Tomas Pfister

✉ tomas@pfister.fi
http://tomas.pfister.fi
in tomaspfister

Education

2011–2015 **Oxford University**, ***PhD Engineering Science***, VGG Research Group.
- Advisor: Prof. Andrew Zisserman.
- Topic: Deep learning for human pose & gesture.

2007–2010 **Cambridge University**, ***BA (Hons) Computer Science***, class 1 (highest grade).
- Dissertation: Automatic emotion detection from speech.

Work Experience

2018– **Head of research, Cloud AI**, ***Google***, Sunnyvale.
- Leader for Google Cloud AI Research, a team of researchers and software engineers working on tackling the most valuable research problems for Google Cloud customers.
- Set the mission, vision and strategy for the organization, revamping it from a basic science-focused organization to a successful research organization that is having significant product impact across Google Cloud (10+ impactful launches generating $xxM revenue and leading to $xxxM deals) as well as publishing 15 high quality papers in top ML conferences.
- Led a complex 100 Googler cross-PA effort involving 8 Google VPs and 3 SVPs to develop a highly accurate COVID-19 forecasting model that is used by the federal government, US states, healthcare organizations and Japanese government to determine public policy decisions, distribute personal protective equipment and plan hospital capacity.
- Launched 4 recommendations and time series forecasting models that improved Google Cloud AI Retail customer results by 40% and generating significant additional customer revenue.
- Launched manufacturing defect detection method that improves performance on key customer metrics by 70%.
- Launched new active learning method that improves results by 40%+ into production for document understanding and manufacturing defect detection production models.
- Launched novel end-to-end graph neural network document understanding model that is key for a $xxxM deal.
- Launched a novel attention-based tabular data learning system (featured in Google Cloud Next 2020) which is used by very large Google Cloud customers and has helped seal several deals.

2015–2018 **Staff Research Scientist & Cofounder**, ***Apple***, Cupertino.
- Published Apple's first deep learning paper that won the Best Paper Award at CVPR'17.
- Recipient of the Forbes 30 Under 30 in Science 2018 award.
- Initiated & invented the CVPR'17 project and took it from its inception through many hurdles to getting the award. Key technical and strategic contributor. Presenter at Apple CVPR'17 event. Wrote Apple's first Machine Learning Journal post.
- Now cofounder of Apple's new central ML research group. Leading recruiting efforts for research scientists & research engineers. Leading & contributing to several research projects on GANs, unsupervised learning, multimodal learning and general deep learning.
- Key contributor to Apple's AI strategy. Led the development of a series of machine learning organizational development proposals to SVPs and VPs that have now been implemented.
- Designed neural networks for Face ID facial landmark estimation & anti-spoofing in iPhone X.
- Deep learning research advisor for a variety of Apple teams.
- Advised Apple SVP on program challenges.
- Technical evaluator for two key startup acquisitions.
- Apple applications impacted: Face ID, autonomous systems, learning from simulators, object detection, keypoint estimation, multimodal learning, speech recognition, speech generation, cross-modal image retrieval, gaze estimation, hand tracking.

2015
Deep Learning Research Scientist, ***Lighthouse***, Palo Alto.
- Developed an accurate RGB-D video object detector & tracker neural network.
- Advised home monitoring startup on its AI strategy.

2012–2015
Deep Learning Researcher, ***VGG, University of Oxford***.
- Deep learning research in computer vision problems: human pose estimation, gesture recognition.
- Published in CVPR, ICCV, ECCV, IJCV; best paper honorable mention & best poster awards.

2013–2015
Machine Learning Lead, ***Wakelet***, Manchester, UK.
- Led machine learning efforts for a 15-person, $1.5M funded startup.
- Proposed new machine learning features, prototyped and tested them on users.
- Developed new algorithms for search, recommender systems and content quality prediction.
- Worked with five engineers to build and launch new features.

2009–2015
Chief Technology Officer (CTO), ***SJR Host***, Finland.
- CTO of a 4-person web hosting startup with a 60 server data centre and $500K annual revenue.
- Formulated business (branding, marketing) and technical (system architecture) strategies.
- 24/7 final point of contact for emergencies (full responsibility of operations & data centre).
- Led, supported and guided two full-time staff in person, by phone and e-mail.

2014–2015
Machine Learning Consultant, ***Cytora***, London, UK.
- Advisor to CTO on machine learning for automatic political risk prediction.
- Devised new machine learning & natural language processing algorithms.

2012–2015
HPC Consultant, ***University of Oxford***.
- Advised Oxford on High Performance Computing for CPU+GPU supercomputer.

2011
Engineering Intern, ***Google***, Mountain View.
- YouTube computer vision research for a new project using the Google's vision framework.
- Database engineering (Python load testing tool & 1,000+ lines open-sourced MySQL C code).

2010–2011
Computer Vision Researcher, ***University of Oulu***.
- Developed first system for automatically recognizing facial micro-expressions (for lie detection).
- Work published in ICCV & covered by *Wired* and other mainstream media.

2009–2010
Machine Learning Researcher, ***University of Cambridge***.
- Developed method for real-time prediction of emotions & public speaking skills from speech.

2009
Human Computer Interaction Researcher, ***University of Cambridge***.
- Developed an open-source robot controller that was demoed at HCI'09 & BBC News.

2008
Software Engineer Intern, ***AlertMe***, Cambridge, UK.
- Implemented database support in C backend & redesigned PHP customer control panel.

2004–2006
Java Project Maintainer and Developer, ***TWcore***.
- Maintained large open-source Java project for online game with >10,000 active users.

2001–2009
Chief architect, ***SJR Host***, Finland.
- Engineered large-scale web and systems for the company & clients.

Selected Publications & Patents

[1] S. Arik and T. Pfister, "Tabnet: Attentive interpretable tabular learning," in *AAAI*, 2021.

[2] K. Sohn, C.-L. Li, J. Yoon, M. Jin, and T. Pfister, "Learning and evaluating representations for deep one-class classification," in *ICLR*, 2021.

[3] Y. Zou, Z. Zhang, H. Zhang, C.-L. Li, X. Bian, J.-B. Huang, and T. Pfister, "Pseudoseg: Designing pseudo labels for semantic segmentation," in *ICLR*, 2021.

[4] S. O. Arik, C.-L. Li, J. Yoon, R. Sinha, A. Epshteyn, L. T. Le, V. Menon, S. Singh, L. Zhang, N. Yoder, M. Nikoltchev, Y. Sonthalia, H. Nakhost, E. Kanal, and T. Pfister, "Interpretable sequence learning for covid-19 forecasting," in *NeurIPS*, 2020. ***Spotlight***.

[5] J. Yoon, S. O. Arik, and T. Pfister, "Data valuation using reinforcement learning," in *ICML*, 2020.

[6] C. Xing, S. Arik, Z. Zhang, and T. Pfister, "Distance-based learning from errors for confidence calibration," in *ICLR*, 2020.

[7] L. Zhu, S. Arik, Y. Yang, and T. Pfister, "Learning to transfer learn," in *ECCV*, 2020.

[8] M. Gao, Z. Zhang, G. Yu, S. O. Arik, L. S. Davis, and T. Pfister, "Consistency-based semi-supervised active learning: Towards minimizing labeling budget," in *ECCV*, 2020.

[9] Z. Zhang, H. Zhang, S. Arik, H. Lee, and T. Pfister, "Distilling effective supervision from severe label noise," in *CVPR*, 2020.

[10] C.-K. Yeh, B. Kim, S. Arik, C.-L. Li, T. Pfister, and P. Ravikumar, "On concept-based explanations in deep neural networks," in *NeurIPS*, 2020.

[11] Y. Xie, H. Dai, M. Chen, B. Dai, T. Zhao, H. Zha, W. Wei, and T. Pfister, "Differentiable top-k operator with optimal transport," in *NeurIPS*, 2020.

[12] B. Lim, S. Arik, N. Loeff, and T. Pfister, "Temporal fusion transformers for interpretable multi-horizon time series forecasting," in *arXiv*, 2020.

[13] S. Arik and T. Pfister, "Protoattend: Attention-based prototypical learning," in *JMLR*, 2020.

[14] B. Lim, S. Arik, N. Loeff, and T. Pfister, "Time series interpretability using temporal fusion transformers," in *WHI*, 2020.

[15] K. Sohn, Z. Zhang, C.-L. Li, H. Zhang, C.-Y. Lee, and T. Pfister, "A simple semi-supervised learning framework for object detection," in *arXiv*, 2020.

[16] R. Zhang, T. Pfister, and L.-J. Li, "Harmonic unpaired image-to-image translation," in *ICLR*, 2019.

[17] L. Liu, M. Muelly, J. Deng, T. Pfister, and L.-J. Li, "Generative modeling for small-data object detection," in *ICCV*, 2019.

[18] D. Lee, T. Pfister, and M.-H. Yang, "Inserting videos into videos," in *CVPR*, 2019.

[19] S. Arik and T. Pfister, "Tabnet: Attentive interpretable tabular learning," in *arXiv*, 2019.

[20] J. Yoon, S. O. Arik, and T. Pfister, "Rl-lim: Reinforcement learning-based locally interpretable modeling," in *arXiv*, 2019.

[21] S. Arik and T. Pfister, "Protoattend: Attention-based prototypical learning," in *arXiv*, 2019.

[22] C.-K. Yeh, B. Kim, S. Arik, C.-L. Li, P. Ravikumar, and T. Pfister, "On concept-based explanations in deep neural networks," in *arXiv*, 2019.

[23] Y. Lu, J. Zhu, and T. Pfister, "A simple yet effective baseline for robust deep learning with noisy labels," in *arXiv*, 2019.

[24] A. Shrivastava, T. Pfister, O. Tuzel, J. Susskind, W. Wang, and R. Webb, "Learning from simulated and unsupervised images through adversarial training," in *CVPR*, 2017. ***Oral Presentation. Best Paper Award.***

[25] J. Charles, T. Pfister, D. Magee, D. Hogg, and A. Zisserman, "Personalizing video pose estimation," in *CVPR*, 2016. ***Oral presentation.***

[26] T. Pfister, J. Charles, and A. Zisserman, "Flowing ConvNets for human pose estimation in videos," in *ICCV*, 2015.

[27] T. Pfister, *Advancing Human Pose and Gesture Recognition.* PhD thesis, Oxford University, 2015.

[28] X. Li, X. Hong, A. Moilanen, X. Huang, T. Pfister, G. Zhao, and M. Pietikäinen, "Reading hidden emotions: Spontaneous micro-expression spotting and recognition," in *TAFFC*, 2015.

[29] T. Pfister, K. Simonyan, J. Charles, and A. Zisserman, "Deep convolutional neural networks for efficient pose estimation in gesture videos," in *ACCV*, 2014.

[30] T. Pfister, J. Charles, and A. Zisserman, "Domain-adaptive discriminative one-shot learning of gestures," in *ECCV*, 2014.

[31] J. Charles, T. Pfister, D. Magee, D. Hogg, and A. Zisserman, "Upper body pose estimation with temporal sequential forests," in *BMVC*, 2014. ***Best poster award***.

[32] T. Pfister, J. Charles, and A. Zisserman, "Large-scale learning of sign language by watching TV (using co-occurrences)," in *BMVC*, 2013.

[33] J. Charles, T. Pfister, M. Everingham, and A. Zisserman, "Automatic and efficient human pose estimation for sign language videos," *IJCV*, 2013.

[34] J. Charles, T. Pfister, D. Magee, D. Hogg, and A. Zisserman, "Domain adaptation for upper body pose tracking in signed tv broadcasts," in *BMVC*, 2013, ***Oral presentation***.

[35] X. Li, T. Pfister, X. Huang, G. Zhao, and M. Pietikäinen, "A spontaneous micro facial expression database: Inducement, collection and baseline," in *FG*, 2013.

[36] T. Pfister, J. Charles, M. Everingham, and A. Zisserman, "Automatic and efficient long term arm and hand tracking for continuous sign language TV broadcasts," in *BMVC*, 2012. ***Best video award. Best paper award honorable mention***.

[37] G. Zhao, X. Li, M. Pietikäinen, X. Huang, and T. Pfister, "Computer vision research for expression and micro-expression recognition.," *International Journal of Psychology*, 2012.

[38] T. Pfister et al, "Automated recognition algorithm for detecting facial expressions." Patent, 2012.

[39] T. Pfister, X. Li, G. Zhao, and M. Pietikäinen, "Recognising spontaneous facial micro-expressions," in *ICCV*, 2011. ***Most remembered poster***.

[40] T. Pfister, X. Li, G. Zhao, and M. Pietikäinen, "Differentiating spontaneous from posed facial expressions within a generic facial expression recognition framework," in *ICCV Workshops*, 2011.

[41] T. Pfister and P. Robinson, "Real-time recognition of affective states from non-verbal features of speech and its application for public speaking skill analysis," in *TAFFC*, 2011.

[42] T. Pfister and P. Robinson, "Speech emotion classification and public speaking skill assessment," in *ICPR Workshops*, 2010.

[43] T. Pfister, "Emotion detection from speech." Cambridge University Dissertation, 2010.

Selected Awards

- 2018 ○ Forbes 30 Under 30 in Science award.
- 2017 ○ Best Paper Award at Computer Vision & Pattern Recognition (CVPR) conference.
- 2011–2015 ○ Oxford University EPSRC Research Studentship & Centenary Year Graduate Scholarship.
- 2011–2015 ○ 20+ personal research awards worth over €180,000 for outstanding AI research (Osk Huttunen Foundation, Nokia Foundation, Ehrnrooth Foundation, Walter Ahlstrom Foundation, Tauno Tonning Foundation, Kaute Foundation, Foundation for Tech Advancement & HPY Foundation).
- 2015 ○ Oxford University Young Leader of Tomorrow at St Gallen Symposium (top 5 students).
- 2014 ○ Best Poster award at British Machine Vision Conference.
- 2012 ○ Oxford University TATA Idea Idol Startup Business Plan competition finalist.
- 2012 ○ Best Paper honorable mention & Best Video award at British Machine Vision Conference.
- 2011 ○ €250,000 in research funding from TEKES.
- 2007–2010 ○ Cambridge University European Trust scholar.
- 2010 ○ Cambridge University Outstanding Academic Achievement award.

2010 ○ Cambridge University Outstanding Dissertation award.
2010 ○ Cambridge University Computer Lab nominee for European Student of the Year award.
2007 ○ Academy of Finland Outstanding Dissertation award.
2007 ○ International Baccalaureate Top Ranking in Maths award.

Skills, Languages & Service

Machine Learning	12 years experience in computer vision & machine learning.	Systems	20 years of professional experience in UNIX systems & networks.
Programming	Mostly Python now, in the past C++.	Leadership	Leadership experience in 3 startups.
Languages	*Fluent:* English, Swedish, Finnish.	*Intermediate:*	German, Chinese.
Service	Reviewer for top machine learning publications (NeurIPS, ICML, CVPR, ECCV, ICCV).		