CA, Apr 2018
- **On Disentangling Motion and Content for Video Generation**, NIPS Learning Disentangled Representations: from Perception to Control, Los Angeles, CA, 2017
- **Head Pose Estimation and Facial Expression Recognition under a Wide Range of Head Poses**, Robotics institute, Carnegie Mellon University, PA, Nov 2014

Selected publications

[1] A. Siarohin, O. Woodford, J. Ren, M. Chai, S. Tulyakov *Motion Representations for Articulated Animation.* Computer Vision and Pattern Recognition, CVPR, 2021.

[2] J. Ren, M. Chai, O. Woodford, K. Olszewski, S. Tulyakov *Flow Guided Transformable Bottleneck Networks for Motion Retargeting.* Computer Vision and Pattern Recognition, CVPR, 2021.

[3] **[Oral]** W. Menapace, S. LathuiliÃĺre, S. Tulyakov, A. Siarohin, E. Ricci *Playable video generation.* Computer Vision and Pattern Recognition, CVPR, 2021.

[4] Q. Jin, J. Ren, O. Woodford, J. Wang, G. Yuan, Y. Wang, S. Tulyakov *Teachers Do More Than Teach: Compressing Image-to-Image Models.* Computer Vision and Pattern Recognition, CVPR, 2021.

[5] **[Spotlight]** Y. Tian, J. Ren, M. Chai, K. Olszewski, X. Peng, D. Metaxas, S. Tulyakov *A good image generator is what you need for high-resolution video synthesis.* International Conference on Learning Representations, ICLR, 2021.

[6] A. Siarohin, S. Lathuilière, S. Tulyakov, E. Ricci, and N. Sebe. *First Order Motion Model for Image Animation.* Neural Information Processing Systems, NeurIPS, 2019.

[7] **[Oral]** K. Olszewski, S. Tulyakov, O. Woodford, H. Li and L. Luo *Transformable Bottleneck Networks.* International Conference on Computer Vision, ICCV, 2019.

[9] J. Robinson, Y. Li, N, Zhang, Y. Fu and S. Tulyakov *Laplace landmark localization.* International Conference on Computer Vision, ICCV, 2019.

[10] **[Oral]** A. Siarohin, S. Lathuilière, S. Tulyakov, E. Ricci, and N. Sebe. *Animating Arbitrary Objects via Deep Motion Transfer.* Computer Vision and Pattern Recognition, CVPR, 2019.

[11] Z. Geng, C. Cao, S. Tulyakov. *3D Guided Fine-Grained Face Manipulation.* Computer Vision and Pattern Recognition, CVPR, 2019.

[12] S. Tulyakov, M.-Y. Liu, X. Yang, and J. Kautz. *MoCoGAN: Decomposing Motion and Content for Video Generation.* Computer Vision and Pattern Recognition, CVPR, 2018.

[13] W. Wang, S. Tulyakov, N. Sebe. *Recurrent Convolutional Shape Regressions.* Transactions on Pattern Analysis and Machine Intelligence, TPAMI, 2018.

[14] S. Tulyakov, A. Fitzgibbon, and S. Nowozin *Hybrid-VAE: Improving Deep Generative Models using Partial Observations.* Bayesian Deep Learning Workshop, NIPSW, 2017.

[15] László A. Jeni, Sergey Tulyakov, Lijun Yin, Nicu Sebe, and Jeffrey F. Cohn *The First 3D Face Alignment in the Wild (3DFAW) Challenge.* European Conference on Computer Vision, ECCV, 2016.

[16] **[Oral]** W. Wang, S. Tulyakov, N. Sebe. *Recurrent Convolutional Face Alignment.* Asian Conference on Computer Vision, ACCV, 2016.

[17] S. Tulyakov, L. A. Jeni, N. Sebe, and J. Cohn. *Viewpoint-consistent 3D Face Alignment.* Transactions on Pattern Analysis and Machine Intelligence, TPAMI.

[18] **[Oral]** S. Tulyakov, X. Alameda-Pineda, E. Ricci, L. Yin, N. Sebe, and J. Cohn. *Self-Adaptive Matrix Completion for Heart Rate Estimation from Face Videos under Realistic Conditions.* Computer Vision and Pattern Recognition, CVPR, 2016.

[19] S. Tulyakov and N. Sebe. *Regressing a 3D Face Shape from a Single Image.* In International Conference on Computer Vision, ICCV, 2015.

[20] S. Tulyakov, R. L. Vieriu, E. Sangineto and N. Sebe. *FaceCept3D: Real Time 3D Face Tracking and Analysis.* In International Conference on Computer Vision Workshops, ICCVW, 2015.

[21] R. L. Vieriu, S. Tulyakov, E. Sangineto, S. Semeniuta, and N. Sebe. *Facial Expression Recognition under a Wide Range of Head Poses.* In Face and Gesture Recognition, FG, 2015.

[22] S. Tulyakov, R. L. Vieriu, S. Semeniuta, and N. Sebe. *Robust Real-Time Extreme Head Pose Estimation.* In International Conference on Pattern Recognition, ICPR, 2014.

Patents

[1] C. Crutchfield, M. Gusarov, S. Korolev, S. Kotcur, A. Stoliar, S. Tulyakov and A. Wan *Generative Neural Network Distillation.* USPTO Patent Application.

[2] C. Cao, Z. Geng, and S. Tulyakov *Image Face Manipulation.* USPTO Patent Application.

[3] E. Buehl, J. Hurwitz, S. Vij and S. Tulyakov *Cloud Based Machine Learning.* USPTO Patent Application.

[4] S. Tulyakov, M-Y. Liu, X. Yang, and J. Kautz. *System and method for content and motion controlled action video generation.* Patent Application No: 62/354,475.

[5] S. Tulyakov, X. Alameda-Pineda, E. Ricci, L. Yin, N. Sebe, and J. Cohn. *Self-Adaptive Matrix Completion for Heart Rate Estimation from Face Videos under Realistic Conditions.* Patent Application No: 62/480,094.