ROBERT ELLIOT ENGLISH

eenglish@gmail.com +1 6507993617 San Francisco Bay Area, California, USA elliot-english-a821b230

EDUCATION

Stanford University - *Ph.D. Computer Science* September 2008 - June 2013

"A Scalable and Adaptive Discretization for Free Surface Incompressible Flow Simulation using Overlapping Cartesian Grids"

University of British Columbia - *BSc. Honours Computer Science* September 2003 - May 2008

"Animating Developable Surfaces using Nonconforming Elements"

WORK EXPERIENCE

Syntiant Corp - *Fellow* September 2022 to present

I am a technical lead at Syntiant Corporation where I direct machine learning related software efforts.

Key responsibilities:

- Researching new methods to accelerate large language models using software-only strategies and custom hardware.
- Developing new neural network model architectures for computer vision and audio processing that maximize accuracy while minimizing latency, static footprint, and runtime memory usage.
- Architecting the model inference software stack included in the Neural Decision Processor (NDP) firmware. This includes adding support for heterogeneous computing, improving power efficiency, adding support for more network operators, and implementing new software/hardware acceleration strategies.
- Adding support for new hardware platforms to the software-only model inference stack.
- Implementing new ML tasks (e.g. CV: classification, localization, segmentation, reconstruction, identification, Audio: classification, identification) using both machine learning, and heuristic approaches.
- Architecting the ML ops/model training and evaluation software infrastructure.

Pilot AI Labs, Inc - *Cofounder & Chief Technical Officer* August 2015 to August 2022

I led the technical efforts at Pilot AI Labs, developing machine learning based computer vision solutions for embedded devices in a wide range of applications. The work resulted in millions of units deployed with a variety of security and analytics applications employing computer vision, along with a successful exit.

Key responsibilities:

- Managed an agile engineering organization with 25 engineers from highly diverse backgrounds.
- Directly engaged with customers from initial meetings, to development and production support. Supported applications from a wide variety of spaces such as consumer home automation, industrial security and analytics, and government/defense.
- Designed compact neural network architectures for each core ML model (image classification, detection, segmentation, pose estimation, embedding vector identification). This also included developing methods for multiscale problems with large resolution imagery (100M pixels+) and small feature scales (5-10 pixels).
- Developed the core inference platform architecture with a flexible and modular user API. With the same codebase, the platform scales from Nvidia GPUs to midtier NPUs, to ARM CPUs to microcontrollers with severe memory and instruction limits.
- Invented novel methods for at-runtime data-sensitive dynamic sparse evaluation methods for off the shelf models.
- Invented novel training objectives and regularization techniques to improve model sparsity.
- Implemented neural architecture search (NAS) with objectives for model accuracy and compactness.
- Optimized high performance operator kernel code using intrinsics and assembly (e.g. Intel CPUs (AVX/SSE), ARM CPUs (NEON, ARMv5/6 eDSP), Qualcomm Hexagon DSPs (HVX), Ingenic CPUS (MIPS/custom)).
- Integrated vendor NPU SDKs for kernel level acceleration (e.g. Nvidia CuDNN/CuBLAS, Senselab NNAPI) and model level acceleration (e.g. VeriSilicon NPUs/OpenVX, Ambarella CVFlow, Qualcomm SNPE, Nvidia TensorRT).
- Developed low latency image preprocessing algorithms for processing raw camera buffers (e.g. YUV420p to RGB + crop/resize).
- Implemented a custom model training stack using Nvidia CUDA/CuDNN for fast model updates, with online data augmentation and importance sampling.
- Implemented an ML ops system with continuous training and analysis of production models. Also developed a Mongo/MySQL/Ceph backed sample data storage and query service.
- Developed object tracking algorithms using ML and traditional algorithms, with short term spatial tracking and long term visual reidentification.
- Developed a camera calibration and object 3D localization stack for multicamera tracking.
- Developed visual analytics heuristics for implementing business logic and pattern of life analysis.
- Developed a video conferencing autoframing system using visual object tracking and multispeaker audio direction of arrival.
- Developed a drone flight control system with target object state estimation using Kalman and other custom filters and object following control logic.

MetaMind, Inc - *Senior Data Scientist* September 2014 to July 2015

I implemented a model training system (pre-tensorflow/pytorch availability) using CUDA and CuBLAS acceleration. In parallel, I developed models for NLP tasks such as sentiment classification and question answering using RNNs, recursive models, and memory networks.

Lawrence Berkeley National Laboratory - *Postdoctoral Scholar* September 2013 to August 2014

Under the supervision of Drs. Phillip Colella and Hans Johansen, I developed high order finite volume methods on mapped grids for climate modelling and viscous compressible flow simulation.

Feinberg Day Alberti & Thompson LLP - *Technical Consultant* February 2012 to March 2014

I provided claim analysis and code review for intellectual property litigation.

- Matter Intellectual Ventures v. Capital One Inc. - Banking systems.
- Matter Intellectual Ventures v. PNC Bank Inc. - Banking systems.
- Matter Intellectual Ventures v. Motorola Inc. - Android Linux kernel, framework and application code.
- Matter HTC Corp. v. Apple Inc. - Android Linux kernel, framework and application code.

Stanford University, Computer Science - *Research Assistant* September 2008 to August 2013

Under the supervision of Dr. Ronald Fedkiw, I developed an adaptive method for solving the incompressible Navier-Stokes equations using overlapping Cartesian grids that scales to thousands of cores. I additionally developed algorithms for monolithic fluid-structure interaction problems, developed continuous collision detection and handling methods for rigid body dynamics, and explored fast multi-level solvers for contact problems.

University of British Columbia - *Honours Thesis Research Assistant* May 2007 to August 2008

Under the supervision of Dr. Robert Bridson, I developed a method for simulating inextensible materials using nonconforming finite elements in order to avoid the locking problem.

University of British Columbia - *Undergraduate Reseach Assistant* May 2006 to August 2007

With Drs. John Lloyd and Sidney Fels, I was part of the team developing the biomechanical simulation software Artisynth. I developed robust methods for collision detection and handling between rigid and deformable bodies. I also implemented methods for incompressible finite element simulation and incomplete matrix factorizations.

NGRAIN Corporation - *Software Engineer* January 2005 to August 2005

I worked as part of the product development team developing the user interface for interacting with and editing volumetric 3D models. I developed a simulation environment for familiarizing technicians with component assembly and disassembly. In addition to writing automated testing scripts, I participated in regression testing and QA done at the end of each release cycle.

INVITED TALKS

- *Deep Learning based Computer Vision for Autonomous Control Systems*, IEEE Control System Society Seminar Series, Santa Clara Valley Chapter, San Jose, CA, 2017.
- *Deep Learning for Natural Language Processing: Efficient Implementations and GPUS*, CS224D, Stanford University, CA, 2015.
- *Deep Learning for Everybody*, GPU Technology Conference, Santa Clara, CA, 2015.
- *Simulating Beyond* 10^9 *Degrees of Freedom*, IEEE Control System Society Seminar Series, Santa Clara Valley Chapter, San Jose, CA, 2013.
- *A Scalable Adaptive Spatial Discretization for Incompressible Flow using Overset Cartesian Grids*, Lawrence Berkeley National Laboratory, Berkeley, CA, 2013.
- *A Scalable Adaptive Spatial Discretization for Incompressible Flow using Overset Cartesian Grids*, NASA Applied Modeling & Simulation Seminar Series, Mountain View, CA, 2013.
- *Numerical Physical Simulation*, IEEE Control System Society Seminar Series, Santa Clara Valley Chapter, Sunnyvale, CA, 2011.

PUBLICATIONS & PATENTS

- (Patent) B. Pierce, E. English, A. Kumar, J. Su. System and Method for Improved General Object Detection Using Neural Networks. US20150170002A1.
- 6 additional patents submitted.
- A. Kumar, O. Irsoy, J. Su, J. Bradbury, R. English, B. Pierce, P. Ondruska, M. Iyyer, I. Gulrajani, R. Socher, *Ask Me Anything: Dynamic Memory Networks for Natural Language Processing*, arXiv (2015), http://arxiv.org/abs/1506.07285v2
- R. E. English, L. Qiu, Y. Yu, and R. Fedkiw, *Chimera Grids for Water Simulation*, ACM SIGGRAPH Eurographics Symposium on Computer Animation (SCA), 85-94 (2013).
- R. E. English, L. Qiu, Y. Yu, and R. Fedkiw, *An adaptive discretization of incompressible flow using a multitude of moving Cartesian grids*, J. Comp. Phys. 254, 107-154 (2013).
- R. E. English, M. Lentine, R. Fedkiw, *Interpenetration Free Simulation of Thin Shell Rigid Bodies*, IEEE TVCG 19, 991-1004 (2013).
- A. Robinson-Mosher, R. E. English, and R. Fedkiw, *Accurate Tangential Velocities for Solid Fluid Coupling*, ACM SIGGRAPH/Eurographics Symposium on Computer Animation (SCA), edited by Eitan Grinspun and Jessica Hodgins, 227-236 (2009).
- R. E. English and R. Bridson, *Animating Developable Surfaces Using Nonconforming Elements*, Proc. ACM SIGGRAPH (2008).

- K. van den Doel, F. Vogt, R. E. English and S. S. Fels, *Towards Articulatory Speech Synthesis with a Dynamic 3D Finite Element Tongue Model*, In Proceedings of ISSP 06, pages 59-66, (2006).

TECHNICAL EXPERTISE

- Computer vision: Image classification, object detection (e.g. SSD), object tracking, object identification (Face recognition), 3D reconstruction (SfM, SLAM), CNN based models (e.g. BiFPN), RoI based models.
- Audio processing: Audio classification, feature extraction (STFFT/Melbin), CNN based models, RNN based models, DOA and VAD state estimation.
- Natural language processing: Large Language Models for Q&A and unstructured symbolic processing, using both sparse causal transformer and RNN (e.g. RWKV) type models.
- Robotic control: PD control, Kalman filtering, deep reinforcement learning, dynamic modeling.
- Computational solutions to partial differential equations: e.g. computational fluid dynamics and solid mechanics, strongly coupled multiphysics systems, continuous contact and collision detection/handling, fracture. Spatially and temporally adaptive methods (e.g. AMR, tetrahedral unstructured, ALE, overset grid methods).
- Numerical optimization: large scale iterative solvers with linear and nonlinear constraints, interior point optimization methods.
- Numerical linear algebra: parallel solutions to sparse systems with greater than 10^9 variables using preconditioned iterative methods, eigenvalue/SVD problem algorithms, multigrid solvers.
- Computational geometry: spatial acceleration structures (e.g. octree, kdtree, bsptree), Delaunay triangulation, direct Voronoi diagram generation, remeshing algorithms.
- Computer graphics: Scanline rasterization, raytracing, radiosity/global illumination, octree based NeRF.
- Parallel Computation: MPI, OpenCL, OpenMP, pthreads, CUDA.
- Programming Languages: C/C++, Python, C#, Java, Perl, Fortran, COBOL, Bash/Zsh.
- Web Technologies: HTML 5, Javascript, PHP, React.JS, Angular.JS, REST, SOAP.
- Database Technologies: MySQL, Postgres, MongoDB.
- Programming Environments: Linux/GNU Toolset, Android App/NDK/Platform, Mac OSX/Darwin, Windows API.
- Machine Learning Toolkits: Tensorflow, TFLite, PyTorch.
- Scientific Libraries: PETSc, Eigen, Chombo, PhysBAM, Ceres.
- Cloud Environments: AWS (EC2, S3, Document DB, EFS), GCP (Storage, Vertex AI).

INTERESTS

- Mountain Biking, Hang Gliding, Windsurfing, Cooking, Gardening, Music (Guitar, Trumpet).
- Sustainable Energy Development, Material Science, Aerospace.