Daniel **Michelsanti**

Speech Processing and Computer Vision Researcher

michelsanti.daniel@gmail.com | danmic.github.io | danielmichelsanti | Daniel Michelsanti

Deep learning enthusiast. Dealing with speech processing and computer vision while enjoying hygge in Denmark. Currently investigating cutting-edge technology for next-generation hearing aids.

Experience

Demant Enterprise and Aalborg University (AAU)
Smørum and Aalborg, DENMARK

Industrial PostDoc — *Apr 2021 - Today*

- Project: Vision-assisted hearing aid systems.

Aalborg University (AAU)
Aalborg, DENMARK

Research Assistant — *Sep 2020 - Feb 2021*

- Project: Analysis of acoustic signals for industry 4.0 applications.

Universitat Pompeu Fabra (UPF)
Barcelona, SPAIN

Visiting Researcher — *Sep 2019 - Dec 2019*

- Collaboration with the image processing and the music technology groups.
- Research on deep learning applied to speech reconstruction from silent videos.

Aalborg University (AAU)
Aalborg, DENMARK

Teaching Assistant — *Sep 2018 - Dec 2018*

- Course: Machine learning.

Aalborg University (AAU)
Hjørring, DENMARK

Student Teacher — *Jan 2016 & Jan 2017*

- Organisation of the workshops and the activities for the students interested in learning the Italian language during the SprogCamp 2016 and the SprogCamp 2017 at Hjørring Gymnasium.

Bang and Olufsen A/S (B&O)
Struer, DENMARK

Machine Learning Intern — *Sep 2016 - Dec 2016*

- Design and implementation of a multimodal identification system based on machine learning techniques.

Aspasiel
Terni, ITALY

Algorithm Development Intern — *Oct 2012 - Dec 2012*

- Implementation of an algorithm for production scheduling to be used at Acciai Speciali Terni, the market leader of flat rolled stainless steel products in Italy.

Education

Aalborg University (AAU)
Aalborg, DENMARK

PhD Fellow — *Sep 2017 - Aug 2020*

- Project: Audio-visual speech enhancement based on deep learning.
- Supervisors: Prof. Zheng-Hua Tan and Prof. Jesper Jensen.

Aalborg University (AAU)
Aalborg, DENMARK

MSc in Vision, Graphics and Interactive Systems - GPA: 11.5/12 — *Sep 2015 - Jun 2017*

- Thesis Title: Generative adversarial networks for speech processing.
- Supervisor: Prof. Zheng-Hua Tan.

Università degli Studi di Perugia
Perugia, ITALY

BSc in Computer and Electronic Engineering - Final Grade: 110/110 cum laude — *Sep 2009 - Feb 2014*

- Thesis Title: Sorting algorithm implementation to optimise the input sequence of the annealing and pickling line at Acciai Speciali Terni.
- Supervisor: Prof. Emilio Di Giacomo.

Liceo Classico Jacopone da Todi
Todi, ITALY

Liceo Classico (High School Equivalent) - Final Grade: 100/100 — *Sep 2004 - Jul 2009*

- Humanistic studies.

Publications

Morrone, G., **Michelsanti, D.**, Tan, Z.-H. and Jensen, J. "Audio-visual speech inpainting with deep learning". Proceedings of ICASSP (Accepted). 2021.

Michelsanti, D., Tan, Z.-H., Zhang, S.-X., Xu, Y., Yu, M., Yu, D. and Jensen, J. "An overview of deep-learning-based audio-visual speech enhancement and separation". IEEE/ACM Transactions on Audio, Speech, and Language Processing. 2021.

Michelsanti, D., Slizovskaia, O., Haro, G., Gómez, E., Tan, Z.-H. and Jensen, J. "Vocoder-based speech synthesis from silent videos". Proceedings of INTERSPEECH. 2020.

Michelsanti, D., Tan, Z.-H., Sigurdsson, S. and Jensen, J. "Deep-learning-based audio-visual speech enhancement in presence of Lombard effect". Speech Communication, 115, pp.38–50. 2019.

Michelsanti, D., Tan, Z.-H., Sigurdsson, S. and Jensen, J. "Effects of Lombard reflex on the performance of deep-learning-based audio-visual speech enhancement systems". Proceedings of ICASSP, pp.6615–6619. 2019.

Michelsanti, D., Tan, Z.-H., Sigurdsson, S. and Jensen, J. "On training targets and objective functions for deep-learning-based audio-visual speech enhancement". Proceedings of ICASSP, pp.8077–8081. 2019.

Michelsanti, D., Tan, Z.-H., Sigurdsson, S. and Jensen, J. "Conditional generative adversarial networks for speech enhancement and noise-robust speaker verification". Proceedings of INTERSPEECH, pp.2008–2012. 2017.

Michelsanti, D., Ene, A.-D., Guichi, Y., Stef, R., Nasrollahi, K. and Moeslund, T. B. "Fast fingerprint classification with deep neural networks". Proceedings of VISIGRAPP, pp.202–209. 2017.

Arbués-Sangüesa, A., Ene, A.-D., Jørgensen, N. K., Larsen, C. A., **Michelsanti, D.** and Kraus, M. "Pyramid algorithm framework for real-time image effects in game engines". Interactivity, Game Creation, Design, Learning, and Innovation, pp.289–296. Springer, Cham. 2016.

Conferences, Workshops and Summer Schools

2020	**INTERSPEECH,** Oral presentation (virtual).	*Shanghai, CHINA*
2020	**MLSP,** Participation (virtual).	*Espoo, FINLAND*
2020	**ICASSP,** Participation (virtual).	*Barcelona, SPAIN*
2019	**DLBCN,** Poster presentation.	*Barcelona, SPAIN*
2019	**ICASSP,** Poster and oral presentation.	*Brighton, U.K.*
2019	**AMLD,** Poster presentation.	*Lausanne, SWITZERLAND*
2018	**MLSP,** Participation.	*Aalborg, DENMARK*
2018	**Summer school on advanced topics in machine learning,** Participation.	*Copenhagen, DENMARK*
2018	**ICVSS,** 150 selected out of 557. Poster presentation. Winner of the reading group competition.	*Punta Sampieri, ITALY*
2017	**Winter school on signal processing for hearing assistive devices,** Participation.	*Aalborg, DENMARK*
2017	**INTERSPEECH,** Poster presentation.	*Stockholm, SWEDEN*
2017	**VISIGRAPP,** Oral presentation.	*Porto, PORTUGAL*
2016	**SPLINE,** Participation.	*Aalborg, DENMARK*
2016	**The Innovation Camp powered by Bang and Olufsen,** Participation.	*Struer, DENMARK*
2016	**ArtsIT,** Oral presentation.	*Esbjerg, DENMARK*

Talks and Lectures

2021	**Speech reconstruction from silent videos using a vocoder,** Deeptails Seminar (INRIA).	*Remote*
2020	**Audio-visual speech enhancement based on deep fusion,** Aalborg University.	*Aalborg, DENMARK*
2019	**Audio-visual speech enhancement for hearing assistive devices,** Universitat Pompeu Fabra.	*Barcelona, SPAIN*
2018	**Audio-visual speech enhancement for hearing assistive devices,** Oticon A/S.	*Copenhagen, DENMARK*
2018	**Audio-visual speech enhancement based on deep fusion,** Aalborg University.	*Aalborg, DENMARK*

Academic Service

Reviewer for: Computer Modeling in Engineering & Sciences, Tech Science Press; Computer Speech & Language, Elsevier; IEEE Access; IEEE/ACM Transactions on Audio, Speech, and Language Processing; IEEE Signal Processing Letters; Neural Networks, Elsevier.

Additional Information

Languages	Italian (Native Speaker), English (Full Professional Proficiency), Danish (Elementary Proficiency).
Programming	Python (PyTorch, Tensorflow), MATLAB.
Writing and Presentation	LaTeX, Word, PowerPoint.
Professional Membership	IEEE - Denmark Section (Member). ISCA Speech (Member). Auditory-Visual Speech Association (Member).
Photo and Video Editing	Affinity Photo, OmniGraffle, iMovie.
Hobbies	Photography, reading, cooking, board games, movies.