Juan Pablo Zuluaga-Gomez

Speech & Audio Processing Research Group
Idiap Research Institute
Rue Marconi 19, 1920 Martigny
(+41) 027 721 77 11
Website: juanpzuluaga.github.io
Emails: juan.zuluaga@eu4m.eu juan-pablo.zuluaga@idiap.ch

EDUCATION

PhD Candidate, Electrical Engineering & Computer Science
École polytechnique fédérale de Lausanne, Vaud, Switzerland *January 2020 - January 2024*
THESIS - Speech and text-based technologies applied to air traffic control communications
Supervisor: Petr Motlicek, PhD.

Master of Science, Mechatronic Engineering
Universidad de Oviedo, Spain & ENSMM, France *September 2017 - September 2019*
THESIS - Breast cancer diagnosis based on computer vision
Score 89/100
Supervisor: Noureddine Zerhouni, PhD.

Bachelor of Science, Mechatronic Engineering
Universidad Autonoma del Caribe, Barranquilla, Colombia *January 2011 - December 2015*
THESIS - Precordial signal detection system by Seismocardiography
Score 91/100
Supervisor: Pablo Bonaveri, PhD.

PROFESSIONAL EXPERIENCE

Apple, AI/ML Team, Cambridge, MA July – September 2023
Machine Learning Engineer—Internship

- Working on discriminative training of language models to improve automatic speech recognition (ASR) performance on tail named-entity data
- Transformer-based language modeling for production-level ASR systems

Amazon, Amazon Web Services (AWS), Seattle, WA April – July 2023
Applied Scientist—Internship

- Member of the AWS AI Transcribe & Translate teams
- Research on dual speech-to-text Translation (ST) and Transcription (ASR) for conversational speech
- Serialized output training (conditioned with special tokens, akin to Whisper) for robust multilingual ST and ASR
- Our system is aware of speaker turns and overlapped speech, improving BLEU and WER performance

Idiap Research Institute, Valais, Switzerland January 2020 -
Doctor of Philosophy (Ph.D.) - Candidate

- Automatic speech recognition (ASR) for air traffic control (ATC): ATCO2 EU-H2020
- Implemented innovative semi-supervised techniques for ASR in air-traffic control (low resource task)
- Led the integration of natural language processing (NLP) techniques. 50% improvement in named-entity recognition from ASR transcripts (breakthrough)
- Developed systems for speaker role and speaker change detection based on ASR transcripts
- Participated at several venues: INTERSPEECH, ICASSP, OpenSky Network Symposium (7 conf.)

- Implemented a streaming ASR system for ATC communications: collaboration with industrial partners
- Participation on industrial projects: spoken language understanding (use case: call-centers)

Research Institute Femto-ST, Besancon, France February 2019 - October 2019
Master of Science Thesis

- Participated: SBRA-"Smart BRA" project, financed by INTERREG (France-Suisse)
- Developed a system for breast cancer diagnosis based on thermal images
- Early research in multi-modal techniques (vision & signal) for breast cancer diagnosis
- Published two journal papers
- Master Thesis: Breast Cancer Diagnosis Using Machine Learning

Universidad Autonoma del Caribe, Barranquilla, Colombia
Mechatronic Research Group Member UAC September 2014 -

- Participation in national and international events (7), co-authorship in publications (4), 2 patents
- Active member of the GIIM research group of mechatronic, as a senior research student for three consecutive years and then an active member

Bachelor Student January 2011 – December 2015

- Main topics covered: automatic control, electronics, mechanical systems, robotics, nanotechnology, machine learning and computer science
- Research on Titanium dioxide (TiO_2) for wastewater decontaminaton: work as student on the GIIM research group in Mechatronics
- Thesis: developed a system to capture and analyze precordial signals by Seismocardiography

PUBLICATIONS (JOURNAL, PEER REVIEWED)

6. Zuluaga-Gomez, J., Nigmatulina, I., Prasad, A., Motlicek, P., Khalil, D., Madikeri, S., Tart, A., Szoke, I., Lenders, V., Rigault, M., et al. (2023). Lessons learned in atco2: 5000 hours of air traffic control communications for robust automatic speech recognition and understanding. *arXiv preprint arXiv:2305.01155 (under review at Aerospace Journal).*
5. Ahrenhold, N., Helmke, H., Mühlhausen, T., Ohneiser, O., Kleinert, M., Ehr, H., Klamert, L., & **Zuluaga-Gomez, Juan**. (2023). Validating Automatic Speech Recognition and Understanding for Pre-Filling Radar Labels – Increasing Safety While Reducing Air Traffic Controllers' Workload. *Aerospace*, *10*(6). https://doi.org/10.3390/aerospace10060538.
4. **Zuluaga-Gomez, Juan**, Prasad, A., Nigmatulina, I., Motlicek, P., & Kleinert, M. (2023). A virtual simulation-pilot agent for training of air traffic controllers. *Aerospace*, *10*(5). https://doi.org/10.3390/aerospace10050490.
3. **Zuluaga-Gomez, Juan**, Veselý, K., Szöke, I., Motlicek, P., et al. (2022). ATCO2 corpus: A Large-Scale Dataset for Research on Automatic Speech Recognition and Natural Language Understanding of Air Traffic Control Communications. *Under review at Computer Speech & Language, arXiv preprint arXiv:2211.04054.*
2. Zhan, Q., Xie, X., Hu, C., **Zuluaga-Gomez, Juan**, et al. (2021). Domain-Adversarial Based Model with Phonological Knowledge for Cross-Lingual Speech Recognition. *Electronics*, *10*(24), 3172.
1. **Zuluaga-Gomez, Juan** et al. (2021a). A CNN-based methodology for breast cancer diagnosis using thermal images. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, *9*(2).

PUBLICATIONS (JOURNAL, REVIEW PAPER, PEER REVIEWED)

2. **Zuluaga-Gomez, Juan**, Bonaveri, P., Zuluaga, D., et al. (2020). Techniques for water disinfection, decontamination, and desalinization: A review. *Desalination And Water Treatment.*

1. **Zuluaga-Gomez, Juan** et al. (2019). A survey of breast cancer screening techniques: Thermography and electrical impedance tomography. *Journal of medical engineering & technology*, *43*(5).

PUBLICATIONS (CONFERENCE, PEER REVIEWED)

21. **Zuluaga-Gomez, Juan**, Ahmed, S., Visockas, D., & Subakan, C. (2023). CommonAccent: Exploring Large Acoustic Pretrained Models for Accent Classification Based on Common Voice. *Proc. Interspeech 2023 **[Nominated as best student paper award]**.*

20. *Mai, F., ***Zuluaga-Gomez, Juan**, Parcollet, T., & Motlicek, P. (2023). HyperConformer: Multi-head HyperMixer for Efficient Speech Recognition. *Proc. Interspeech 2023.*

19. Nigmatulina, I., Madikeri, S., Villatoro-Tello, E., Motliček, P., **Zuluaga-Gomez, Juan**, Pandia, K., & Ganapathiraju, A. (2022). Implementing contextual biasing in GPU decoder for online ASR. *Proc. Interspeech 2023.*

18. Villatoro-Tello, E., Madikeri, S., **Zuluaga-Gomez, Juan**, Sharma, B., Sarfjoo, S. S., Nigmatulina, I., Motlicek, P., Ivanov, A. V., & Ganapathiraju, A. (2023). Effectiveness of text, acoustic, and lattice-based representations in spoken language understanding tasks. *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 1–5.

17. Helmke, H., Kleinert, M., Ahrenhold, N., Ehr, H., Mühlhausen, T., Ohneiser, O., Klamert, L., Motlicek, P., Prasad, A., **Zuluaga Gomez, Juan**, et al. (2023). Automatic Speech Recognition and Understanding for Radar Label Maintenance Support Increases Safety and Reduces Air Traffic Controllers' Workload. *Proceedings of the 15th USA/Europe Air Traffic Management Research and Development Seminar (ATM2023), Savannah, GA, USA*, *17*.

16. **Zuluaga-Gomez, Juan**, Prasad, A., Nigmatulina, I., et al. (2022). How Does Pre-trained Wav2Vec 2.0 Perform on Domain Shifted ASR? An Extensive Benchmark on Air Traffic Control Communications. *2022 IEEE Spoken Language Technology Workshop (SLT), Doha, Qatar.*

15. **Zuluaga-Gomez, Juan**, Sarfjoo, S. S. et al. (2022). BERTraffic: BERT-based Joint Speaker Role and Speaker Change Detection for Air Traffic Control Communications. *2022 IEEE Spoken Language Technology Workshop (SLT), Doha, Qatar.*

14. *Prasad, A., ***Zuluaga-Gomez, Juan** et al. (2022). Speech and Natural Language Processing Technologies for Pseudo-Pilot Simulator. *12th SESAR Innovation Days.*

13. Helmke, H., Ondřej, K., Shetty, S., Arilíusson, H., Simiganoschi, T., Kleinert, M., Ohneiser, O., Ehr, H., **Zuluaga-Gomez, Juan**, & Smrz, P. (2022). Readback Error Detection by Automatic Speech Recognition and Understanding – Results of HAAWAII Project for Isavia's Enroute Airspace. *12th SESAR Innovation Days.*

12. Prasad, A., **Zuluaga-Gomez, Juan**, Motlicek, P., et al. (2022). Grammar Based Speaker Role Identification for Air Traffic Control Speech Recognition. *12th SESAR Innovation Days.*

11. Nigmatulina, I., **Zuluaga-Gomez, Juan**, Prasad, A., et al. (2022). A two-step approach to leverage contextual data: speech recognition in air-traffic communications. *ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 6282–6286.

10. **Zuluaga-Gomez, Juan** et al. (2021b). Contextual Semi-Supervised Learning: An Approach to Leverage Air-Surveillance and Untranscribed ATC Data in ASR Systems. *Proc. Interspeech 2021.*

9. Kocour, M., Veselý, K., Szoke, I., Kesiraju, S., **Zuluaga-Gomez, Juan**, Blatt, A., et al. (2021). Automatic processing pipeline for collecting and annotating air-traffic voice communication data. *Engineering Proceedings*, *13*(1), 8.

8. Kocour, M., Veselý, K., Blatt, A., **Juan Zuluaga-Gomez**, et al. (2021). Boosting of Contextual Information in ASR for Air-Traffic Call-Sign Recognition. *Proc. Interspeech 2021*.
7. **Zuluaga-Gomez, Juan**, Veselý, K. et al. (2020). Automatic call sign detection: Matching air surveillance data with air traffic spoken communications. *Multidisciplinary Digital Publishing Institute Proceedings*, *59*(1), 14.
6. **Zuluaga-Gomez, Juan**, Motlicek, P. et al. (2020). Automatic Speech Recognition Benchmark for Air-Traffic Communications. *Interspeech*, 2297–2301. https://doi.org/10.21437/Interspeech.2020-2173.
5. Ma, J., Shang, P., Lu, C., Meraghni, S., Benaggoune, K., **Zuluaga-Gomez, Juan**, Zerhouni, N., Devalland, C., & Al Masry, Z. (2019). A portable breast cancer detection system based on smartphone with infrared camera. *Vibroengineering Procedia*, *26*, 57–63.
4. Bonaveri, P., Barrios, M., & **Zuluaga-Gomez, Juan**. (2017). Diseño y construcción de un sistema basado en acelerometría para la captación y análisis en matlab de señales precordiales usando sismocardiografía 3d. *Memorias del Congreso Nacional de Ingeniería Biomédica*, *2*(1), 153–156.
3. **Zuluaga-Gomez, Juan** et al. (2018). Aprendizaje orientado a proyectos integradores y perfeccionamiento del trabajo en equipo: Caso máster erasmus mundus en ingenieria mecatronica. *XXVI Congreso Universitario de Innovación Educativa en las Enseñanzas Técnicas*.
2. **Zuluaga-Gomez, Juan**, & Bonaveri, P. (2016). Sistema para la detección de señales precordiales mediante sismocardiografia. *Prospectiva*, *14*(1), 89–95.
1. Corredor, S., Valbuena, M., **Zuluaga-Gomez, Juan**, & Barrios, M. (2014). Design and construction a measurer of total body water, fat mass and fat free mass using labview. *2014 III International Congress of Engineering Mechatronics and Automation (CIIMA)*, 1–4.

PUBLICATIONS (BOOK/BOOK CHAPTER)

1. **Zuluaga-Gomez, J** et al. (2017). Tratamiento de aguas residuales mediante el proceso de fotocatalisis con dioxido de titanio (tio2). In U. A. del Caribe (Ed.). Uniautonoma, ISBN: 9789585431010.

WORKSHOPS (CONFERENCE, PEER REVIEWED)

2. Burdisso, S., **Zuluaga-Gomez, Juan**, Fajcik, M., et al. (2022). IDIAPers @ causal news corpus 2022: Causal relation identification using a few-shot and prompt-based fine-tuning of language models. *The 5th Workshop on Challenges and Applications of Automated Extraction of Socio-political Events from Text (CASE @ EMNLP 2022)*.
1. Fajcik, M., Singh, M., **Zuluaga-Gomez, Juan**, et al. (2022). Idiapers @ causal news corpus 2022: Extracting cause-effect-signal triplets via pre-trained autoregressive language model. *The 5th Workshop on Challenges and Applications of Automated Extraction of Socio-political Events from Text (CASE @ EMNLP 2022)*.

PUBLICATIONS (PRE-PRINT)

2. Nigmatulina, I., Braun, R., **Zuluaga-Gomez, Juan**, & Motlicek, P. (2021). Improving callsign recognition with air-surveillance data in air-traffic communication. *arXiv preprint arXiv:2108.12156*.
1. Madikeri, S., Tong, S., **Zuluaga-Gomez, Juan**, Vyas, A., Motlicek, P., & Bourlard, H. (2020). Pkwrap: A pytorch package for lf-mmi training of acoustic models. *arXiv preprint arXiv:2010.03466*.

PARTICIPATION IN CONFERENCES

- IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Year: 2022, 2023
- Interspeech. Year: 2020, 2021, 2023
- OpenSky Network (OSN) Symposium. Year: 2021, 2020
- XXVI CUIEET National Congress, Spain, 2018
- XVIII National and XII International Research Meeting - Colciencias, Colombia, 2015
- XII Departmental meeting of Research, Colombia, 2015
- Biomedical Engineering National Congress, Mexico, 2015
- IV International Mechatronics and Automation Congress, Colombia, 2015
- III International Mechatronics and Automation Congress, Colombia, 2014
- II International Mechatronics and Automation Congress, Colombia, 2013

REVIEWER IN CONFERENCES/JOURNALS

- Interspeech, 2023
- IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2023
- EMNLP @ CASE workshop, 2022

RESEARCH PROJECTS

- ATCO2 EU-funded Horizon 2020 project. Website.
- HAAWAII EU-funded Horizon 2020 project. Website.
- Minerva CQ company: research on spoken language understanding and robust ASR.

PROGRAMMING SKILLS

- Experienced in Bash scripting and Git (Github/Gitlab)
- Experienced in Python, PyTorch, Numpy, Google Colab and Jupyter Notebooks
- Speech Recognition toolkit: Kaldi, HuggingFace, SpeechBrain & ESPNet
- Natural Language Processing toolkit: PyTorch, HuggingFace
- Parallel experimentation with Sungrid Engine and Weight & Biases

SOFTWARE

- **PkWrap**: Python library for LF-MMI training of acoustic models for automatic speech recognition with Pytorch, https://github.com/idiap/pkwrap
- **Wav2Vec 2.0 for air traffic control communications**, https://github.com/idiap/w2v2-air-traffic
- **BerTraffic:** Bert-based speaker Diarization, https://github.com/idiap/bert-text-diarization-atc
- **CommonAccent:** accent ID with CommonVoice, https://github.com/JuanPZuluaga/accent-recog-slt2022
- **HyperConfomer:** efficient automatic speech recognition with Conformer and HyperMixer, https://github.com/JuanPZuluaga/hypermixer-speech
- **ATCO2:** a 5000-hours corpus for research on Automatic Speech Recognition and Understanding of Air Traffic Control communications. Baselines and code in: https://github.com/JuanPZuluaga/atco2-corpus

TEACHING

Teaching Assistant:

- Deep Learning Course, EE-559, EPFL (Prof. François Fleuret), Spring 2022

TALKS

- Paper presentation. Unsupersived speech recognition (abs, pdf). At *ECCS Seminar: Advanced Topics in Machine Learning*, 2022.
- Keynote: An introduction to speech-based technologies for Natural Language Processing applications. *Mexican NLP Summer School 2021*, Ciudad de Mexico, Mexico, 2021

AWARD, NOTABLE ACHIEVEMENT

1. "CommonAccent: Exploring Large Acoustic Pretrained Models for Accent Classification Based on Common Voice" paper nominated as best student paper award at Interspeech 2023
2. Ranked 2nd and 3rd place at OLR-2021 challenge task 3 & 4, 2021
3. Scholarship: Erasmus Mundus - European Union, EACEA, 2017
4. Scholarship: to attend XVI World Summit of Nobel Peace Laureates in Colombia, 2017
5. Scholarship: to attend XV World Summit of Nobel Peace Laureates in Spain, 2015
6. Distinction: rank obtained during bachelor studies: ranked 1st out of 7, 2015
7. Distinction: six distinctions and scholarships (GPA during bachelor studies), 2012-2015
8. Scholarship: by DAAD (Germany), visiting student September-October, 2014

LANGUAGE SKILLS

- Spanish: Native
- English: Bilingual Proficiency
- French: Limited Working Proficiency

MEMBERSHIPS

- Member of the International Speech Communication Association (ISCA), since 2020
- Graduate Student Member of the Institute of Electrical and Electronics Engineers (IEEE), since 2022

PATENTS

Device for Cardiac Signals Detection, ***Granted*** September 2019

- Seismocardiography system for Cardiac Signals Detection
- Registration number: NC16-175508
- Phase: Granted 2020
- Financing partners: Commerce Chamber of Barranquilla - CIENTECH

Robot for Martial Arts Training - RobPam, ***Granted*** October 2020

- Robotic humanoid to practice martial arts
- Registration number: NC201-0007622

- Phase: Granted 2020
- Financing partners: Commerce Chamber of Barranquilla - CIENTECH

SUPERVISION ACTIVITIES

Universidad Autonoma del Caribe, Barranquilla, Colombia

Mechatronic Engineering Undergraduate Program October 2016

- Development of a biomedical instrument and mobile APP for cardiac signals (SCG) and pulse oximetry monitoring in older people. **Students**: Cristhian Escalona, Dario Garcia.
- Development of a biomedical system to capture, process and visualize impedance cardiographm and electrocardiogram signals in a web page. **Student**: Juan Villalobos, Daniel Castaneda.

INTERESTS

Cooking, Hiking and Traveling, Reading (recommended books), Coffee Brewing.