SGER: Grid-to-grid neural networks for innovative pose invariant face recognition

Khan M. Iftekharuddin[1], Elizabeth L. Threlkeld[2], and William E. White III[3]
University of Memphis, Intelligent Systems and Image Processing Lab, Memphis, TN 38152, USA. *[1]E-mail: iftekhar@memphis.edu* *[2]E-mail: ethrlkld@memphis.edu*
[3]E-mail: wewhite@memphis.edu

ABSTRACT: This paper is a trip report for the 2008 IREE trip taken by The University of Memphis. For three months, Dr. Iftekharuddin and two of his students, Elizabeth Threlkeld and William White, studied at The University of Wollongong in Wollongong, NSW, Australia. They continued their work on pose invariant facial recognition and were able to learn new techniques to apply to their existing work. This report discusses specific accomplishments by the faculty member, students, interactions with the host lab, as well as broader impacts of this international collaboration.

INTRODUCTION

This grant was awarded to The University of Memphis, as a supplement to our existing grant for research on facial recognition. In our current SGER award we are investigating new classes of practical tools such as the cellular simultaneous recurrent net (CSRN), which learn mappings from data defined over an image grid to outputs defined over a grid or to global summary variables. In the ongoing SGER research, more powerful mapping of image data will be adapted so as to maximize performance in the pose-related face recognition task. Simultaneously, time permitting, our secondary goal is to investigate how brains process pose invariance for face recognition and relate that knowledge to the inner working principles of CSRN.

Under our IREE award, we studied the brain-like biologically-inspired approach to pose invariant face recognition by building an international collaborative research with the *Visual Signal and Information Processing* (VSIP), laboratory http://www.snrc.uow.edu.au/centres/VSIPRG/, within the *Institute for Telecommunications and Information Technology Research* (TITR), at University of Wollongong, Australia. The primate brains perform distortion-invariant image recognition quite easily in addition to hundreds of other complex tasks. In principle, one may capture this unfolding wonderful working model of the vision system in a distributed intelligent image processing framework for more robust pattern recognition. Thus, the objective of *this proposed IREE research* was to develop an understanding of biologically-inspired approach to pattern recognition in general and pose invariant face recognition in particular.

Our *Intelligent Systems and Image Processing* (ISIP) group https://umdrive.memphis.edu/iftekhar/faculty/research.html at University of Memphis and the VSIP laboratory in Australia have been following each others research on pattern recognition, biologically inspired vision, and image processing for some time. Our foreign partner, Dr. Bouzerdoum, and Dr. Iftekharuddin had already met at several international conferences and discussed common research issues. Prof. Bouzerdoum has been investigating, modeling and implementing vision systems and visual information and image processing algorithms since 1988. He has published over 200 technical articles in international journals, book and Encyclopedia chapters, and conference proceedings. He has attracted, in collaboration with other colleagues, more than AU$ 7.5 million total funding for commercial and basic research. He has supervised 15 PhD dissertations, 6 Master theses, and more than sixty honors theses. He has been lecturing in signal processing, image processing and pattern recognition since 1992. He developed or contributed to a number of new subjects in optimum signal processing, image and video processing, multimedia signal processing, and pattern recognition.

Our ongoing and/or anticipated research and education outcomes for both the faculty member and students are as follows.

1) Improved understanding of bio-motivated pattern recognition.
2) Better formulation of ongoing research techniques and methods in pose related face recognition
3) Initiation of collaborative research and joint proposal activities with our Australian counterpart to target funding agencies such as NSF and Australian Research Council (ARC)
4) Broader perspective and understanding on international research, culture and environment
5) Appreciation of diversity and cooperation
6) An MS thesis and an honors undergraduate thesis
7) Joint publications with host laboratory

We chose to work with VSIP lab in TITR, at the University of Wollongong, Australia. TITR is the largest university-based information and communications technology research center in the Southern Hemisphere [1]. The travelers on the IREE trip were Dr. Khan Iftekharuddin, Elizabeth Threlkeld, and William White. Dr. Iftekharuddin is an associate professor at the University of Memphis. Elizabeth is a graduate student while William is an undergraduate student in the same department. Dr. Iftekharuddin traveled between May 15, 2008 and August 5, 2008. He had to come back to USA little early to conduct a technical conference, chair a session and present paper at the SPIE annual meeting held from August 10 – August 15 in San Diego. Both Elizabeth and William traveled from May 15, 2008 to August 15, 2008.

RESEARCH ACTIVITIES AND ACCOMPLISHMENTS OF THE INTERNATIONAL COOPERATION

After surviving close to twenty-four hours of traveling to make it from Memphis, TN to Wollongong, Australia and recovering from jet lag somewhat, we got right to work at our host laboratory. Dr. Iftekharuddin was given an office and Elizabeth and William were given places to work in a laboratory. Unfortunately there was not space in the laboratory the rest of Dr. Bouzerdoum's students worked in, so they were in a nearby building. Being in a separate room slightly hampered interactions between the students initially, but this was overcome with daily lunches in the break room. These less formal daily interactions balanced nicely with the more formal group meetings that happened roughly once a month. In the more formal meetings different students would make presentations on their current work. We were given a chance to present at the meetings. In the first meeting, our presentation was an introduction to who were, where we were from, and work currently going on in our ISIP lab. In a subsequent meeting, Dr. Iftekharuddin gave a more formal and much longer presentation on some of his recent works. In the final group meeting we attended, Elizabeth and William both gave presentations on the work they had accomplished during their visit. In addition to lunches and formal group meetings, Elizabeth and William met on a more individual basis at least once a week with other students whose research directly pertained to their work. Dr. Iftekharuddin generally met with Dr. Bouzerdoum once a day, and with several other faculty members on a weekly basis.

Research conducted on site at the University of Wollongong can be divided into two main categories, improvements with the CRSN network, and shunting inhibitory convolutional networks. We were able to improve CSRN code to increase the number of training samples possible to use. We also clustered the mazes created that program, and are now working on using these clusters to reduce the number of training data needed. Quite a bit was learned about shunting-inhibitory neurons, and we are planning on incorporating this knowledge into the CSRN.

William's focus while at The University of Wollongong was improving our Cellular Simultaneous Recurrent Neural Network (CSRN). In searching for performance improvements to the CSRN, he first began by simply looking at ways to improve the MATLAB code we were working with. By analyzing the publicly available CSRN MATLAB code, he was able to find a few major areas to improve. These improvements made a greater difference when working with higher dimensional data such as images. One challenge faced was that the maze traversing problem the CSRN was originally designed for was not nearly as complex as an image of a face. As a result, the code had to be changed to allow for higher dimensional data by optimizing the memory usage of the program and changing some system variables. After the improvements the computer was able to handle 50% more training samples.

Even after these improvements, the problem was we were still unable to handle as much image data as we would like to process. To alleviate this problem, we began discussing a technique our host VSIP lab had used, which was training networks with weighted clustered data [2]. This allowed them to train their networks with less data but achieve performance similar to training with much larger original dataset. In order to explore this technique, we decided to study the technique with the original application of the current published CSRN code e.g. maze traversal problem. However, we ran into a few problems. First, we were unsure of how to create useful clusters from mazes until we found a published article by Mannor et al [3]. Using a similar approach we were able to create meaningful clusters from a maze. Our preliminary results on clustering a simple maze are shown in Figures 1 and 2.

However, we then ran into a problem of how to input our clustered data to the CSRN network. After further study, we finally concluded that the current CSRN network would be unable to successfully use our clustered data in its current structure. From there, we began looking into ways to change the structure of CSRN network to allow the use of our clustered data. We are currently working on re-designing the publicly available CSRN code to fit our needs of large scale image data.

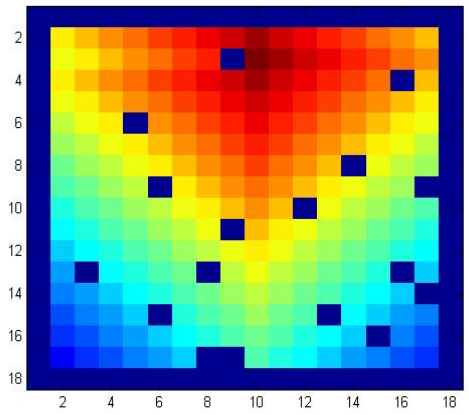

Figure 1: A known maze used to train the CSRN. The dark blue cells represent obstacles and the dark red cell represents the goal. The rest of the cells are color codes based on the number of steps to the goal from that cell.

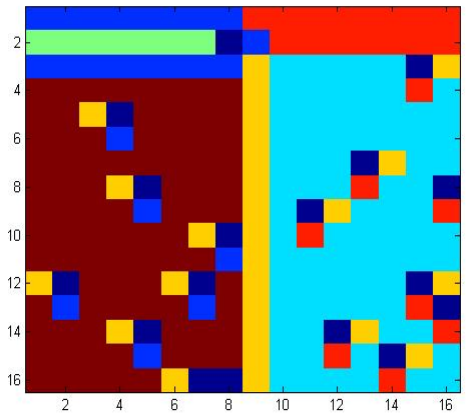

Figure 2: The same maze, clustered according to distance and direction to the goal. The obstacles are dark blue again. Cells around the obstacle often belong to a different cluster than its neighbors because the direction had a heavy influence in clustering

Elizabeth spent the majority of her time focusing on convolutional neural networks (CoNN), and more specifically the integrated shunting inhibition neuron in CoNN, as studied by our host Dr. Bouzerdoum and his group [4]. In nature, neurons receive both excitatory inputs as well as inhibitory ones. The shunting neuron attempts at mimicking both these activities in a neural network. CoNNs in general mimic a mammalian visual cortex, and their output is multiple lower dimensional images known as feature maps. After gaining a through knowledge of this specific network, Elizabeth began trying to use individual feature maps to classify facial images. Using a Principle Component Analysis (PCA) based approach [5] on the standard probes in the FERET database [6], she preformed classifications. Initially the images were run through the PCA classifier before putting them through the CoNN. Next each of the CoNN output feature maps were individually classified. The highest performing feature maps where then concatenated together to see if this further improved accuracy. Figure 3 shows the classification accuracy of an individual feature map on the *fafb* FERET probe set, which deals with the aging of subjects. At rank 200 we obtain a classification rate of around 90%. This can be compared to Figure 4 which is the classification accuracy of three feature maps concatenated together. Note in Fig. 4, three maps slightly outperform the classification rate of the single feature map. Additionally, initial results so far confirm hypothesis of Dr. Bouzerdoum, that feature maps are a good way to deal with illumination problems. Each individual feature map outperformed the original images on the FERET illumination probe set (*fafc* probe). Current future plans for this work include using preprocessing techniques on the images before they are divided into feature maps by the CoNN, and to investigate other classifiers than just PCA. Further plan includes exploring inclusion of inhibitory neuron ideas into CSRN to improve its pattern classification performance.

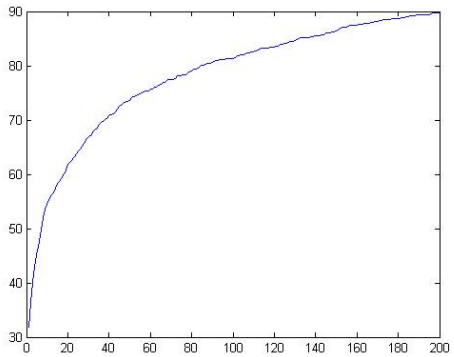

Figure 3: Classification rate of an individual feature map up to rank 200 on the FERET fafb probe set

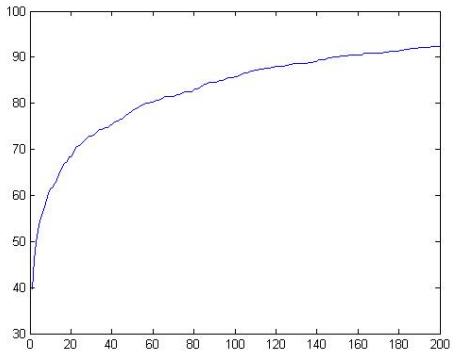

Figure 4: Classification rate of three concatenated feature maps up to rank 200 on the FERET fafb probe set.

BROADER IMPACTS OF THE INTERNATIONAL COOPERATION

The IREE award promoted diversity in the ISIP lab right from the start by bringing an undergraduate into the lab. It expanded the scope by bringing in new ideas from the host VSIP lab that we are working on incorporating with our CSRN. The two main ideas are the use of clustering as a form a preprocessing, and the use of a shunting inhibition neuron to improve large scale image data handling and pattern classification of CSRN. Dr. Bouzerdoum and his students have found that clustering training data offers two advantages such as i) the NN pattern classification performance for huge amount of data can be as good as that of a reasonable amount of data; and ii) such NN can also train on imbalance data. William worked with this idea and has clustered the values in the mazes created by the CSRN. Elizabeth worked with the shunting inhibition neuron concept. This closer mimics biology and in theory will improve accuracy in the CSRN network once it is incorporated with it.

Future interactions between the ISIP and VSIP lab are already being planned. Drs. Bouzerdoum and Iftekharuddin applied for and received a small grant from the University of Wollongong which will allow them to visit each others lab and continue research collaboration, joint proposal writing and other activities in 2009. Further, another junior faculty member from Dr. Bouzerdoum's lab in Australia is planning to spend part of his sabbatical at our lab in Memphis in 2009.

We traveled to Australia expecting most things to be very similar to what we experience in the United States, and for the most part we found this to be true. The most shocking difference we found was internet access. We are used to having unlimited internet access on campus from our labs and found this to not be the case at the University of Wollongong. They have a quota system for student's access which allocates around 400 Megabytes per student per semester. Students can get approval for additional internet use added to their quota, assuming their quota was used up by legitimate academic use. To get additional internet, ones previous internet use is reviewed, and if it was genuinely used for academic purposes more will be granted. This seemed to be a common idea throughout Australia and similar things were found for home internet use. Instead of paying for unlimited high speed internet access like we can in the United States, you had to pick from various plans that were priced based on how much you were allowed to download with them. If you exceeded this amount per month, your internet was dropped down to dial up speeds if not totally blocked.

Traveling on this trip increased Elizabeth and William's familiarity with the metric system. While they were both comfortable with it on an academic level, it was a somewhat challenging experience being expected to use it for daily things like temperature that they were used to thinking of in degrees Fahrenheit instead of degrees Celsius. Neither thinks they would be as comfortable using the metric system for daily units like they are now if they had not lived in another country where they were expected to use metric units.

One similarity we immediately saw was the overwhelmingly large number of foreign graduate students compared to Australian students. In the VSIP lab there was only one Australian citizen, with the others coming from various places in Asia and one student from France. We similarly have very few United States citizens in the ISIP lab at home, and we fit right in as foreign engineering students in their lab. More social conversations definitely had a global feel to them. When discussing how things were different in Australia from the United States we also got to learn about how they differ from various countries in Asia as well as from Europe. Gas prices were a big difference, with ours being much lower than most other places (about half of what they are in Australia).

Australian currency was a welcomed surprise, the one and two dollar coins especially. Paper bills began at five dollars and from there followed increments similar to what we have in the U.S. While we have dollar coins at home, they are not widely used like they were in Australia. The option of larger amounts of money being formed easily by a few coins allows the acceptance of ten dollar bills in most vending machines. Coins consisted of five, ten, twenty, and fifty cent pieces. Not having twenty-five cent piece coins took a little while to get used to, but not having to deal with pennies was nice. Hopefully the U.S. will go with this trend in the future and eliminate them from our currency as well.

While people in Australia also speak English as the main language, there was some confusion at times over specific phrases. Examples include calling strollers prams, calling the trunk of a car the boot of the car, and having to order take away instead of carry out. We also have come back home with a new respect for the ozone layer. From government sponsored TV commercials we learned that about one out of every two Australians will develop cancer at some point in their life because of the very high rates of skin cancer found there. Elizabeth especially was amazed and how much she would get sunburned by being outside for short periods of time during the winter.

DISCUSSION AND SUMMARY

Our most significant accomplishments were learning two new techniques that we are continuing to study further and hope to incorporate into our pose invariant facial recognition research soon. We learned a clustering technique to help reduce the amount of training data needed to fully train our network. We also learned about a shunting inhibitory neuron that we hope will further make CSRN more bio-motivated and robust. Being on site at the University of Wollongong, gave us the opportunity to meet in person with people to get an understanding of their work. It would have either taken much longer or, been almost impossible to gain this level of understanding through emails or over the phone. We are excited about future interactions and collaborations with the VSIP lab. Drs. Bouzerdoum and Iftekharuddin have already begun planning subsequent trips to visit each other in the near future, and they have received a small grant from the University of Wollongong for this purpose. In addition, another faculty member from our host lab in Australia is also planning to spend part of his sabbatical in Memphis working on collaborative research and joint funding proposals. Furthermore, Drs. Iftekharuddin and Bouzerdoum, along with others, are organizing a joint *IEEE Symposium on Computational Intelligence for Multimedia Signal and Vision Processing* in 2009. We hope to start publishing our collaborative works soon.

The May 2008 pretrip conference was helpful, but we were leaving in about two weeks after the conference. So we felt like it was too late for some of the information to be useful. The breakout group sessions based on where you were going where the most helpful. We met Mike Roddewig, a student at the Michigan Tech, in our breakout sessions and learned he was planning on visiting the University of New South Whales in Sydney, Australia, which is about an hour away from the University of Wollongong. We exchanged email addresses at the conference, and then met with him several times while in Australia. Most of our interactions with him were more social ones, but we spent one day touring the lab he was visiting, and learned about his research on field programmable gate arrays (FPGAs). He in turn came down and spent a day visiting our lab at the University of Wollongong. Meeting with Mike gave us a good opportunity to compare our similar experiences of studying at a school in Australia. We strongly encourage future IREE participants to visit other participants in the program while abroad in their respective countries.

ACKNOWLEDGEMENTS

We would like to thank our NSF program officer, Dr. Paul Werbos. Our current grant award number is ECCS-0715116.

We would also like to thank Dr. Bouzerdoum and everyone else in the host VSIP lab, especially Steve and Giang for all the time they spent meeting with us about their published work.

REFERENCES

1. "Telecommunications and Information Technology Research". 7 February, 2007. The University of Wollongong. 12 September 2008. <http://www.uow.edu.au/research/profile/UOW008927.html>.
2. Nguyen, Giang H., Abdesselam Bouzerdoum, Son Lam Phung, Efficient Supervised Learning with Reduced Training Exemplars", Proceedings of the 2008 IEE World Congress on Computational Intelligence, 1-6 June 2008, Hong Kong, China.
3. Mannor, Shie, Ishai Menache, Amit Hoze, and Uri Klein, "Dynamic abstraction in reinforcement learning via clustering", Proceedings if the twenty-first international conference on Machine Learning, p. 71, 4-8 July 2004, Bannf, Alberta Canada.
4. Fok Hing Chi and Abdesselam Bouzerdoum, "A Brain-Inspired Visual Pattern Recognition Architecture and its Applications." *Pattern Recognition Technologies and Applications: Recent Advances*. Eds. Brijesh Verma and Michael Blumenstein, IGI Global, 2008.
5. Delac K., Grgic M., Grgic S., Independent Comparative Study of PCA, ICA, and LDA on the FERET Data Set, *International Journal of Imaging Systems and Technology*, Vol. 15, Issue 5, 2006, pp. 252-260.
6. Phillips, P. J., H. Moon, P. J. Fauss, and S. Rizvi, "The FERET evaluation methodology for face recognition algorithms", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 10, October 2000.

BRIEF BIOGRAPHIES OF RESEARCHERS

Dr. Khan M. Iftekharuddin is an associate professor in the department of electrical and computer engineering at the University of Memphis (U of M). He also holds a joint appointment with the joint graduate program in biomedical engineering at the U of M and University of Tennessee at Memphis. He further serves on the faculty of bioinformatics program at U of M. He was an assistant professor of computer science and electrical and computer engineering at North Dakota State University (NDSU) before joining to the U of M in the fall of 2000. He obtained an early promotion and tenure at the U of M in the fall of 2003. He is named the Researcher of the Year in college of Engineering and Architecture for 2000 at NDSU. Much of his research had focused on different aspects of computer vision and signal/image processing problems. He is the principal author of more than eighty refereed journal papers and conference proceedings; and multiple book chapters. He is a fellow of SPIE, a senior member of IEEE and a member of IEEE CIS and OSA.

Elizabeth L. Threlkeld received a BSE in Computer Engineering from Tulane University in 2006. At Tulane Elizabeth was IEEE Student Chapter President for the 2005-2006 school year, and was recognized as the chapter's student member of the year for the same time period. She was also awarded the Texas Instruments Award for the Outstanding Senior in the Computing Sciences. In 2005, during the semester Tulane was closed following hurricane Katrina, she worked in the Navy Primary Standards Lab at North Island Naval Air station in San Diego, CA. Currently Elizabeth is a graduate student at the University of Memphis working on a MS in Computer Engineering, where she works under Dr. Iftekharuddin in both the Intelligent Systems Image Processing Lab and the Embedded Systems Lab.

William E. White III received his high school diploma from White Station High School in Memphis, TN in 2004. He is currently working on dual degrees in electrical engineering and computer engineering, as well as a second major in physics, at The University of Memphis. He is on track to graduate in Spring 2009 with departmental and university honors. William is currently the treasurer of his fraternity (Phi Gamma Delta) and is a member of many honors organizations including Phi Kappa Phi, Omicron Delta Kappa, Who's Who in American Colleges and Universities.