



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 assistive devices with the goal of improving the quality of life in people with a hearing loss.

Experience

Industrial PostDoc

AALBORG UNIVERSITY (AAU) AND OTICON

- Project: Vision-assisted hearing aid systems.

Aalborg and Smørum, DENMARK

Apr 2021 - Today

Research Assistant

AALBORG UNIVERSITY (AAU)

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

Aalborg, DENMARK

Sep 2020 - Feb 2021

Visiting Researcher

UNIVERSITAT POMPEU FABRA (UPF)

- Collaboration with the image processing and the music technology groups to conduct research on speech reconstruction from silent videos.

Barcelona, SPAIN

Sep 2019 - Dec 2019

Teaching Assistant

AALBORG UNIVERSITY (AAU)

- Course: Machine learning.

Aalborg, DENMARK

Sep 2018 - Dec 2018

Student Teacher

AALBORG UNIVERSITY (AAU)

- Organisation of the activities for high school students interested in learning the Italian language during the SprogCamp 2016 and 2017.

Hjørring, DENMARK

Jan 2016 & Jan 2017

Machine Learning Intern

BANG AND OLUFSEN A/S (B&O)

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

Struer, DENMARK

Sep 2016 - Dec 2016

Algorithm Development Intern

ASPASIEL

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

Terni, ITALY

Oct 2012 - Dec 2012

Education

PhD Fellow

AALBORG UNIVERSITY (AAU)

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

Aalborg, DENMARK

Sep 2017 - Aug 2020

MSc in Vision, Graphics and Interactive Systems

AALBORG UNIVERSITY (AAU)

- GPA: 11.5/12.
- Thesis Title: Generative adversarial networks for speech processing.
- Supervisor: Prof. Zheng-Hua Tan.

Aalborg, DENMARK

Sep 2015 - Jun 2017

BSc in Computer and Electronic Engineering

UNIVERSITÀ DEGLI STUDI DI PERUGIA (UNIPG)

- Final Grade: 110/110 cum laude.
- 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.

Perugia, ITALY

Sep 2009 - Feb 2014

Liceo Classico (High School Equivalent)

LICEO CLASSICO JACOPONE DA TODI

- Humanistic studies. - Final Grade: 100/100.

Todi, ITALY

Sep 2004 - Jul 2009

Publications

Common metrics on the impact of my research can be found in [Google Scholar](#) and [Scopus](#).

Kalør, A. E., **Michelsanti, D.**, Chiariotti, F., Tan, Z.-H. and Popovski, P. "Remote anomaly detection in industry 4.0 using resource-constrained devices". Proceedings of SPAWC. 2021.

Morrone, G., **Michelsanti, D.**, Tan, Z.-H. and Jensen, J. "Audio-visual speech inpainting with deep learning". Proceedings of ICASSP. 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, 29, pp.1368–1396. 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.

Selected Conferences, Workshops and Summer Schools

2021	INTERSPEECH , Participation.	<i>Brno, CZECH REPUBLIC</i>
2021	ICASSP , Poster presentation (virtual).	<i>Toronto, CANADA</i>
2020	INTERSPEECH , Oral presentation (virtual).	<i>Shanghai, CHINA</i>
2019	DLBCN , Poster presentation.	<i>Barcelona, SPAIN</i>
2019	ICASSP , Poster and oral presentation.	<i>Brighton, U.K.</i>
2019	AML , Poster presentation.	<i>Lausanne, SWITZERLAND</i>
2018	Summer school on advanced topics in machine learning , Participation.	<i>Copenhagen, DENMARK</i>
2018	ICVSS , 150 selected out of 557. Poster presentation. Winner of the reading group competition.	<i>Punta Sempieri, ITALY</i>
2017	Winter school on signal processing for hearing assistive devices , Participation.	<i>Aalborg, DENMARK</i>
2017	INTERSPEECH , Poster presentation.	<i>Stockholm, SWEDEN</i>
2017	VISIGRAPP , Oral presentation.	<i>Porto, PORTUGAL</i>
2016	The Innovation Camp powered by Bang and Olufsen , Participation.	<i>Struer, DENMARK</i>
2016	ArtsIT , Oral presentation.	<i>Esbjerg, DENMARK</i>

Talks and Lectures

2021	Audio-visual speech enhancement and separation based on deep learning , ODSC Training.	<i>Remote</i>
2021	Audio-visual speech enhancement and separation based on deep learning , ICASSP Tutorial.	<i>Remote</i>
2021	Speech reconstruction from silent videos using a vocoder , Deeptails Seminar (INRIA).	<i>Remote</i>
2019	Audio-visual speech enhancement for hearing assistive devices , Universitat Pompeu Fabra.	<i>Barcelona, SPAIN</i>
2018	Audio-visual speech enhancement based on deep fusion , Aalborg University.	<i>Aalborg, DENMARK</i>

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.

Honors and Awards

Christian Benoit Award 2021, International Speech Communication Association and Association Francophone de la Communication Parlee. Award given biannually to a promising young scientist in the domain of speech communication.

ICVSS Reading Group Competition 2018, ICVSS Organizing Committee. Winner of the competition.

Merit-Based Scholarship 2010, Fondazione Franco Todini. Best high school student of the year 2009 at Liceo Classico Jacopone da Todi.