Compiled using LaTeX

Yan Yang

Graduate Studies @ KAUST
Machine Learning R&D Engineer @ Deepcell Inc.

yy4jobs@gmail.com
+1(202)876-8456

Interests

- Machine Learning, Deep Learning
- Informatics, Information Theory
- Image Processing, Visualization
- CS Topics in Interdisciplinary Complex Systems (Cognitive Science, Social Science, Neuroscience, Biology, etc.)

Skills

Machine Learning:

- Convolutional NN ●●●●○
- NLP ●●●○○
- Deep Reinforcement L. ●●●○○
- Tensorflow ●●●●○
- Keras ●●●●●
- PyTorch ●●●●○
- Chainer ●●●○○
- Numpy ●●●●●

Data Processing and Visualization:

- Dimension Reduction ●●●●○
- Coloring ●●○○○
- SQL ●●●●●
- Pandas ●●●●●
- Apache Arrow ●●○○○
- Matlab ●●○○○
- Excel ●●●●●
- Matplotlib ●●●●○
- WebCanvas & Chart.js ●●●●○

General Programming:

- C/C++ ●●●●○
- Python ●●●●●
- Java ●●●○○

Web Development:

- JavaScript ●●●●●
- Flask ●●●●●

Formal Education / Degree

2023 – 2024 **Ph.D. in Computer Science** — **KAUST**
Courses: Deep Learning for Visual Computing | Spatial Data Science with R | Modeling and Simulation of Biosystems **Supervisor**: Prof. Dominik Michels. **CGPA:** 3.550
Dynamical Systems | Computer Vision | Speech Recognition

2021 – 2022 **Ph.D. in Computer Science and Engineering** — **U. of Nevada, Reno**
Courses: Neurobiology | CS Topic: Bioinformatics | Independent Study: Computer Vision | Elements of Research Computing | CS Topic: Mass Detection in Mammograms | CS Topic: High-Performance Networking Systems **Supervisors**: Prof. Mircea Nicolescu and Prof. George Bebis. **CGPA:** 4.000
Computer Vision | Medical Imaging

2017 – 2019 **M.Sc. in Computer Science** — **Georgetown U.**
Courses: Algorithms | Computer Hardware & System Architecture | Comp. Corpus Linguistics | Empirical Methods in NLP | Stat. Machine Translation | Adv Semantic Representation | Automatic Reasoning | Machine Learning | Deep Reinforcement Learning | Independent Study: Advanced Analytics | Statistic Machine Learning (Ph.D. seminar, approved by instructor) **Supervisor**: Prof. Grace Hui Yang **CGPA:** 3.701
Machine Learning | NLP | Information Retrieval

2002 – 2006 **B.E. in Information Engineering** — **Southeast U., China**
Courses: Linear Algebra & Space Analytic Geometry **(scholarship endowed)** | Advanced Mathematics | Probability & Statistics | Numerical Computing & Modeling | Digital Image Processing | Speech Signal Processing | Signals & Linear System | Electromagnetic Field & Waves | Digital Systems & Course Design | Electronic Circuits & Comprehensive Experiment | Microcomputer Systems & Interfaces | Fundamentals of Circuit | Analog Electronic Circuits | Modern Psychology | Modern Biology | College Physics | Music Theory | Practice of MATLAB | etc. **CGPA:** 2.974

Complementary Education / Certification

2024 - now	Elements of Information Theory (Ongoing)	**Book by Thomas Cover**
2024 - now	Introduction to Functional Analysis (Ongoing)	**MIT 18.102**
2022	Human Research (Social Behavior Research Investigators and Key Personnel Group, Basic)	**CITI Program - U. of Nevada, Reno**
2020	Certified Generalist Software Engineer (Top 3%)	**TripleByte**
2017	Gender Study Summer School (writing: A+; other: A)	**Harvard U.**
—-	**—Coursera—**	—-
2023	Probabilistic Graphical Models 2 (94/100)	**Stanford U.**
2023	Probabilistic Graphical Models 1 (100/100)	**Stanford U.**
2022	Computational Neuroscience (99.57/100)	**U. of Washington**
2021	Convolutional Neural Networks (100/100)	**Deeplearning.AI**
2021	Mathematical Biostatistics Boot Camp I (100/100)	**Johns Hopkins U.**
2020	Genomics: Decoding the Universal Language of Life (Unlicensed)	**UIUC**
2020	Finding Hidden Messages in DNA (Unlicensed)	**UC San Diego**

Publications

May, 2025 **Modeling Limits in Phenomenological Contexts - A Statement on *Improving Searcher Struggle Detection via the Reversal Theory.*** **(Preprint)**
Yan Yang *ResearchGate*

Dec, 2024 **Improving Searcher Struggle Detection via the Reversal Theory.** **(Article)**
Luo Jiyun, **Yan Yang**, Valerie Nayak, Grace Hui Yang *Discover Computing*

Apr, 2022 **Two-Step Data Augmentation for Masked Face Detection and Recognition: Turning Fake Masks to Real.** **(Conference Paper)**
Yan Yang, George Bebis, Mircea Nicolescu *SCITEPRESS*

Dec, 2021 **Feature Modulation to Improve Struggle Detection in Web Search: A Psychological Approach.** **(Preprint)**
Luo Jiyun, **Yan Yang**, Valerie Nayak, Grace Hui Yang *arXiv*

- RESTful
- React.js
- GoLang

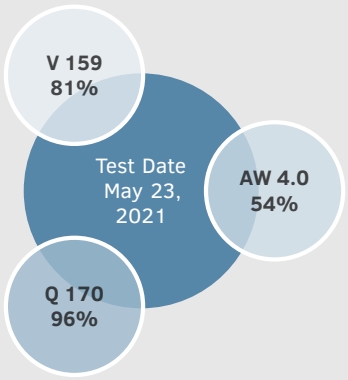
Project Management:

- Agile
- Docker
- LaTeX
- Git

Computing Resource Management:

- Nginx
- Google Cloud
- Slurm
- Kubernetes
- Linux

Test Scores



Profiles

Languages

Yue Chinese (Mother Tongue)

Mandarin (Mother Tongue)

English (Proficient)

Cantonese (Elementary)

French (Elementary)

Research and Teaching Work Experience

Sep, 2023 – Apr, 2024 **Research Fellow** — **Computational Sciences Group (KAUST)**
- Researched to automate the **synthesis of diseased leaf images** used by **classification** models in agriculture. Explored methods to learn **diffusion-reaction systems** (i.e., PDE parameters) from data.
- Proposed ideas on Arabic **Speech Recognition**, including **acoustic source separation** and data augmentation in the **spectrogram**.

Jul, 2022 – Sep, 2022 **Volunteer Research Assistant** — **Visual Perception Lab (U. of Nevada, Reno)**
- Trained for **BrainVoyager** usage and **human research ethics**.
- Helped with **software and package installations**.
- Read papers on **Representational Similarity Analysis**, **Multivoxel Pattern Analysis**, human brain's **face/object perception**, **fMRI**, etc., and discussed research ideas with other lab members based on their existing code and data.

May, 2021 – Aug, 2022 **R&D Intern** — **Deepcell Inc.**
- Trained **GAN models**, supplementing data for cell image classifiers.
- Led the development of a comprehensive **CNN embedding layer visualization** tool, used company-wide by **data** and **bioinformatic scientists** for clustering, correlation analysis, etc., and greatly improved their productivity.

Jan, 2018 – May, 2019 **Research Assistant** — **Georgetown U.**
- InfoSense Lab: 1) Translated **psychological reversal theory** into statistic languages to analyze search engine **user behaviors** (**search logs**) and improve **user struggle detection**; 2) Revamped search engine's **VR interface** for **behavioral data collection**, and provided **web technique** support for its SIGIR '18 conference demonstration;
- Massive Data Institute: "Victims of Violence" tweets analytics: **occupational analysis** and **database management.**

Jan, 2021 – Dec, 2021 **Teaching Assistant** — **U. of Nevada, Reno**
- CS 302 (Spring 2021), Data Structure in C++.
- MATH 127 (Spring 2022), Precalculus II (**recitation classes**).

Jun, 2016 – Aug, 2016 **Web Development Mentor** — **DataMesh (Chinese tech company)**
Mentored new employees on webpage building.

Jun, 2006 – Oct, 2008 **Volunteer Psychology Supporter (Community Service)** — **Ayaohelp.com**
- Provided moral **support** & helped with life event strategic analyses.
- Posted essays/poems on **personal development** journeys, which were listed by the website host to attract site visitors.

Scholarship, Presentation, and Peer Review

2024 — Reviewed one **Computer Vision** article for two rounds. — *IEEE TNNLS*

Sep 2023 — Doctoral fellowship. — KAUST

Jan 2003 — Course scholarship for **Linear Algebra & Space Analytic Geometry**. — Southeast U. (China)

Apr 2022 — **Conference IMPROVE 2022 Talk** — *SCITEPRESS*
Presentation for *Two-Step Data Augmentation for Masked Face Detection and Recognition: Turning Fake Masks to Real*.

Course Projects (*blue* ⋆= research project)

NLP / Computational Linguistics
- Shift-Reduce **semantic parser** (natural language) using **deep reinforcement learning**; (link) ⋆
- A Neural-Network **linguistic morphology cutter (Keras)**; (link) ⋆
- Course assignments: **N-gram**, **Perceptron**, Machine translation (e.g. **Estimation-Maximization**, **machine-aided translation**, etc.)

Computer Vision
- Analyzed **data augmentation** methods for **mammogram** mass **segmentation**. Quantitatively applied **adversarial training**; ⋆
- Used **image-to-image translation** GAN models to generate training images for **masked face detection** task; ⋆
- **Deep Reinforcement Learning** video game agents;
- **C++ Image processing** (Gaussian blur, flip, sharpening, etc.)

Analytics /
- Construction of I/O log datasets (Blue Water HPC); Future

Data Crawling	throughput prediction (exact bytes): **AutoRegressive Integrated Moving Average**, **LSTM multi-layer RNN**s, & **Decision Tree Regressor**s; Performance analysis per dataset/model. ⋆ • "#MeToo in Tweets" analytics (topic modeling); (link) ⋆ • **Crawlers**: free books by Springer during COVID-19. (link)
Machine Learning	• Java **standard libraries** only: Build **ML learners** & reproduce paper results; **Automatic reasoning** projects.
Computer Systems	• C++: Multi-thread experiments | Windows MineSweeper | Linux embedded programs; • Assembly hardware experiments: traffic light, etc.; • VHDL: **8-bit CPU**; • Verilog: CPU Branching, MEM, ROM.
Non-CS	• Proposal on **brain vision connections** (BIOL675 Neurobiology) ⋆ • 5-minute **French** (**linguistic features** introduction) • **MATLAB Graphs** for Middle school Algebras

Engineering Work Experience

Oct, 2022 – May, 2023 — **Machine Learning R&D Engineer** — **Deepcell Inc.**
• Led the **system design** of the remodeled **Neural Network embedding** visualization system, which is to be used seamlessly in the Cluster-Based Sorting product. When the imaging phase of an imaging+sorting run is done, we use this design to pause the run and start a **data exploration session** to select sorting criteria.
• My system design covered decisions on **data storage media**, **database schemas**, **API specs**, **serialization/deserialization** options, **caching**, **parallelizing** large calculations, and others for incremental developing plans and long-term goals.

May, 2021 – Aug, 2022 — **R&D Intern** — **Deepcell Inc.**
• For the **CNN embedding layer** visualization tool, independently coordinated **user demand collection**, full-stack development, and **usability verification**.
• The second version refactored the app to **improve security** and separated large-chunk data processing into a **socket streaming service** for multiprocessing.

Jun, 2020 – Jan, 2021 — **Member of Technical Staff** — **Pure Storage**
• Implemented feature: import/restore **volume snapshots** from company disks into Kubernetes.
• Enhancement/bug fixing: **multipath device** problem, upgrading **bash scripts** to Python3, etc.
• Led the **Github release process** of Pure Service Orchestrator v6.0.3.

Jun, 2016 – Aug, 2016 — **Development Engineer** — **DataMesh**
• Led the development of a news publishing system implemented as a **React-Redux** app;
• Solved Apache deployment conflicts among static files, React router hack, and Python service.

Jan, 2016 – Jun, 2016 — **Software Engineer** — **Meican.com**
• Designed and implemented a **message push system** using **WebSocket**, **AngularJS**, and **GoLang Micro Service** for the meal website, which solved user complaints that pages don't sync the latest meal status;
• Code cleanness and reusability: encapsulated **WebSocket** in an **AngularJS Factory**.

Jul, 2015 – Jan, 2016 — **Full Stack Engineer** — **Colorful Clouds Tech Ltd.**
• Independently took charge of a weather report app with **dynamic charts** for precipitation and other weather metrics (AngularJS, Flask, NodeJS, AliCloud, Nginx, MySQL).
• Implemented a compatibility patch for **data API** v2, providing continued support to v1 clients.
• Automatic periodic **collection of API health data** into the Redis database.
• **Precipitation interpolation** using **NumPy**.

Oct, 2014 – Jun, 2015 — **Frontend Developer** — **2 different companies**
• Developed **WeChat minigames** for marketing activities.
(https://github.com/DIMPLY/MiniGame-Zillionaire)
• Built webpages for an information management app, using D3.js for **data reports and charts**.

Sep, 2013 – Oct, 2014 — **Web Development Engineer** — **2 different companies**
Developed **C++-based** Windows apps (ID card verifier, Word Solitaire game) with Linux server.

Other Work Experience

Dec, 2008 – Dec, 2012 — **English-Chinese Translator** — **Self-Employment**
• Published dozens of translation articles on yeeyan.org
• Took charge of the translationorganization of *Women in The 19th Century* (Margaret Fuller 1843).
• Translated & proofread marketing materials, users' & developers' manuals, TV show subtitles, texts in software & games. Customers included Grand Strong and OK Translation.

Apr, 2008 – Jan, 2009 — **Scientific Secretary** — **Tongji University**
Supported department research and student activities as an administrative secretary.