

# Ivan Yakovlev

[ivan@iyakovlev.dev](mailto:ivan@iyakovlev.dev) | [LinkedIn](#) | [Google Scholar](#) | [Kaggle](#) | [@vancheeck](#)

## EDUCATION

---

**Saint Petersburg State University**  
*Bachelor of Science in Quantum Mechanics*

St. Petersburg, Russia  
2012 – 2016

## EXPERIENCE

---

**Senior Research Engineer: Speech Processing**  
*Palabra AI Ltd.*

Mar. 2025 – Present  
*Barcelona, Spain*

- Solo-built production zero-shot multilingual streaming TTS system supporting 8+ languages with voice cloning, serving as the core product for Palabra's voice AI platform
- Designed and implemented cost-efficient data pipeline producing clean pseudo-labeled speech at \$0.10/hour (10–50× below industry standard), collecting 200K+ hours across 20 languages
- Architecting end-to-end speech-to-speech translation system for voice AI agent market
- Researched and developed streaming LLM-ASR model (640ms latency) surpassing conformer-transducer based models in most domains
- Invented ReDimNet2 architecture — second generation of ReDimNet with improved scalability via time-pooled dimension reshaping (B0–B6 model family, 1.1M–12.3M params)

**Sr. Machine Learning Engineer: Speech processing**  
*ID R&D Inc.*

Sep. 2019 – Feb. 2025  
*Barcelona, Spain (Remote)*

- Developed an automated data pipeline generating a 200k-identity dataset, doubling ASV performance across multiple domains
- Invented new SOTA neural network architecture, called Reshape Dimensions Network (ReDimNet), 60% faster, 35–50% smaller and 15–25% more accurate than existing architectures (ResNets, CAM++, ECAPA-TDNN)
- Designed a subnetwork approach for ResNet and similar architectures, improving downstream task performance by over 50% in target metrics
- Achieved top placements in the VoxCeleb Speaker Recognition Challenges (2022, 2023) and SASV Challenge, contributing to three publications
- Optimized speaker verification training pipelines with PyTorch and TPU, achieving a 4× training speedup within the same budget

**Software Engineer**  
*EPAM Systems*

Feb. 2017 – Mar. 2019  
*Saint Petersburg, Russia*

- Developed AR/VR applications integrating deep learning vision models using Unity3D

## ACTIVITIES

---

**VoxCeleb Speaker Recognition Challenge**

2020, 2022, 2023

- 1st place in open track (2023, 2022); 3rd place in closed track (2020)
- Results published at Interspeech workshops; oral presentation in 2023

**Spoofing-Aware Speaker Verification Challenge**

2022

- 1st place; paper accepted to Interspeech

**Kaggle Competitions Master**

2012+

- 🥉 Top 3%: 27/1189 at BirdCLEF 2023 competition
- 🥈 Top 5%: 57/1390 at Cornell Birdcall Identification challenge 2020
- 🥇 Top 1%: 8/2265 place in Deepfake Detection Challenge 2019

## TECHNICAL SKILLS

---

**Programming Languages:** Python

**ML Frameworks:** PyTorch, TensorFlow

**Infrastructure:** Docker, Git, SkyPilot, Redis, RabbitMQ

**Cloud & Compute:** GCP, Nebius, multi-GPU/multi-node distributed training (H100, A100)

**ML Platforms & Monitoring:** MLFlow, Weights & Biases, TensorBoard, Grafana/Prometheus

**Spoken Languages:** English (B2+), Russian (Native)

## PUBLICATIONS

---

- **Ivan Yakovlev**, Anton Okhotnikov, “ReDimNet2: Scaling Speaker Verification via Time-Pooled Dimension Reshaping”, submitted to Interspeech 2026
- **Ivan Yakovlev**, Rostislav Makarov, Andrei Balykin, Pavel Malov, Anton Okhotnikov, Nikita Torgashov, “Reshape Dimensions Network for Speaker Recognition”, Interspeech 2024
- **Ivan Yakovlev**, Anton Okhotnikov, Nikita Torgashov, Rostislav Makarov, Yuri Voevodin, Konstantin Simonchik, “VoxTube: a multilingual speaker recognition dataset”, Interspeech 2023
- Nikita Torgashov, Rostislav Makarov, **Ivan Yakovlev**, Pavel Malov, Andrei Balykin, Anton Okhotnikov, “The ID R&D VoxCeleb Speaker Recognition Challenge 2023 System Description”, VoxSRC-23 at Interspeech 2023
- Alexander Alenin, Nikita Torgashov, Anton Okhotnikov, Rostislav Makarov, **Ivan Yakovlev**, “A subnetwork approach for spoofing aware speaker verification”, Interspeech 2022
- Rostislav Makarov, Nikita Torgashov, Alexander Alenin, **Ivan Yakovlev**, Anton Okhotnikov, “ID R&D System Description to VoxCeleb Speaker Recognition Challenge 2022”, VoxSRC-22 at Interspeech 2022
- **Ivan Yakovlev**, Mikhail Melnikov, Nikita Bukhal, Rostislav Makarov, Alexander Alenin, Nikita Torgashov, Anton Okhotnikov, “LRPD: Large Replay Parallel Dataset”, ICASSP 2022