Amelia **Villegas Morcillo**

✉ amelvimo@gmail.com | 👤 amelvim.github.io | ⬛ amelvim | in amelvim | 🎓 Amelia Villegas Morcillo

My research interests lie in developing computational methods to understand the natural world. So far, I have worked on topics including protein structure and function prediction using representation learning techniques, and spoken language recognition. I am looking to expand my knowledge on these topics in order to build towards the application of deep learning techniques in healthcare research.

Education

University of Granada   *Granada, Spain*

PHD IN INFORMATION AND COMMUNICATION TECHNOLOGIES, COMPUTER SCIENCE   *2017 - 2022*

- Research line: Signal processing and multidisciplinary applications
- Research group: Signal Processing, Multimedia Transmission and Speech/Audio Technologies (SigMAT)
- Doctoral Thesis: "Automatic identification of the protein fold type using representations from the amino acid sequence and deep learning techniques"

Polytechnic University of Madrid   *Madrid, Spain*

MSC IN TELECOMMUNICATIONS ENGINEERING   *2015 - 2017*

- Specialization: Bioengineering
- Master's Thesis: "Language Recognition System Optimization by Merging Phonotactic and Acoustic Information with Deep Neural Networks"

University of Granada   *Granada, Spain*

BSC IN TELECOMMUNICATION TECHNOLOGIES ENGINEERING   *2011 - 2015*

- Specialization: Telecommunication Systems
- Bachelor's Thesis: "Protein Classification based on Physicochemical Amino Acid index"

Experience

Delft University of Technology   *Delft, The Netherlands*

POSTDOCTORAL RESEARCHER   *Mar. 2023 - Present*

- Developing methods for AI-assisted Drug Discovery.

InstaDeep Ltd   *London, UK*

PHD RESEARCH INTERN   *May 2022 - Oct. 2022*

- Developed Deep Learning models for Protein Structure Prediction.
- Set up pipelines to train, fine-tune, and validate models.

Delft University of Technology   *Delft, The Netherlands*

VISITING RESEARCHER   *Aug. 2019 - Dec. 2019*

- Developed Automatic Function Prediction techniques for proteins using machine learning and neural networks.
- Compared performance of hand-crafted representations and learned embeddings of the proteins.

Polytechnic University of Madrid   *Madrid, Spain*

R&D TELECOMMUNICATIONS TECHNICIAN   *Nov. 2015 - July 2017*

- Developed Automatic Speech Recognition techniques based on Deep Neural Networks.
- Prepared specific databases and scripts for open source tools (Kaldi).

Publications

P08   *So ManyFolds, So Little Time: Efficient Protein Structure Prediction With pLMs and MSAs*   *[bioRxiv preprint]*
  TD Barrett, **A Villegas-Morcillo**, B Gaujac, L Robinson, D Adméte, E Saquand, K Beguir, and A Flajolet
  In AI4Science and MLSB workshops, NeurIPS, 2022

P07　*ManyFold: An efficient and flexible library for training and validating protein folding models*　*[paper] [github]*
A Villegas-Morcillo, L Robinson, A Flajolet, and TD Barrett
In Bioinformatics, 39(1), 2023

P06　*An analysis of protein language model embeddings for fold prediction*　*[paper] [github]*
A Villegas-Morcillo, AM Gomez, and V Sanchez
In Briefings in Bioinformatics, 23(3), pp. 1-14, 2022

P05　*FoldHSphere: deep hyperspherical embeddings for protein fold recognition*　*[paper] [github]*
A Villegas-Morcillo, V Sanchez, and AM Gomez
In BMC Bioinformatics, 22(1), pp. 1-21, 2021

P04　*Unsupervised protein embeddings outperform hand-crafted sequence and structure features at predicting molecular function*　*[paper] [github]*
A Villegas-Morcillo*, S Makrodimitris*, RCHJ van Ham, AM Gomez, V Sanchez, and MJT Reinders
In Bioinformatics, 37(2), pp. 162-170, 2021

P03　*Protein Fold Recognition from Sequences using Convolutional and Recurrent Neural Networks*　*[paper] [website]*
A Villegas-Morcillo, AM Gomez, JA Morales-Cordovilla, and V Sanchez
In IEEE/ACM Transactions on Computational Biology and Bioinformatics, 18(6), pp. 2848-2854, 2021

P02　*End-to-end prediction of protein-protein interaction based on embedding and recurrent neural networks*　*[paper]*
F Gonzalez-Lopez, JA Morales-Cordovilla, **A Villegas-Morcillo**, AM Gomez, and V Sanchez
In IEEE International Conference on Bioinformatics and Biomedicine (BIBM), pp. 2344-2350, 2018

P01　*Improved Protein Residue-Residue Contact Prediction Using Image Denoising Methods*　*[paper]*
A Villegas-Morcillo, JA Morales-Cordovilla, AM Gomez, and V Sanchez
In 26th European Signal Processing Conference (EUSIPCO), pp. 1167-1171, 2018

Conferences and Summer Schools

2022	**Oxford Machine Learning Summer School.** ML Fundamentals (27-29 June) and ML for Health (7-10 Aug.)	*Oxford, UK (Online)*
2021	**29th International Conference on Intelligent Systems for Molecular Biology / 20th European Conference on Computational Biology (ISMB/ECCB).** Presented (poster)	*Online*
2020	**28th International Conference on Intelligent Systems for Molecular Biology (ISMB).** Presented (talk)	*Online*
2019	**II Congreso Nacional / IV Jornadas de Investigadores en Formación: Fomentando la Interdisciplinariedad (JIFFI).** Presented (talk)	*Granada, Spain*
2018	**IEEE International Conference on Bioinformatics and Biomedicine (BIBM).** Attended	*Madrid, Spain*
2018	**26th European Signal Processing Conference (EUSIPCO).** Presented (poster)	*Rome, Italy*

Professional Activities & Service

Organizer	Part of Local Committee / Scientific Review Committee in IberSPEECH 2022 (Granada)
Journal Reviewer	Cell Systems 2022 ● Briefings in Bioinformatics 2021 ● Bioinformatics 2021 ● IEEE Access 2020
Conference Reviewer	IberSPEECH 2022 ● IEEE PIMRC Symposium 2018

Honors & Awards

2017	**Best master's academic record in the field of Bioengineering.** ETSIT-UPM awards	*Madrid, Spain*
2015	**Best bachelor thesis from the area of Signal Theory and Communications.** DTSTC-UGR third edition of End-of-Degree Project awards	*Granada, Spain*

Teaching

Communication Theory	University of Granada. 2021 (1.5 ECTS)
Communications I	University of Granada. 2021 (6 ECTS) and 2020 (3 ECTS)
Digital Signals	University of Granada. 2020 (1.5 ECTS)

Skills

Programming	Python, Matlab, Bash Scripting, C, LaTeX
Libraries	Numpy, Scikit-learn, Pytorch, Tensorflow, Jax
Software	Linux, Git, Docker, MS Office
Languages	Spanish (native), English (proficient)
Soft skills	Organization, Responsibility, Willingness to learn