# Valentin **Gabeur**

✉ valentin.gabeur@gmail.com  |  ⌂ gabeur.github.io

📷 gabeur  |  in gabeur  |  🐦 @vgabeur  |  🎓 gabeur

## **Wor**k Experience

### Google Deepmind
SENIOR RESEARCH SCIENTIST

*San Francisco, USA*
*Jan 2025 - Present*

- Working on **Gemini**, Google's foundational AI model.

### Meta AI
POSTDOCTORAL RESEARCHER

*San Francisco, USA*
*Apr 2023 - Dec 2024 (1.5 year)*

- Developed and trained **SAM 2** ↗ , a next-generation AI model for interactive **Video Object Segmentation** and **Image Segmentation**. SAM 2 obtains a **better accuracy**, runs at **6x the speed** and requires **3x less interactions** than the previous state-of-the-art.
- Designed the specifications and supervised the collection quality of SA-V, a **large video segmentation dataset** produced by **hundreds of human annotators** over several months.
- Prepared the **public release** of the dataset, model weights, code, paper and demo.

### Google AI
STUDENT RESEARCHER

*Grenoble, France*
*Feb 2019 - Feb 2022 (3 years)*

- Developed a novel pre-training approach and a cross-modal architecture for the task of **text-to-video** retrieval. Obtained **state-of-the-art** results on 6 public datasets and won **1st place in the CVPR 2020 Video Pentathlon** ↗ (a video retrieval competition).
- Solved **large-scale** parallelism engineering challenges to train models on 200TB (15 years) of videos with 128 TPU v3 cores using the Jax framework. Ran experiments on both Google Cloud Platform (GCP) and Google internal production cluster.
- My research led to **publications** at **ECCV-20**, **WACV-22** and **INTERSPEECH-22** with **open-sourced** releases.

### Inria
RESEARCH INTERN

*Grenoble, France*
*Apr 2018 - Jan 2019 (10 months)*

- Generated a semi-synthetic **dataset** of 3D captures in a virtual environment.
- Proposed a parameter-free **3D shape representation** for estimating human meshes from an image. **Publication** at **ICCV-19**.

### Addition Manufacturing Technologies
MECHANICAL ENGINEER

*San Diego, USA*
*Apr 2013 – Jul 2017 (4.5 years)*

- Responsible for the development of a new robotic tube bender. Study and design of innovative engineering concepts.
- Assembly and setup of robotic tube-bending cells. Optimization of machine design. Kuka robots programming.
- Technical support to sales. Presentation and promotion of the equipments at Fabtech shows. Installation at customers facility.

### Flabeg Holding GmbH
MECHANICAL ENGINEER

*Shanghai, China*
*Oct 2011 – Oct 2012 (1 year)*

- Designed new equipment for the production of automotive mirrors.

## **Edu**cation

### Grenoble-Alpes University / Inria
PH.D. IN COMPUTER SCIENCE

*Grenoble, France*
*Feb 2019 - Oct 2022*

- Thesis subject: Multi-modal learning for video understanding.
- Supervisors: Cordelia Schmid, Karteek Alahari.
- Computer vision, Natural language processing, Machine learning.

### Paul-Sabatier University
M.SC. IN ROBOTICS: DECISION AND CONTROL

*Toulouse, France*
*2017 - 2018*

- Machine-learning, Optimisation & Estimation, 3D Perception, Industrial & Mobile robotics. - **Rank 1/18**

### ICAM
M.SC. IN ENGINEERING

*Lille, France*
*2006 - 2011*

- Mechanical engineering, Automation, Computer science, Manufacturing, Project management. - **Rank 1/96**

# Skill and tools

- **Programming:** Python, C/C++, Matlab.
- **Deep Learning:** PyTorch, Jax, Tensorflow.
- **Tools:** Linux, VSCode, Tmux, Git, LaTeX, AWS, GCP, Blender, Solidworks.
- **Languages Spoken:** French (native), English (fluent), Spanish (fluent), Chinese (basics).

# Publications

🎓 scholar.google.com

[1] SAM 2: Segment Anything in Images and Videos.
Nikhila Ravi*, **Valentin Gabeur***, Yuan-Ting Hu*, Ronghang Hu*, Chaitanya Ryali*, Tengyu Ma*, Haitham Khedr*, Roman Rädle*, Chloe Rolland, Laura Gustafson, Eric Mintun, Junting Pan, Kalyan, Vasudev Alwala, Nicolas Carion, Chao-Yuan Wu, Ross Girshick, Piotr Dollár, Christoph Feichtenhofer*.
arXiv 2024

[2] AVATAR: Unconstrained Audiovisual Speech Recognition.
**Valentin Gabeur***, Paul Hongsuck Seo*, Arsha Nagrani*, Chen Sun, Karteek Alahari and Cordelia Schmid.
INTERSPEECH 2022

[3] Masking Modalities for Cross-modal Video Retrieval.
**Valentin Gabeur**, Arsha Nagrani, Chen Sun, Karteek Alahari and Cordelia Schmid.
WACV 2022

[4] Multi-modal Transformer for Video Retrieval.
**Valentin Gabeur**, Chen Sun, Karteek Alahari and Cordelia Schmid.
ECCV 2020 (Spotlight)

[5] CVPR 2020 video pentathlon challenge: Multi-modal transformer for video retrieval.
**Valentin Gabeur**, Chen Sun, Karteek Alahari and Cordelia Schmid.
CVPR 2020 Video Pentathlon Challenge Workshop - 1st place

[6] Moulding Humans: Non-parametric 3D Human Shape Estimation from Single Images.
**Valentin Gabeur**, Jean-Sebastien Franco, Xavier Martin, Cordelia Schmid and Gregory Rogez.
ICCV 2019