

Sercan Ö. Arik

✉ sercanarik@alumni.stanford.edu • 🌐 www.sercanarik.com

Education

Stanford University

PhD, Electrical Engineering
CGPA: 3.99/4.0

Stanford, CA
September 2011 – September 2016

Stanford University

Master of Science, Electrical Engineering
CGPA: 4.02/4.0

Stanford, CA
September 2011 – April 2013

Bilkent University

Bachelor of Science, Electrical and Electronics Engineering
CGPA: 3.98/4.0

Ankara, Turkey
September 2007 – June 2011

Professional Experience

Baidu Silicon Valley Artificial Intelligence Research Lab

Research Scientist

Sunnyvale, CA
October 2016 – Current

- Researched:
 - Neural network architectures and deep learning techniques, including:
 - Discriminative architectures,
 - Generative models,
 - Sequence-to-sequence models,
 - Attention mechanisms,
 - Distributed representations,
 - Multi-modal learning,
 - Few-shot learning,
 - Transfer learning,
 - Unsupervised learning.
 - Signal processing techniques for information representation and retrieval,
 - Efficient implementation of machine learning algorithms.
- Developed keyword spotting systems for applications in conversational interfaces, demonstrated state-of-the-art detection performance using deep learning and worked on deployment in embedded environments.
- Developed neural artificial speech synthesis systems, achieving state-of-the-art synthesis performance with human-level naturalness and production-level generation capability.
- Developed speaker classification systems using discriminative deep learning for applications in identity detection and personalized speech generation.
- Developed neural voice cloning systems, achieving state-of-the-art capabilities in style transfer and synthesized audio quality towards applications in personalization of human-machine interfaces.
- Provided technical consultancy for venture capital investments and business strategy development in artificial intelligence related areas.
- Received 'Baidu Quarterly Star' award.

Binatix Labs

Research Intern

Palo Alto, CA
July 2015 – September 2015

- Researched:
 - Machine learning techniques (using neural networks, ensemble models, kernel machines etc.) for applications in time-series predictions,
 - Signal processing algorithms and optimization methods for decision making in stochastic environments.
- Developed information retrieval techniques for financial data sets.
- Implemented trading and portfolio optimization algorithms.
- Demonstrated notable performance results for global financial instruments.

Mitsubishi Electric Research Labs*Research Intern***Cambridge, MA***July 2013 - September 2013*

- Researched:
 - Advanced modulation and coding techniques for digital communication systems,
 - Digital signal processing algorithms for compensation of noise and distortion effects.
- Developed transmission schemes based on high-dimensional signal constellations.
- Demonstrated significant achievements towards information-theoretic capacity-approaching long-haul optical networks.

Google*Research Intern***Mountain View, CA***July 2012 - September 2012*

- Researched:
 - Transmission and switching technologies for data centers,
 - Architectures for fiber-to-the-home access networks.
- Developed statistical estimation methods for analysis of optical effects.
- Designed and simulated transmission and networking architectures that contributed to future strategies.

Swiss Federal Institute of Technology (EPFL)*Research Intern***Lausanne, Switzerland***June 2010 - September 2010*

- Researched:
 - Computer vision algorithms for feature detection, image matching, and 3-D reconstruction,
 - Machine learning techniques for semi-supervised classification.
- Developed similarity determination methods for uncalibrated multi-view images and demonstrated high performance for various image classification applications.

National Space Technologies Research Institute*Hardware Engineering Intern***Ankara, Turkey***August 2009 - September 2009*

- Researched:
 - Image compression techniques,
 - Efficient hardware implementation of decoding/encoding algorithms.
- Designed the embedded system for a wireless image communication system based on a media-on-chip microprocessor.

Teaching Experience

Stanford University**Stanford, CA***2012-2016*

- *Instructor*
(Responsibilities included teaching the class for the entire quarter, preparing the curriculum, lecture notes, assignments and exams, and holding office hours.)
 - 'Digital Signal Processing'
 - 'The Fourier Transform and Its Applications'
- *Teaching Assistant*
(Responsibilities included devising assignments and exams, holding office hours; overseeing lab experiments and final projects.)
 - 'Introduction to Photonics'
- *Grading Assistant*
(Responsibilities included grading assignments.)
 - 'Signal Processing and Linear Systems II'
 - 'Digital Communications'
 - 'Advanced Optical Fiber Communications'
 - 'Introduction to Optical Fiber Communications'

Stanford Pre-Collegiate Studies**Stanford, CA***2014*

- *Instructor*
(Responsibilities included teaching the class, preparing the curriculum and lecture notes)
 - 'Information Technologies: from Particles to Bits' - Covered fundamentals of modern information technologies at middle school level (on digital signal processing, computing architectures, communication technologies, information theory, algorithms and machine learning).

Publications

Academic Impact: (according to Google Scholar as of July 2018)

- Citations \geq 651
- h-index \geq 13

Journal Papers:

1. S. Ö. Arik, H. Jun and G. Damos, "Fast spectrogram inversion using multi-head convolutional neural networks", submitted to *IEEE Signal Processing Letters*, Aug. 2018.
2. K. Choutagunta, S. Ö. Arik, K. P. Ho, and J. M. Kahn, "Characterizing mode-dependent loss and gain in multimode components", *Journal of Lightwave Technology*, vol. 36, no. 18, pp. 3815-3823, Sep. 2018.
3. S. Ö. Arik, and J. M. Kahn, "Low complexity implementation of convex optimization based phase retrieval", *Journal of Lightwave Technology*, vol. 36, no. 12, pp. 2358 - 2365, Jun. 2018.
4. O. D. Domingues, D. A. A. Mello, R. Silva, S. Ö. Arik, and J. M. Kahn, "Achievable rates of space-division multiplexed submarine links subject to nonlinearities and power feed constraints", *Journal of Lightwave Technology*, vol. 35, no. 18, pp. 4004 - 4010, Jun. 2017.
5. S. Ö. Arik, B. Ibragimov and L. Xing, "Fully automated quantitative cephalometry using convolutional neural networks", *Journal of Medical Imaging*, vol. 4, no. 1, pp. 014501-014501, Jan. 2017.
6. S. Ö. Arik and J. M. Kahn, "Direct-detection mode-division multiplexing enabled by phase retrieval", *Optics Letters*, vol. 41, no. 18, pp. 4265-4268, May 2016.
7. S. Ö. Arik, K. P. Ho and J. M. Kahn, "Group delay management and multi-input multi-output signal processing in mode-division multiplexing systems", *Journal of Lightwave Technology*, vol. 34, no. 11, pp. 2867 - 2880, Jun. 2016 (Invited).
8. S. Ö. Arik and H. M. Özaktaş, "Optimal representation and processing of optical signals in quadratic-phase systems", *Optics Communications*, vol. 366, pp. 17 - 21, May 2016.
9. S. Ö. Arik, K. P. Ho and J. M. Kahn, "Delay spread reduction in mode-division multiplexing: mode coupling vs. delay compensation", *Journal of Lightwave Technology*, vol. 33, no. 21, pp. 4504 - 4512, Nov. 2015.
10. S. Ö. Arik and J. M. Kahn, "Spectral and spatial aggregation for high-throughput data transmission and networking", *SPIE Newsroom* 10.1117/2.1201501.005757, Feb. 2015 (Invited).
11. S. Ö. Arik, K. P. Ho and J. M. Kahn, "Optical network scaling: roles of spatial and spectral aggregation", *Optics Express*, vol. 22, no. 24, pp. 29868 - 29887, Dec. 2014 (Invited).
12. S. Ö. Arik and J. M. Kahn, "Diversity-multiplexing tradeoff in mode-division multiplexing", *Optics Letters*, vol. 39, no. 11, pp. 3258 - 3261, Jun. 2014.
13. S. Ö. Arik, D. Askarov and J. M. Kahn, "Adaptive frequency domain equalization in mode-division multiplexed systems", *Journal of Lightwave Technology*, vol. 32, no. 10, pp. 1841-1852, May. 2014.
14. D. Millar, T. K.-Akino, S. Ö. Arik, K. Kojima, K. Parsons, T. Yoshida and T. Sugihara, "High-dimensional modulation for coherent optical communications systems", *Optics Express*, vol. 22, no. 7, pp. 8798-8812, Apr. 2014.
15. S. Ö. Arik, J. M. Kahn and K. P. Ho, "MIMO signal processing in mode-division multiplexing", *IEEE Signal Processing Magazine*, vol. 31, no. 2, pp. 25-34, Mar. 2014 (Invited).
16. S. Ö. Arik and J. M. Kahn, "Coupled-core multi-core fibers for spatial multiplexing", *IEEE Photonics Technology Letters*, vol. 25, no. 21, pp. 2054-2057, Nov. 2013.
17. S. Ö. Arik, D. Askarov and J. M. Kahn, "Effect of mode coupling on signal processing complexity in mode-division multiplexing", *Journal of Lightwave Technology*, vol. 31, no. 13, pp. 423-431, Feb. 2013.
18. H. M. Özaktaş, S. Ö. Arik and T. Coşkun, "Fundamental structure of Fresnel diffraction: longitudinal uniformity with respect to fractional Fourier order", *Optics Letters*, vol. 37, no. 1, pp. 103-105, Jan. 2012.
19. H. M. Özaktaş, S. Ö. Arik and T. Coşkun, "Fundamental structure of Fresnel diffraction: natural sampling grid and the fractional Fourier transform", *Optics Letters*, vol. 36, no. 13, pp. 2524-2526, Jul. 2011.

Conference Papers/Presentations:

1. Y. Zhou, S. Ebrahimi, S. Ö. Arik, H. Yu, H. Liu, and G. Damos, "Resource efficient neural architect", submitted to *Conference on Neural Information Processing Systems*, 2018.
2. S. Ö. Arik, J. Chen, K. Peng, and W. Ping, and Y. Zhou, "Neural voice cloning with a few samples", submitted to *Conference on Neural Information Processing Systems*, 2018.
3. W. Ping, K. Peng, A. Gibiansky, S. Ö. Arik, A. Kannan, S. Narang, J. Raiman, and J. Miller, "Deep Voice 3: Scaling text-to-speech with convolutional sequence learning", *International Conference on Learning Representations*, Vancouver, Canada, Oct. 2017.
4. S. Ö. Arik, G. Damos, A. Gibiansky, J. Miller, K. Peng, W. Ping, J. Raiman, and Y. Zhou, "Deep Voice 2: Multi-speaker neural text-to-speech", *Conference on Neural Information Processing Systems*, Long Beach, CA, Dec. 2017 (Spotlight).
5. S. Ö. Arik, M. Chrzanowski, A. Coates, G. Damos, A. Gibiansky, Y. Kang, X. Li, J. Miller, A. Ng, J. Raiman, S. Sengupta, and M. Shoeybi, "Deep Voice: Real-time neural text-to-speech", *International Conference on Machine Learning*, Sydney, Australia, Aug. 2017.
6. S. Ö. Arik, M. Kliegl, R. Child, J. Hestness, A. Gibiansky, C. Fougner, R. Prenger, and A. Coates, "Convolutional

recurrent neural networks for small-footprint keyword spotting", *Interspeech*, Stockholm, Sweden, Aug. 2017.

7. J. M. Kahn and S. Ö. Arik, "Mode-division multiplexing using direct detection and adaptive optical signal processing", *OSA Topical Meeting on Photonic Networks and Devices*, New Orleans, LA, Jul. 2017 (*Invited*).
8. K. Choutagunta, S. Ö. Arik, M. Moradshahi and J. M. Kahn, "Optical MIMO signal processing for direct-detection mode-division multiplexing", *International Conference on Transparent Optical Networks*, Girona, Spain, Jul. 2017 (*Invited*).
9. A. C. J. Neto, C. E. Rothenberg, D. A. A. Mello, S. Ö. Arik, J. M. Kahn, "Scaling optical networks using full-spectrum spatial switching", *International Conference on High Performance Switching and Routing*, Campinas, Brazil, Jun. 2017.
10. O. Domingues, D. A. A. Mello, R. Silva, S. Ö. Arik, J. M. Kahn, "Capacity limits of space-division multiplexed submarine links subject to nonlinearities and power feed constraints", *Optical Fiber Communications Conference*, Los Angeles, CA, Mar. 2017.
11. S. Ö. Arik and J. M. Kahn, "Optical MIMO processing for direct-detection mode-division multiplexing", *Proceedings of European Conference on Optical Communications*, Düsseldorf, Germany, Sep. 2016 (*Invited*).
12. S. Ö. Arik and J. M. Kahn, "Direct-detection mode-division multiplexing enabled by phase retrieval", *IEEE Summer Topicals*, Newport Beach, CA, Jul. 2016.
13. S. Ö. Arik, K.-P. Ho and J. M. Kahn, "Group delay statistics and management in mode-division multiplexing", *Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, Nov. 2015 (*Invited*).
14. J. M. Kahn, S. Ö. Arik and K. P. Ho "MIMO channel statistics and signal processing in mode-division multiplexing systems", *IEEE International Workshop on Signal Processing Advances in Wireless Communications*, Stockholm, Sweden, Jun. 2015 (*Invited*).
15. S. Ö. Arik and J. M. Kahn, "MIMO DSP complexity in mode-division multiplexing systems", *Optical Fiber Communications Conference*, Los Angeles, CA, Mar. 2015 (*Invited*).
16. S. Ö. Arik and J. M. Kahn, "MIMO signal processing in mode-division multiplexing systems", *SPIE Photonics West Opto*, San Francisco, CA, Feb. 2015 (*Invited*).
17. S. Ö. Arik and J. M. Kahn, "Roles of spectral and spatial aggregation in optical network scaling", *SPIE Photonics West Opto*, San Francisco, CA, Feb. 2015 (*Invited*).
18. S. Ö. Arik and J. M. Kahn, "Adaptive MIMO signal processing in mode-division multiplexing", *IEEE Summer Topicals*, Montreal, Canada, Jul. 2014 (*Invited*).
19. S. Ö. Arik, D. Millar, T. K.-Akino, K. Kojima and K. Parsons, "High-dimensional modulation for mode-division multiplexing", *Optical Fiber Communications Conference*, San Francisco, CA, Mar. 2014.
20. D. Millar, T. K.-Akino, S. Ö. Arik, K. Kojima and K. Parsons, "Comparison of quaternary block-coding and sphere-cutting for high-dimensional modulation", *Optical Fiber Communications Conference*, San Francisco, CA, Mar. 2014.
21. S. Ö. Arik, E. Vural and P. Frossard, "Alignment of uncalibrated images for multi-view classification", *IEEE International Conference on Image Processing*, pp. 2413-2416, Brussels, Belgium, Sep. 2011.

Patents:

1. S. Ö. Arik, H. Jun and G. Damos, "Fast spectrogram inversion using multi-head convolutional neural networks" (filed in Aug. 2018).
2. S. Ö. Arik, H. Jun, E. Undersander, and G. Damos, "Spectrogram to waveform synthesis using convolutional generative adversarial networks" (filed in May 2018).
3. S. Ö. Arik, J. Chen, K. Peng, and W. Ping "Neural voice cloning with a few samples" (filed in Feb. 2018).
4. S. Ö. Arik, W. Ping, K. Peng, A. Gibiansky, A. Kannan, S. Narang, J. Raiman, and J. Miller, "Systems and methods for neural text-to-speech using convolutional sequence learning" (filed in Oct. 2017).
5. S. Ö. Arik, G. Damos, A. Gibiansky, J. Miller, K. Peng, W. Ping, J. Raiman, and Y. Zhou, "Systems and methods for multi-speaker neural text-to-speech" (filed in May 2017).
6. S. Ö. Arik, M. Kliegl, R. Child, J. Hestness, A. Gibiansky, C. Fougner, R. Prenger, and A. Coates, "Convolutional recurrent neural networks for small-footprint keyword spotting" (filed in March 2017).
7. S. Ö. Arik, M. Chrzanowski, A. Coates, G. Damos, A. Gibiansky, J. Miller, A. Ng, J. Raiman, S. Sengupta, M. Shoeybi, "Systems and methods for real-time neural text-to-speech" (filed in February 2017).

Computer Skills

Programming Languages: Python, Matlab, R, C/C++, Java

Special Expertise: Machine learning and big data frameworks

Achievements and Awards

- Received *Stanford Graduate Fellowship* - the most prestigious award offered by Stanford University for PhD studies (2011 - 2015)
- Received awards in trading algorithm development competitions by Quantiacs (2014-2015)

- Received *SPIE Scholarship in Optics and Photonics* - for achievements and potential contributions in the research fields of optics and photonics (2012)
- Ranked 2nd among the students of engineering faculty upon graduation from Bilkent University (2011)
- Ranked 13th in Turkey University Entrance Exam among 1.7M students and received numerous national fellowships (2007)

Major Media Appearances

- "Baidu's voice cloning AI can swap genders and remove accents", *The Next Web* (Feb. 2018)
- "Who wanted a future in which AI can copy your voice and say things you never uttered? Who?!", *The Register* (Feb. 2018)
- "Creepy technology can mimic your voice with just 60 seconds of audio", *IFL Science* (Feb. 2018)
- "Baidu can clone your voice after hearing just a minute of audio", *New Scientist* (Feb. 2018)
- "AI voice cloning from a few seconds of voice sampling is real and rapidly improving", *Next Big Future* (Feb. 2018)
- "China's Google equivalent can clone voices after seconds of listening", *Futurism* (Feb. 2018)
- "Baidu's new system can learn to imitate every accent", *The Verge* (Oct. 2017)
- "Baidu's new text-to-speech system can master hundreds of accents", *The Verge* (May 2017)
- "Baidu's text-to-speech system mimics a variety of accents perfectly", *Engadget* (May 2017)
- "Baidu's Deep Voice can quickly synthesize realistic human speech", *Engadget* (Mar. 2017)
- "A groundbreaking new AI taught itself to speak in just a few hours", *Futurism* (Mar. 2017)
- "Baidu's artificial intelligence lab unveils synthetic speech system", *MIT Technology Review* (Mar. 2017)

Other Professional Activities

- Ad-hoc journal reviewing:
 - IEEE Access
 - IEEE Transactions on Signal Processing
 - IEEE Journal of Biomedical and Health Informatics
 - IEEE Transactions on Biomedical Engineering
 - International Journal of Electrical Power and Energy Systems
 - IEEE Transactions on Information Forensics and Security
 - Optics Letters
 - Optics Express
 - Applied Optics
 - Journal of Optical Society of America A
 - Journal of Lightwave Technology
 - Journal of Optical Communications and Networking
 - IEEE Photonics Technology Letters
 - IEEE Transactions on Wireless Communications
 - IEEE Transactions on Communications
- Invited panelist at *Blockchain Based Artificial Intelligence, Cryptocurrency and Blockchain Venture Summit* (Mar. 2018)