Mingyu Cha
#204, 144-1, Gangseo-ro, Gangseo-gu
Seoul, Republic of Korea
+82(10) 6375-6954
mgcha85@gmail.com

EDUCATION

University of Central Florida – Orlando, FL
Ph.D. in Computer Science (Incomplete) – December 2020
GPA: 3.34/4.0

- Set multiple histone database and developed miRNA TSS detection algorithm by using deep learning
- Retrieved RNA-seq and CAGE tags on multiple cell lines and tissues and analyzed the pattern between miRNAs and protein coding genes from the large-scale database
- Built database query methods at GPU level using CUDA, facilitating efficient data handling for large-scale datasets.
- Programmed GUI Road defect detection software by using deep learning to automate labor work
- Inspected ground penetrating radar signal to detect cavity signal underground
- Developed groundwater level measurement device circuit and software by using quantitative measurement toolkit to prevent sinkhole

Gwangju Institute of Science and Technology – Gwangju, South Korea
M.S. in Electrical Engineering – February 2014
GPA: 3.54/4.5

- Assisted Agency for Defense Development to develop an LPI (Low Probability of Intercept) radar
- Corrected errors of the automatically scanning machine for radar and radiometer in the lab
- Developed RADAR signal processing and entire simulation tool by using MATLAB OOP
- Devised automatic gear shift for bicycle with EMG sensors
- Observed the Digital Integrator with modified Romberg integration

Hankuk University of Foreign Languages – Yongin, South Korea
B.S. in Computer Science – February 2012
GPA: 3.93/4.5

- Minor in Business Administration, incorporating the integration of technology solutions in business processes.

EXPERIENCE

LG Display – Seoul, Korea
Research Engineer (Generative AI and NLP) – October 2023 – Present

- Innovative AI Solutions: Developed and implemented advanced Generative AI applications, significantly improving predictive accuracy and operational efficiency.
- Large Scale Data Management: Managed extensive datasets using AWS cloud services, ensuring efficient data manipulation, visualization, and storage.
- Cross-functional Leadership: Led a team of engineers and analysts, aligning Generative AI initiatives with strategic business objectives.
- Recognized as a key talent by the CEO of LG Display in 2023 and 2024, resulting in a salary increase
- Spearheaded the development of the "Statistical Map Control" project, a real-time monitoring system designed to track process status, report foreign object presence, and maintain equipment functionality, using XGboost, SHAP, correlation techniques
- Initiated the use of NVIDIA DGX clusters for deep learning model training, achieving remarkable improvements in training times and model accuracy.
- Developed an automatic web app service for statistical analysis, which involved preprocessing tasks such as coordinate conversion and duplicate removal of inspection machine data to analyze extensive foreign material data within the process. CUDA is used for efficient preprocessing of large datasets, pre-matching large quantities of data, and enabling user-defined matching criteria, using Dask and CUDA for rapid data processing, delivering near real-time analysis results to users
- Delivered lectures in the AI Big Data Room on topics such as accelerating Python analysis and query optimization.
- Developing a machine learning task automation framework utilizing PySpark
- Developed visualization functions using Superset and CanvasJS for web-based result display and facilitated email notifications for analysis updates.
- Implemented the MLOps concept, developing a solution on a distributed server and currently conducting a pilot operation
- Architected and implemented an ML platform for real-time data processing and analysis of semiconductor thickness, utilizing 2D interpolation and CUDA for efficient GPU resource utilization, enhancing the scalability and performance of ML model training and deployment.
- Conducted in-house training on Git and Visual Studio Code usage, while also preparing Docker for configuration management purposes
- Optimized ML platform performance, focusing on GPU cluster management, through advanced techniques in data ingestion, model training, and deployment scalability
- Developed the Python package 'lgddata', featuring functions for querying comprehensive data sets and offering unique utilities and information not available in the database
- Developed API for overall defect analysis package by using Django
- By optimizing the existing inefficient code with the concepts of numba, CUDA, algorithms, and computer architecture, the speed is improved by more than 10 times
- Development of statistical analysis methods lacking in existing route analysis
- Color coordinate prediction model using deep learning and XGboost
- Proposal and development of analysis methods for real-time sensor data (f-domain, filtering)
- Development of simulation tool for RGB value optimization
- First use and development of analytics server in-house using Spark and Dask
- Developed of a new algorithm to analyze long-term trends in real-time sensor data using low-pass filters and regression
- Mentored three research engineers, guiding their project work and professional development, enhancing team productivity and innovation

LexisNexis Risk Solutions – Boca Raton, FL
Intern – May 2020 – July 2020

- Developed CI/CD DevOps pipelines and automated error reporting systems using Python and C++, ensuring high availability and error tracking efficiency.

TJ Media – Seoul, South Korea
Software Engineer – January 2012 – July 2012

- Developed embedded systems and optimized C-based code for commercial software products, improving performance and reducing resource consumption.
- Focused on developing algorithms for UI animations and data handling.

TECHNICAL SKILLS

- **Programming Languages:** C++, Python, Rust, C/C++, Java, Mojo
- **Libraries and Tools:** Boost, OpenCV, Leptonica, CUDA, OpenCL, TensorFlow, PyTorch
- **Development Tools:** Git, Docker, Visual Studio Code, Linux, Pandas, Spark, SQL
- **Specializations:** Algorithm Design, Data Structures, AI Acceleration on GPUs, Deep Learning Optimization
- **Languages:** Korean (native), English (fluent)

KEY STRENGTHS

- Strong experience with high-performance C++ development, especially in algorithmic design and data structures.
- Expertise in developing scalable systems and optimizing computational efficiency using modern C++ standards and libraries like Boost.
- Familiarity with reverse engineering and debugging at the assembly level for performance tuning and system improvements.
- Extensive work with GPUs (CUDA, OpenCL) for parallel computing and optimization of large data systems.

LEADERSHIP & PROJECTS

- **Mentored engineering teams** at LG Display, providing technical guidance and oversight on project deliverables and innovation strategies.
- **Founder, Cha Data Research Corporation** – Led a startup providing financial data services using advanced web scraping and data aggregation algorithms.

PUBLICATIONS

- Guhathakurta, S., Adams, L., Sivakumar, A., **Cha, M**; Fiadeiro, M, Hu, H, Kim, Y (2021). Precise epigenomic editing by a newly developed modular epigenetic toolkit. (under review)
- Talukder, A., Wang, S**., Cha, M**., Li, X., Hu, H., (2019). Computational annotation of miRNA transcription start sites (Editing)
- **Cha, M**., Barham, C., Li, X., Hu, H., (2019). D-miRT: A Two-Steam Convolutional Neural Network for MicroRNA Transcription Start Site Feature Integration and Identification (under review)
- Barham, C., **Cha, M**., Li, X., Hu, H., (2019). Application of Deep Learning Models to microRNA Transcription Start Site Identification, ICBCB 2019 conference
- Jung, W., **Cha, M**. (2018), New Weapon Requirements Recommendation Using M&S and Big Data Analytics, IISE Annual Conference

- Mokhtari, S., Wu, L., **Cha, M**. (2016), Comparison of Supervised Classficiation Techniques for Vision-Based Pavement Crack Detection, Transpotation Research Board, Washington, D.C.
- **Cha, M**., Kim, Y. (2015), A low probability of intercept (LPI) radar signal analysis using Independent component analysis (ICA), Korea Electromagnetic Engineering Society Conference
- **Cha, M**., Kim, Y. (2015), WVD, STFT, wavelet, QMF, and CSA algorithm comparison and analysis to detect an LPI radar signal, Korea Electromagnetic Engineering Society ConferenceGuhathakurta, S., Adams, L., Sivakumar, A., **Cha, M**; Fiadeiro, M, Hu, H, Kim, Y (2021). Precise epigenomic editing by a newly developed modular epigenetic toolkit. (under review)
- Talukder, A., Wang, S**., Cha, M**., Li, X., Hu, H., (2019). Computational annotation of miRNA transcription start sites (Editing)
- **Cha, M**., Barham, C., Li, X., Hu, H., (2019). D-miRT: A Two-Steam Convolutional Neural Network for MicroRNA Transcription Start Site Feature Integration and Identification (under review)
- Barham, C., **Cha, M**., Li, X., Hu, H., (2019). Application of Deep Learning Models to microRNA Transcription Start Site Identification, ICBCB 2019 conference
- Jung, W., **Cha, M**. (2018), New Weapon Requirements Recommendation Using M&S and Big Data Analytics, IISE Annual Conference

- Mokhtari, S., Wu, L., **Cha, M**. (2016), Comparison of Supervised Classficiation Techniques for Vision-Based Pavement Crack Detection, Transpotation Research Board, Washington, D.C.
- **Cha, M**., Kim, Y. (2015), A low probability of intercept (LPI) radar signal analysis using Independent component analysis (ICA), Korea Electromagnetic Engineering Society Conference
- **Cha, M**., Kim, Y. (2015), WVD, STFT, wavelet, QMF, and CSA algorithm comparison and analysis to detect an LPI radar signal, Korea Electromagnetic Engineering Society Conference