Northwestern | THE GRADUATE SCHOOL

Application for Admission

App Type **New Student**

Submitted Date **12-26-2018**

App ID# **77505519**

Intended Status **Full-time**

Entry Quarter **Fall 2019**

Prior TGS Applicant (Program)

Last Name **Lu**

First **Yingjing**

Middle

Gender Pronouns (US only)

Birthdate **01-19-1997**

Gender **Male**

Program **Computer Science: MS**

Specialization/Area of Interest **Artificial Intelligence and Machine Learning**

Cluster

Secondary PhD (MEAS Only)

MS Consideration (MEAS Only)

JD/PhD **No**

DPT/PhD **No**

Fee Waiver

US Vet/Active Forces

Ethnicity **Asian**

Hispanic **No**

Citizenship **CHINA**

Visa

Citizenship Status **International Student**

Country of Birth **CHINA**

Green Card #

Current Address
1635 Denniston Street
Apt2 Floor 2
Pittsburgh, PA, 15217
UNITED STATES

Current Phone **4125194745**

Cell Phone

Email Address **yingjing_lu@163.com**

Permanent Address
Building 2 Unit 3 Yun Qu Yuan 2 Qu
Hui Long Guan, Changping District
Beijing, 102208
CHINA

Permanent Phone

Preferred Phone Number **Current Phone Number**

Previous Institution	From	To	Field of Study	Level	Degree	Date
Carnegie Mellon University	08-31-2015	05-20-2019	Information Systems		US Bachelor of Science	05-20-2019

Cumulative UG GPA	**3.47**	UG Junior/Senior Year GPA	**3.68**
Cumulative UG GPA - Unconverted		Max UG GPA Scale	
Cumulative Grad GPA			
Cumulative Grad GPA - Unconverted		Max Grad GPA Scale	

Letters of Recommendation

1. **Roger Dannenberg** **rbd@cs.cmu.edu**
2. **Larry Heimann** **profh@cmu.edu**
3. **Kang Guo** **kangguoroo@gmail.com**
4.
5.

Are you interested in studying with specific faculty members? (List names below)

1. First Name **Ian** Last Name **Horswill**
2. First Name **Kenneth** Last Name **Forbus**
3. First Name Last Name
4. First Name Last Name

Please indicate the highest level of education completed by your parent(s) or guardian(s) (the one or two people most responsible for raising you)

First individual's highest level of education completed: **Graduate or professional degree**

If other, please explain:

Second individual's highest level of education completed: **Bachelor's degree or equivalent**

If other, please explain:

Language	Reading	Writing	Speaking

Self-Reported Test Scores

GRE Gen	08-04-2018	Verbal	157	76	Quant	170	96	A.W.	4	59				
GRE Sub						LSAT								
TOEFL	07-12-2014	Ovr	110	Read	29	List	27	Speak	26	Writ	28	IELTS		Ovr
GMAT		Tot			Verb			Quant			A.W.			I.R.
MCAT		Bioscience			Verbal			Physical Science						

Please list any honors you have been awarded

Dean's List, High Honor - 2018
Dean's List 2016, 2018
Midwest Trading Competition - Case 2 First Place - 2018
MIT Trading Competition - Top 20% - 2017
Hack Princeton - Best First Hack - 2016

Have you applied for or been awarded an external fellowship?

Yes ○ No ⊗ If yes, please specify;

Please describe your plans for the future.

I want to enter industry as an researcher or developer in the field of computer perception.

Other Universities Applied (in preferred rank order)

1. School Drop Down	**Carnegie Mellon**	5. School "other"	
2. School Drop Down	**Stanford University**	6. School "other"	
3. School Drop Down	**University of California-Berkeley**	7. School "other"	
4. School Drop Down	**University of Michigan**	8. School "other"	
	University of Illinois-Urbana		

Academic misconduct? Yes ○ No ⊗ Convicted of crime? Yes ○ No ⊗

If answered yes, applicant is asked to upload explanation. If uploaded, explanation will be attached to end of application PDF.

Statement of Purpose

My long term interest involves building scalable systems that extend human perception. To be more capable of building systems as such, I need the ability to research and implement models in computer perceptions and more extensive knowledge in distributed computer systems. On one hand, I want to acquire more knowledge to bring more high quality contributions in computer perceptions to the research community. On the other hand, I want more courseworks and project opportunities in distributed computer systems to convert research contributions into real world products. The M.S. CS program at Northwestern University fits my interests as it allows me to get opportunities to reach professors related to computer perceptions, and offers me various courses in computer systems that I can apply to build large scale systems.

In my undergraduate study, I was able to dive into areas of generative models for vision understandings through individual research projects. I also delivered original research contributions at major conferences such as AAAI and IEEE ICDM. I proposed an architecture that utilized variational autoencoders to encode information into latent spaces and used GANs as latent space mappers to transport information in a predefined knowledge graph. I encountered many instances of mode collapse in both variational autoencoder and GAN training. To solve training problems, I learned various normalization and exponential decay methodologies and achieved desired results. I presented my work in the IEEE ICDM 2018 Workshop. Another project was on using structural similarity measures as image reconstruction loss functions. Structural similarities have been used as loss functions but limited by their insensitivity to luminance resulting in color loss in multi channel image generations. Preexisting works resolved this issue through combining this metric with other losses such as L1, but with little theoretical justification. I was able to explain where the limitation came from and proposed a modified structural similarity loss function with both theoretical and empirical justifications. I published my preliminary results in the AAAI 2019 Student Track; I hope to research on new loss functions that enhance training quality for high dimensional data.

In addition to my individual projects, I worked with professor C.F. Larry Heimann on anomaly detections. I was able to dive into different online learning algorithms and got exposed to the phenomena of concept drift (changes in patterns). I built a novel model with multimodal loss for detecting concept drifts. The algorithm was able to detect 80% changes in attack pattern in the KDD dataset. This experience allowed me to learn to model concept drifts by modeling them as a multimodal problem. My current research with professor Roger Dannenberg allows me to apply my learned concept drift models into a new area to detect changing patterns in music onsets. By modeling music onsets as a multimodal concept drift problem, I was able to build models that outperformed the baseline classifier. This ongoing onset detection project allows me to explore ways to engineer spectrogram features with one dimensional convolutional neural networks. I intend to bring my feature engineering skills in spectrograms into future projects.

In my graduate study, I aspire to further extend my research interest in computer perception to human level reasoning and interaction. The M.S. program at Northwestern University fits my interest by offering me opportunities to work with faculty members that align with my interest through thesis track. In particular, I am interested in working with professor Kenneth Forbus and professor Ian Horswill. Professor Forbus's new work on *Action Recognition from Skeleton Data via Analogical Generalization over Qualitative Representations* provides me with a novel perspective to learn explainable sketches for human action classification. Wanting to extend on the direction for interpretable machine learning further, I am interested in working on future projects that use this paradigm to disentangle and interpret decision boundaries for deep learning models. Another direction I would like to work on is to create intelligent agents that could quantize and reason human emotions and philosophies. In this direction, I found the series work done by professor Ian Horswill on interactive narrative a great match. I am interested in working on future projects that leverage the power of intelligent agent for art creation and interactive assistance.

Additionally, the M.S. CS program provides me with fundamental courses in distributed computer systems and parallel computation that I am excited to extend my background in. These courses will compliment the skills I started developing as an intern, and these skills are essential to my future professional career. I worked as a software engineering intern at Goldman Sachs, ZPU Data and Carnegie Mellon University. I was able to quickly apply my software development background that I learned through my undergraduate courses and quickly delivered successful, working products. More importantly, I also learned that scalability is a crucial part of software design for real world businesses that depend not only on advanced algorithms, but also on large amount of data and high throughput. To be able to convert a research model into into real world applications, I need to build systems to be extensively parallelized.

The M.S. CS program offers me various courseworks in distributed systems, parallel programming to help me to further extend my engineering capability. I want to utilize those courses to extend the knowledge in computer systems I obtained from graduate study to better prepare me for tackling problems in industries after graduation. I would also want to utilize the system engineering perspective to propose new models from my research that can be more tolerant and scalable in real world distributed use cases.

I envision that during my graduate study, I will be able to deliver more extensive and solid research results. I also aspire to start building working systems from my research results that benefit real world use cases.

Official Academic Transcript from:
CARNEGIE MELLON UNIVERSITY
UNIVERSITY REGISTRAR`S OFFICE
5000 FORBES AVE.
PITTSBURGH, PA 15213-3890

TELEPHONE: 412-268-4138

Document Type: THIRD-PARTY SECURE PDF

Official Academic Transcript of:
YINGJING LU
Transcript Created: 20-Dec-2018

Intended Recipient:
YINGJING LU
5032 FORBES AVE
PITTSBURGH, PA 15289-2000

E-Mail: yingjing_lu@163.com

Requested by:
YINGJING LU
5032 FORBES AVE
PITTSBURGH, PA 15289-2000

E-Mail: yingjing_lu@163.com

Delivered by:
CREDENTIALS SOLUTIONS, LLC / TRANSCRIPTSNETWORK
Under Contract To:
CARNEGIE MELLON UNIVERSITY
Order Number: 8MV777794- 1
Telephone: (847) 716-3005

Statement of Authenticity
This Official Academic Transcript in Portable Document Format (PDF) was requested by the individual identified above in compliance with the provisions of the Family Educational Rights and Privacy Act of 1974 as Amended and in conformance with the prescribed ordering procedures of Carnegie Mellon University who has contracted with Credentials Solutions, LLC of Northfield, IL for electronic delivery of Official Academic Transcripts in PDF form via Credentials' TranscriptsNetwork(tm). You may verify the authenticity of our relationship with Carnegie Mellon University by visiting their website at http://www.cmu.edu/hub/records/transcripts/.

Colleges and Universities
If you are an accredited post-secondary academic institution, please be advised that you are receiving this transcript as a "Third-Party" receiver. Since you are not registered in our TranscriptsNetwork or eScrip-Safe network, additional security provisions have been added to this document to prevent content copying or alteration. You also are not permitted to print the document without watermark protections or add notations to the document when saving to your student information system. Should you wish to receive future documents through our electronic networks without these additional security features, please register your institution at https://escrip-safe.com/signup. Please note that there is no cost to be a Receiving Institution.

Privacy and Other Information
This Official Academic Transcript is for delivery to the above-named "Intended Recipient". If you are not the "Intended Recipient", please notify the University Registrar's Office at Carnegie Mellon University. You are not permitted to copy or alter this document. You may not forward this document or disclose its contents to any person or organization other than the "Intended Recipient" without the express written permission of the student. If this document is copied or printed, the words "PRINTED COPY" will appear in the replicated transcript image. You may verify the authenticity of this electronic document and have us independently certify that the document has not been altered since its creation by going to https://www.credentials-inc.com/cgi-bin/cicgipdf.pgm?VALID and following the instructions for transcript certification.

In the interest of security and privacy, we delete this Official Academic Transcript from our server 48 hours after it is initially downloaded excluding weekends and holidays. If a replacement is subsequently needed, the requesting party must order another transcript from Carnegie Mellon University. If you have any questions about this document or require further assistance, please contact Credentials Customer Service at (847) 716-3005. Our operators are available from 7:00 am to 7:00 pm Monday through Thursday and 7:00 am to 5:00 pm on Fridays (Central Time).

Carnegie Mellon University

Student Name: Yingjing Lu

Birth Date/Month: 19-Jan

College: Dietrich College of H&SS

Department: Dietrich College Interdisciplinary

Major: Information Systems

Beginning of Undergraduate Record

ADVANCED PLACEMENT CREDIT

DPT	CRS #	COURSE TITLE	UNITS	FINAL GRADE
MSC	21120	DIFFERENTIAL INT CAL	10.0	AD
MSC	21122	INTEGRTN & APPROX	10.0	AD
STA	36201	STATS REASON PRACTCE	9.0	AD
	38012	AP ENVIRONMENTAL SCI	9.0	AD
ML	82231	INTERMED CHINESE I	12.0	AD
ML	82236	INTS CHS LNG CLT INT	9.0	AD
		TOTAL AP / TRANSFER CREDIT	**59.0**	

Fall 2015

DPT	CRS #	COURSE TITLE	UNITS	FINAL GRADE	QUALITY POINTS
CS	15110	PRINCPLS OF COMPUTNG	10.0	A	40.0
MSC	21127	CONCEPTS OF MATHMTCS	10.0	A	40.0
ISH	67100	INFO SYS FRSH WRKSHP	1.0	P	0.0
ENG	76100	READ WRITE ACAD CNTX	9.0	B	27.0
ENG	76130	IMMIGRANT FICTIONS	9.0	B	27.0
PSY	85102	INTRO TO PSYCHOLOGY	9.0	B	27.0
CMU	99101	CMPTNG CARNEGIE MELL	3.0	P	0.0

	UNITS PASSED	UNITS FACTORED	FINAL QPA	TOTAL POINTS
Semester	**51.0**	**47.0**	**3.43**	**161.0**
Cumulative	**110.0**	**47.0**	**3.43**	**161.0**

Spring 2016

Dean's List

DPT	CRS #	COURSE TITLE	UNITS	FINAL GRADE	QUALITY POINTS
CS	15112	FNDMTLS OF PGMG & CS	12.0	A	48.0
MSC	21259	CALCULUS IN 3-D	9.0	B	27.0
ISH	67250	INFO SYSTEMS MILIEUX	9.0	A	36.0
ISH	67353	IT & ENVIRON SUSTAIN	6.0	A	24.0
ENG	76101	INTERPRETN & ARGMNT	9.0	A	36.0
HIS	79104	GLOBAL HISTORIES	9.0	B	27.0

	UNITS PASSED	UNITS FACTORED	FINAL QPA	TOTAL POINTS
Semester	**54.0**	**54.0**	**3.67**	**198.0**
Cumulative	**164.0**	**101.0**	**3.55**	**359.0**

Fall 2016

Dean's List

DPT	CRS #	COURSE TITLE	UNITS	FINAL GRADE	QUALITY POINTS
CS	15122	PRIN IMPRTV COMPTATN	10.0	B	30.0
ROB	16456	REALITY COMPUTING	12.0	A	48.0
DES	51261	COMM DES FUNDMNTLS	9.0	A	36.0
MUS	57337	SOUND RECORDING	6.0	A	24.0
ISH	67262	DTABASE DSGN/DEV	9.0	A	36.0
BA	70311	ORGNZTN BEHAVIOR	9.0	B	27.0

	UNITS PASSED	UNITS FACTORED	FINAL QPA	TOTAL POINTS
Semester	**55.0**	**55.0**	**3.65**	**201.0**
Cumulative	**219.0**	**156.0**	**3.59**	**560.0**

Spring 2017

DPT	CRS #	COURSE TITLE	UNITS	FINAL GRADE	QUALITY POINTS
CS	15213	INTR CMPUTER SYSTEMS	12.0	C	24.0
MSC	21241	MATRC & LINR TRNSF	10.0	C	20.0
MSC	21270	INTRO MATH FINANCE	9.0	B	27.0
STA	36217	PROB THEO RNDM PROC	9.0	B	27.0
MUS	57338	SOUND EDITING & MSTR	6.0	B	18.0
ISH	67272	APPLCTN DESGN & DEV	9.0	C	18.0
ML	82273	INTRO JAP LANG/CULT	9.0	A	36.0

	UNITS PASSED	UNITS FACTORED	FINAL QPA	TOTAL POINTS
Semester	**64.0**	**64.0**	**2.66**	**170.0**
Cumulative	**283.0**	**220.0**	**3.32**	**730.0**

Fall 2017

DPT	CRS #	COURSE TITLE	UNITS	FINAL GRADE	QUALITY POINTS
MLG	10701	MACHINE LEARNING	12.0	B	36.0
LTI	11641	MACH LRNG TXT MINING	12.0	B	36.0
EPP	19452	EPP PROJECT	12.0	A	48.0
MSC	21260	DIFFRENTL EQUATIONS	9.0	B	27.0
MSC	21370	DSCRETE TIME FINANCE	9.0	B	27.0

	UNITS PASSED	UNITS FACTORED	FINAL QPA	TOTAL POINTS
Semester	**54.0**	**54.0**	**3.22**	**174.0**
Cumulative	**337.0**	**274.0**	**3.30**	**904.0**

Spring 2018

Dean's List, High Honors

DPT	CRS #	COURSE TITLE	UNITS	FINAL GRADE	QUALITY POINTS
MLG	10708	PROB GRAPHCL MODELS	12.0	A	48.0
CS	15323	COMPTR MUSIC SYSTEMS	9.0	A	36.0
ISH	67373	IS CONSULTING PROJ	12.0	A	48.0
ISH	67495	ADV PROJ INFO SYSTMS	12.0	A	48.0
ML	82234	CHINA AGE REFRM	9.0	A	36.0
ML	82333	INTRO CHNSE LNG CUL	12.0	A	48.0

	UNITS PASSED	UNITS FACTORED	FINAL QPA	TOTAL POINTS
Semester	**66.0**	**66.0**	**4.00**	**264.0**
Cumulative	**403.0**	**340.0**	**3.44**	**1168.0**

Summer 1/All 2018

DPT	CRS #	COURSE TITLE	UNITS	FINAL GRADE	QUALITY POINTS
ISH	67505	INFO SYSTMS INTRNSHP	3.0	P	0.0

	UNITS PASSED	UNITS FACTORED	FINAL QPA	TOTAL POINTS
Semester	**3.0**	**0.0**	**0.00**	**0.0**
Cumulative	**406.0**	**340.0**	**3.44**	**1168.0**

John R. Papinchak
University Registrar

Carnegie Mellon University

Student Name: Yingjing Lu

Birth Date/Month: 19-Jan

Fall 2018

DPT	CRS #	COURSE TITLE	UNITS	FINAL GRADE	QUALITY POINTS
MLG	10500	SENIOR RESERCH PRJCT	12.0	A	48.0
MLG	10703	DEEP RL & CONTROL	12.0	A	48.0
LTI	11663	APPLIED MACH. LRNG.	12.0	A	48.0
LTI	11731	MACHINE TRANSLATION	12.0	B	36.0

	UNITS PASSED	UNITS FACTORED	FINAL QPA	TOTAL POINTS
Semester	**48.0**	**48.0**	**3.75**	**180.0**
Cumulative	**454.0**	**388.0**	**3.47**	**1348.0**

End of Undergraduate Record

John R. Papinchak
University Registrar

Carnegie Mellon University

Transcript Information

This information is also available at *http://www.cmu.edu/hub/transcript-legend.html*

Contact Information

Please direct all questions to The University Registrar's Office:

Carnegie Mellon University	Phone: 412-268-4138
University Registrar's Office	Fax:412-268-6651
5000 Forbes Avenue	Email:uro-transcrtipts@andrew.cmu.edu
Pittsburgh, PA 15213-3890	Website: http://www.cmu.edu/hub

Family Education Rights and Privacy Act (FERPA) of 1974 as Amended

This record is released on the condition that the student information contained herein will not be transferred to a third party without the written consent of the student.

Accreditation

Carnegie Mellon University is an accredited member of the Middle States Commission on Higher Education, 3624 Market St., Philadelphia, PA 19104.

College and department accreditations include:

- Accreditation Board for Engineering Technology (ABET) – Chemical Engineering, Civil Engineering, Electrical and Computer Engineering, Engineering and Public Policy, Mechanical Engineering, and Materials Science and Engineering
- National Architectural Accrediting Board (NAAB) – Architecture
- National Association of Schools of Art and Design (NASAD) – Art and Design
- National Association of Schools of Music (NASM) – Music
- American Assembly of Collegiate Schools of Business (AACSB) and Middle Atlantic Association of College of Business Administration (MAACBA) – David A. Tepper School of Business; American Chemical Society (ACS) – Chemistry
- National Association of Schools of Public Affairs and Administration (NAPAA) – H. John Heinz III College.

Language of Instruction

All Carnegie Mellon University courses are taught in English.

Teaching Location

This transcript reflects all Carnegie Mellon coursework, independent of campus or teaching location.

Calendar

Carnegie Mellon observes the semester system with two six-week summer sessions. Courses may be taught in other shortened sessions.

Course Numbers

Each Carnegie Mellon University course number begins with a two-digit prefix which designates the department offering the course (76-xxx courses are offered by the Department of English, etc.) Although each department maintains its own course numbering practices, typically the first digit after the prefix indicates the class level: xx-1xx courses are freshmen-level, xx-2xx courses are sophomore level, etc. xx-6xx courses may be either undergraduate senior-level or graduate-level, depending on the department. Xx-7xx courses and higher are graduate-level.

Degree Requirements

Degrees are awarded upon satisfactory completion of residence requirements, all requirements in the approved curriculum(a) and by recommendation for degree(s) by the faculty of the appropriate college(s).

Units of Work vs. Credit Hours

Three units equal one semester hour of credit.

Quality Point Average (QPA) Calculations

Carnegie Mellon University defines a quality point as a point value times units for a given course. QPAs are calculated according to the following formula:

Semester QPA: quality points divided by factorable units.
Cumulative QPA: total quality points divided by total factorable units.

Undergraduate courses may not be factorable into the QPA for graduate students, depending on the student's college.

Units Carried vs. Passed vs. Factored

Units Carried refers to the total number of units for which the student is registered.

Units Passed is the total of all units that have a passing grade. Grades of 'R','N', and 'I' are not included in units passed.

Units Factored is the total of all units factored into QPA. See Grading Standards for grades not factorable into QPA. Undergraduate courses are not factorable into QPA for graduate students.

Cumulative QPA

A separate cumulative QPA is calculated for undergraduate and graduate records. If a student attends for a combination of undergraduate, graduate studies, and/or non-degree studies, a cumulative QPA may not be calculated.

Cumulative QPA may not be available on transcripts for semesters prior to 1989.

4.0 Grading Standards

Undergraduate students (and graduate students in CFA, CIT, CMU, MCS, and SCS who entered before Fall 1995):

Grade	*Point Value*	*Description*
A	4.0	Excellent
B	3.0	Good
C	2.0	Satisfactory
D	1.0	Passing
R	0.0	Failure
X	0.0	Conditional Failure
S	non-factorable	Satisfactory
P	non-factorable	Passing
N	non-factorable	Failure in Pass/Fail Course
O	non-factorable	Audit
W	non-factorable	Withdrawal
I	non-factorable	Incomplete
AD	non-factorable	Credit by examination
TR	non-factorable	Transfer credit

4+ Grading Standards / 9.0 Grading Standards

The 4+ grading scale is applicable to graduate students who entered in and after Fall 1995. The 9.0 grading scale is applicable only to certain graduate students who entered before Fall 1995: Students in the Graduate School of Industrial Administration (GSIA, now Tepper School of Business (TSB)), the School of Urban and Public Affairs (SUPA, now Heinz College (HC)), and graduate students who were admitted to the College of Humanities and Social Sciences (H&SS, now Dietrich College (DC)) after August 1986.

Grade	*Point Value (4+)*	*Point Value (9.0)*	*Description*
A+ *	4.33	9	
A	4.00	8	
A-	3.66	7	
B+	3.33	6	
B	3.00	5	
B-	2.67	4	
C+	2.33	3	
C	2.00	2	
C-	1.67	1	
D+ *	1.33	-	
D *	1.0	0	
R	0.0	0	Failure
X	0.0	-	Conditional Failure
S	non-factorable	non-factorable	Satisfactory
P	non-factorable	non-factorable	Passing
N	non-factorable	-	Failure in Pass/Fail Course
O	non-factorable	-	Audit
W	non-factorable	-	Withdrawal
I	non-factorable	-	Incomplete
AD	non-factorable	-	Credit by examination
TR	non-factorable	-	Transfer credit

* DC and CIT graduate students are not permitted to receive an A+. TSB and HC graduate students do not receive D or D+ grades.

Grading Scale & QPA Conversion

We are unable to provide conversion to other grading scales, such as percentage. A guide to our grading scale is provided on the back of all official transcripts. While we do provide overall semester QPA and cumulative QPA, we are unable to provide an "in-major" QPA (i.e. QPA for courses only in your major).

Physical Education Courses

Physical Education Courses are considered Units Passed in the student's overall semester and cumulative QPA; they are not considered Units Factorable and are not used in calculating the student's overall semester QPA, rank in class or for academic actions.

Statement Of Assurance

Carnegie Mellon University does not discriminate in admission, employment, or administration of its programs or activities on the basis of race, color, national origin, sex, handicap or disability, age, sexual orientation, gender identity, religion, creed, ancestry, belief, veteran status, or genetic information. Furthermore, Carnegie Mellon University does not discriminate and is required not to discriminate in violation of federal, state, or local laws or executive orders.

Inquiries concerning the application of and compliance with this statement should be directed to the Vice President for Campus Affairs, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213, telephone 412-268-2056.

Obtain general information about Carnegie Mellon University by calling 412-268-2000.

Yingjing Lu

yingjing_lu@163.com (412) 519-4745

EDUCATION

Carnegie Mellon University May 2019
BS in Information Systems Major QPA 3.63/4.00
Minor in Machine Learning, Computational Finance Overall QPA 3.47/4.00

PUBLICATIONS

The Level Weighted Structural Similarity Loss: A Step away from Mean Square Error
A novel local structural similarity loss function for training generative models. The model addresses traditional SSIM function family's insensitivity to luminance and contrast issue

- Yingjing Lu.
- Published at AAAI 2019 Student Track

Mining Connections between Domains through Latent Space Mapping
A GAN variant that can map latent spaces between different domains in a semi-supervised manner. The model can be used for cross domain generation, classification in a label efficient manner

- Yingjing Lu.
- Published at IEEE ICDM 2018 HDM Workshop; Oral Presentation

Cross Domain Image Generation through Latent Space Exploration with Adversarial Loss
A framework that can generate images across different domains utilizing VAE and GAN in a semi-supervised manner

- Yingjing Lu.
- *arXiv preprint* arXiv:1805.10130 (2018)

OTHER RESEARCH PROJECTS

Music Onset Detection and Retrieval through Multimodal CNN
A model that treats onset location as a multi-modal learning problem and utilizes convolutional neural networks to extract spectrogram features to locate the onset points.

- Yingjing Lu, Roger B. Dannenberg.
- Machine Learning Senior Thesis Project

Anomaly: Online Anomaly Detection with Server Logs
A novel paired ELM model that can learn to detect server anomalies and adapt concept drifts of attacks in an online manner

- Yingjing Lu, C.F. Larry Heimann.
- Information Systems Senior Thesis Project

Combating Mode Collapse: Bayesian Wasserstein Learning for GAN
A novel Wasserstein Bayesian GAN. Unified Bayesian GAN, and MMD GAN with SGHMC sampling method with artificial frictions.

- Boxiang Lyu, Yingjing Lu, Xiang Si.
- Probabilistic Graphical Model Research Project

WORK EXPERIENCES

Goldman Sachs New York – Software Engineering Intern June 2018 – Aug 2018

Worked in core development team to develop tools for engineers across the firm

- Implemented a node.js server that hosted core functionalities with optimized cache structures
- Successfully defended in production review and pushed into production

ZpuData - Software Engineer & Manager May – Aug 2016

Worked in team to design and developed BETA version of a travel recommendation system

- Modeled MySQL database and Python data analytics framework APIs
- Developed application models and data visualization dashboards using Ruby on Rails
- Created dev docs and trained developers to continue from existing modules

Carnegie Mellon University - Software Developer May 2017 – Sep 2017

Collaborated with marketing and housing team to build applications and data analytics tools

- Wrote a data analytics framework with Python, ColdFusion, and JavaScript
- Optimized SQL database structures and queries to reduce server load

HONORS & ACHIEVEMENTS

Dean's List, High Honor 2018

Dean's List 2016

Midwest Trading Competition – Case 2 First Place 2018

MIT Trading Competition – Top 20% 2017

Hack Princeton – Best First Hack 2016

RESEARCH INTERESTS

Computer Vision, Generative Models, Perception, Signal Processing, Distributed Machine Learning

DEVELOPMENT SKILLS:

Python; SQL; C, C++; HTML, CSS, JavaScript; MongoDB; Ruby; Java; C#;

LIBRARIES & TOOLKITS

TensorFlow; PyTorch; Torch; Django; Rails; Node.js; Scipy; Cython

Yingjing Lu’s List of Publications

Lu, Yingjing. “The Level Weighted Structural Similarity Loss: A Step away from Mean Square Error.” *Association for Advances in Artificial Intelligence 2019* Student Track

Lu, Yingjing. “Mining Connections between Domains through Latent Space Mapping.” *IEEE International Conference on Data Mining 2018* HDM Workshop; Oral Presentation

Lu, Yingjing. “Cross Domain Image Generation through Latent Space Exploration with Adversarial Loss.” *arXiv preprint* arXiv:1805.10130 (2018)

TOEFL

TOEFL® (Test of English as a Foreign Language™)
Internet-based Test (TOEFL iBT™)
Examinee Score Report

ETS Security Guard

Name: LU, YINGJING
Last (Family/Surname) Name, First (Given) Name Middle Name

Email: wuyi5105131@163.com

Gender: M

Date of Birth: 19 Jan 1997

Registration Number: 0000 0000 2137 1583

Test Date: 12 Jul 2014

55C

LU, YINGJING
100089 北京 车道沟南里30号楼5门502 卢应京 收
beijing, Beijing 100089
China

TOEFL Scaled Scores

Reading	29
Listening	27
Speaking	26
Writing	28
Total Score	**110**

30

Country of Birth: China

Native Language: CHINESE

Sponsor Code: **Test Center Code:** STN80055C

Test Center Country: China

Inst. Code	**Dept. Code**

Security Identification

ID Type: National ID **ID No.:** 110108199701193136 **Issuing Country:** China

Reading Skills	Level	Your Performance
Reading	High	Test takers who receive a score at the **HIGH** level, as you did, typically understand academic texts in English that require a wide range of reading abilities regardless of the difficulty of the texts. Test takers who score at the **HIGH** level, typically • have a very good command of academic vocabulary and grammatical structure; • can understand and connect information, make appropriate inferences, and synthesize ideas, even when the text is conceptually dense and the language is complex; • can recognize the expository organization of a text and the role that specific information serves within the larger text, even when the text is conceptually dense; and • can abstract major ideas from a text, even when the text is conceptually dense and contains complex language.

Listening Skills	Level	Your Performance
Listening	High	Test takers who receive a score at the **HIGH level**, as you did, typically understand conversations and lectures in English that present a wide range of listening demands. These demands can include difficult vocabulary (uncommon terms, or colloquial or figurative language), complex grammatical structures, abstract or complex ideas, and/or making sense of unexpected or seemingly contradictory information. When listening to lectures and conversations like these, test takers at the **HIGH** level typically can • understand main ideas and important details, whether they are stated or implied; • distinguish more important ideas from less important ones; • understand how information is being used (for example, to provide evidence for a claim or describe a step in a complex process); • recognize how pieces of information are connected (for example, in a cause-and-effect relationship); • understand many different ways that speakers use language for purposes other than to give information (for example, to emphasize a point, express agreement or disagreement, or convey intentions indirectly); and • synthesize information, even when it is not presented in sequence, and make correct inferences on the basis of that information.

0673700018354957 1

ETSPA01 92673-49264 • SR912R1600 • Printed in U.S.A. I.N. 759004

GRE

TEST TAKER SCORE REPORT

Note: This report is not valid for transmission of scores to an institution.

Yingjing Lu

Address: 5032 Forbes Ave, SMC 4457, Pittsburgh, PA, 15289 United States

Email: yingjing_lu@163.com
Phone: 1-4125194745
Date of Birth: January 19, 1997
Social Security Number (Last Four Digits):
Gender: Male
Intended Graduate Major: Information Sciences/Studies (0404)

Most Recent Test Date: August 4, 2018

Registration Number: 3432462
Print Date: November 24, 2018

Your Scores for the General Test Taken on August 4, 2018

Verbal Reasoning	Quantitative Reasoning	Analytical Writing
Your Scaled Score: 157	Your Scaled Score: 170	Your Score: 4.0
130–170	130–170	0–6
76th Percentile	96th Percentile	59th Percentile

Your Test Score History

General Test Scores

	Verbal Reasoning		Quantitative Reasoning		Analytical Writing	
Test Date	**Scaled Score**	**Percentile**	**Scaled Score**	**Percentile**	**Score**	**Percentile**
August 4, 2018	157	76	170	96	4.0	59
July 7, 2018	156	73	166	90	4.0	59
October 28, 2017	Not Available	---	Not Available	---	Not Available	---

• **Not Available** - Scores are currently not available. Please allow 10-15 days after a computer-delivered test or 5 weeks after a paper-delivered test for your scores to be reported.

Subject Test Scores

You do not have reportable test scores at this time.

Your Score Recipient(s)

Undergraduate Institution

Report Date	Institution (Code)	Department (Code)	Test Title	Test Date

Yingjing Lu

Most Recent Test Date: August 4, 2018

Date of Birth: January 19, 1997

Registration Number: 3432462
Print Date: November 24, 2018

Designated Score Recipient(s)

Report Date	Score Recipient (Code)	Department (Code)	Test Title	Test Date
November 21, 2018	Carnegie Mellon University (2074)	COMPUTER SCIENCE (0402)	General Test	August 4, 2018
November 21, 2018	STANFORD UNIVERSITY (4704)	COMPUTER SCIENCE (0402)	General Test	August 4, 2018
August 15, 2018	U CA SAN DIEGO (4836)	COMPUTER SCIENCE (0402)	General Test	August 4, 2018
August 15, 2018	U CA SANTA BARBARA (4835)	COMPUTER SCIENCE (0402)	General Test	August 4, 2018
August 15, 2018	U PENNSYLVANIA GRAD EDUCATN (2943)	COMPUTER SCIENCE (0402)	General Test	August 4, 2018
August 15, 2018	UNIV WISCONSIN MADISON (1846)	COMPUTER SCIENCE (0402)	General Test	August 4, 2018

About Your *GRE*® Score Report

Score Reporting Policies

With the *ScoreSelect*® option, you can decide which test scores to send to the institutions you designate. There are three options to choose from:

- Most Recent option – Send your scores from your most recent test administration
- All option – Send your scores from all administrations in the last five years
- Any option – Send your scores from one OR as many test administrations in the last five years (this option is not available on test day when you select up to four FREE score reports)

Scores for a test administration must be reported in their entirety. Institutions will receive score reports that show only the scores that you selected to send to them. There will be no special indication if you have taken additional GRE tests. See the *GRE® Information Bulletin* for details. The policies and procedures explained in the Bulletin for the current testing year supersede previous policies and procedures in previous bulletins.

Scores will be sent to designated score recipients approximately 10-15 days after a computer-delivered test and 5 weeks after a paper-delivered test. If your scores are not available for any reason, you will see "Not Available" in Your Test Score History.

GRE test scores are reportable according to the following policies:

- For tests taken prior to July 1, 2016, scores are reportable for five (5) years following the testing year in which you tested (July 1 – June 30). For example, scores for a test taken on May 15, 2015, are reportable through June 30, 2020. GRE scores earned prior to August 2011 are no longer reportable.
- For tests taken on or after July 1, 2016, scores are reportable for five (5) years following your test date. For example, scores for a test taken on July 3, 2016, are reportable through July 2, 2021.

Note: Score recipients will only receive scores from test administrations that you have selected to send to them.

Percentile Rank (% Below)

A percentile rank for a test score indicates the percentage of test takers who took that test and received a lower score. Regardless of when the reported scores were earned, the percentile ranks for General Test and Subject Test scores are based on the scores of all test takers who tested within the most recent three-year period.

Yingjing Lu

Most Recent Test Date: August 4, 2018

Date of Birth: January 19, 1997

Registration Number: 3432462
Print Date: November 24, 2018

Retaking a GRE Test

You can take the *GRE*® **General Test** *once every 21 days*, up to *five times* within any continuous rolling 12-month period (365 days). This applies even if you canceled your scores on a test taken previously. You can take the paper-delivered GRE General Test and *GRE*® **Subject** Tests as often as they are offered.

Note: This policy will be enforced even if a violation is not immediately identified (e.g., inconsistent registration information) and test scores have been reported. In such cases, the invalid scores will be canceled and score recipients will be notified of the cancellation. Test fees will be forfeited.

For More Information

For information about interpreting your scores, see *Interpreting Your GRE Scores* at www.ets.org/gre/understand.

For detailed information about your performance on the Verbal Reasoning and Quantitative Reasoning sections of the computer-delivered GRE General Test, access the free GRE Diagnostic Service from your ETS account. This service includes a description of the types of questions you answered right and wrong, the difficulty level of each question, and the time spent on each question. This service is available approximately 15 days after your test administration and for six months following your test administration.

If you have any questions concerning your score report, email GRE Services at **gre-info@ets.org** or call 1-609-771-7670 or 1-866-473-4373 (toll free for test takers in the U.S., U.S. Territories and Canada) between 8 a.m. and 7:45 p.m. (New York Time).

Mining Connections between Domains through Latent Space Mapping

Yingjing Lu
Carnegie Mellon University
Pittsburgh, USA
yingjinl@andrew.cmu.edu

***Abstract*—Exploring ways to connect data is crucial to building knowledge graphs to associate data from different domains together. Humans, for example, can learn to associate flour with bread because bread is made of flour so that they can recall information of flour given a piece of bread even though bread and flour have few common features. In data mining, this ability can be translated to the way to connect images, texts, audios from different classes or domains together. Most works so far assume shared feature representations between domains we want to connect together. Another limitation yet to be improved is that for each defined mapping scheme, we often have to train a new model end-to-end among all sample data, which is often expensive. In this work, we present a model that aims to simultaneously address the two limitations. We use unconditionally trained Variational Autoencoders(VAEs) to project high dimensional data into the latent space and present a novel generative model that transfer latent representation of data from one domain to another by any custom schema. The model makes no assumption on any shared representation among different domains. The VAEs that encodes entire datasets, being the largest training overhead in this model, can be reused to support any new mapping schema without any retraining.**

***Keywords*-latent space, cross domain, generative models**

I. Introduction

One advantage humans have over current computer learning algorithms is that humans are able to associate knowledge and representations from different domains together easily without any requirement on any common features those domains have in common. For example, humans can learn to associate images with number ”2” with images with text ”Two” even though they may not have explicit visual features in common. In this case, associations among domains are defined by human and are subjected to change. This ability is crucial in modern knowledge learning algorithms to adopt semantics not obvious to be naturally extracted by prevalent learning frameworks such as convolutional neural networks. The most common and efficient way to build association is to explore the relationships of samples in its latent representations since latent representations can have lower dimensionality and contain distilled information that cannot be explicitly defined such as style of writing. Some previous works have been done to address mapping between domains through latent spaces but have some limitations to be improved. One problem is that many work assume there are shared features among different domains or the latent spaces of those domains has similar distributions so that they can be mapped by finding the best way to make the distribution overlap with each other. Another problem is that to map two domains, almost all the models require to train end to end on at least one of the domains. In other words, to connect domain B with domain A under a new mapping scheme, most models can reuse a pre-trained part for domain A but require training on domain B end to end to accommodate the new A-B mapping scheme or on both parts of domain B and A have to be retrained end-to-end to accommodate new mapping scheme. Given large sample counts and high dimensionality of the data, computation is usually expensive. There are currently two directions trying to address the above two limitations:

Domain representation transfer learning is a methodology presented by [1] that utilizes a pretrained classification model in one domain as a base to train a classification model in a new domain by mapping manifold of classes in new domain to classes in old domain. This direction is similar to ours in that it does not make assumption on any commonalities between classes of the two domains but still suffers from training the new domain classifier end-to-end if mapping scheme that connects the two domains gets changed.

Post-hoc latent representation learning refers to the process to learning the latent manifold from an existing model.In this way, the model such as an unconditionally trained VAE can be used to generate images by arbitrarily defined conditionals without retraining. In the paper [2], the authors show effective generation of faces with specific conditioned features from a VAE that is trained to generate faces unconditionally. However, they only show model working on conditionals in the form of one-hot vectors. In this case, the model performance becomes dependent on specific feature engineering. In comparison, our model automatically extract latent conditional representations and does not require any feature engineering.

We base our work on the two directions above and present a model that can both learn to map data between arbitrary domain without assuming shared features and at the same time, does not require end-to-end retraining every time we want to learn a new mapping scheme. Further, we also design our model to be relatively robust to learning new mapping through only a small portion of data to make the learning of new mapping scheme more efficient. Our major contributions are summarized as follows:

- To the best of our knowledge, we are the first to combine domain representation transfer problems with post-hoc latent representation learning to formulate a data mining problem that explores representations of data from one domain to the data on another through arbitrarily defined mapping scheme in latent spaces of those data.
- We propose a model that can learn to connect data from one domain to another that does not require to retrain entire model every time.
- We use frequently used datasets to formulate multiple cross domain learning setups and evaluate our model's performance and robustness on those setups and present performance, limitations and future directions.

II. Background

Our model utilizes two trending generative models to perform mapping. The Variational Autoencoder serves as data encoder that converts data into its latent representation. It is also able to reconstruct data from latent encodes from either its home domain or from codes mapped from another domain. Generative Adversarial Networks is the other model we use to transform latent representation of data from one domain to another. The next two session will be dedicated to introduce background of the two generative models.

A. Variational Autoencoder

Variational Autoencoder(VAE) is a generative model that tries to approximate the true distribution $p(\mathbf{x})$ through N given data points $\{\mathbf{x}\}_{l=0}^{N}$ and latent random variables $\{\mathbf{z}\}_{l=0}^{N}$ that are sampled from a pre-defined prior simple distribution $p(\mathbf{z})$, most often Gaussian. This can be expressed as a marginalizing process as:

$$p(\mathbf{x}) = \int_{\mathbf{z}\sim p(\mathbf{z})} p(\mathbf{x} \mid \mathbf{z})p(\mathbf{z})d\mathbf{z} \tag{1}$$

In [3], the authors point out that $p(\mathbf{x}|\mathbf{z})$ is often intractable and $p(\mathbf{z})$ is, in most cases, often not the true latent probability space that the sample space can be represented in. Thus the authors propose to find an encoding model that can approximate the distribution of the sample space. The model consists of two parts, encoder and decoder. The encoder aims to approximate the intractable distribution $p(\mathbf{x}|\mathbf{z})$ with distribution $q_\theta(\mathbf{z}|\mathbf{x})$ with a set of parameters denoted as ϕ. This is the encoding process that 'encodes' given data samples into the latent probability space $q_\phi(\mathbf{z})$ and represents the encoded sample as a random variable $\mathbf{z}$. On the other hand, the decoder aims to recover the true distribution by approximating $p_\theta(\mathbf{x}|\mathbf{z})$ with a set of parameters denoted as θ. In other words, the decoder aims to reconstruct a sample in the probability distribution of training data given a latent random variable $\mathbf{z}$. Both encoder and decode are implemented as neural networks.The objective function of the VAE is formulated by minimizing the KL divergence between the approximated $q_\phi(\mathbf{z}|\mathbf{x})$ and $p(\mathbf{z}|\mathbf{x})$ which is shown in [3] by maximizing the estimated lower bound(ELBO):

$$\begin{aligned} ELBO &= \mathbb{E}_q[\log p_\phi(\mathbf{x}|\mathbf{z})] - KL_q[q_\theta(\mathbf{z}|\mathbf{x})||p(\mathbf{z})] \\ &\leq \log p(\mathbf{x}) \end{aligned} \tag{2}$$

Here $p(\mathbf{x})$ represent true distribution of the domain dataset. KL represents the Kullback-Leibler divergence. Foot note q represents $q_\theta(\mathbf{z}|\mathbf{x})$. Encoder and decoder functions are parametrized by ϕ and θ respectively.

B. Generative Adversarial Networks

Generative Adversarial Network[4](GAN) is another class of generative model that learns to map a random variable $\mathbf{z}$ from a simple prior distribution $p(\mathbf{z})$ directly to the given distribution $p(\mathbf{x})$ of the dataset. The model consists of a generator G that takes in a latent noise sample $\mathbf{z}$ from prior distribution and generate a sample $G(\mathbf{z})$ that "looks like" a sample from the actual sample distribution. The other part, discriminator D takes in a sample and output 1 if the sample input comes from true sample space, and 0 if the sample is generated by G. The loss function can be expressed as playing a minimax game between G and D in that G tries to "fool" D and D tries to prevent itself from being "fooled":

$$\begin{aligned} \min_G \max_D \mathcal{L}_{GAN}(D,G) &= E_{\mathbf{x}\sim p(\mathbf{x})}[\log(D(\mathbf{x}))] \\ &+ E_{\mathbf{z}\sim p(\mathbf{z})}[log(1 - D(G(\mathbf{z})))] \end{aligned} \tag{3}$$

In the conditional setting, the generation with conditional variable $\mathbf{y}$ can be expressed as a joint expression of random variable and the conditional [5]:

$$\begin{aligned} \min_G \max_D \mathcal{L}_{GAN}(D,G) &= \mathbb{E}_{\mathbf{x}\sim p_(\mathbf{x})}[\log(D(\mathbf{x},\mathbf{y}))] \\ &+ \mathbb{E}_{\mathbf{z}\sim p(\mathbf{z})}[\log(1 - D(G(\mathbf{z},\mathbf{y}),\mathbf{y}))] \end{aligned} \tag{4}$$

III. Proposed Model

We will use foot indices such as $\mathcal{D}_1$ to define variables from corresponding domain 1 and head notes such as $x_1^{(i)}$ to define ith sample data within domain 1. We also use bold alphabets to express random variables such as $\mathbf{x}$. Our proposed framework lies on conditionally generating images from two different domains $\mathcal{D}_1$ with probability distribution $p_1(\mathbf{x}_1)$, $\mathcal{D}_2$ with probability distribution $p_2(\mathbf{x}_2)$ for images from domain 1 and domain 2 respectively. $\mathbf{x}_1$ and $\mathbf{x}_2$ here being i.i.d within each domain. We define $\gamma_1^i \rightarrow \gamma_2^j$ to be the user defined arbitrary class mapping scheme that maps class i from domain 1 to class j from domain 2. One example would be to map images of 1s in domain 1 to associate with images of 2s in domain 2(Images of 1 and 2 visually have no explicit features in common). In training the user

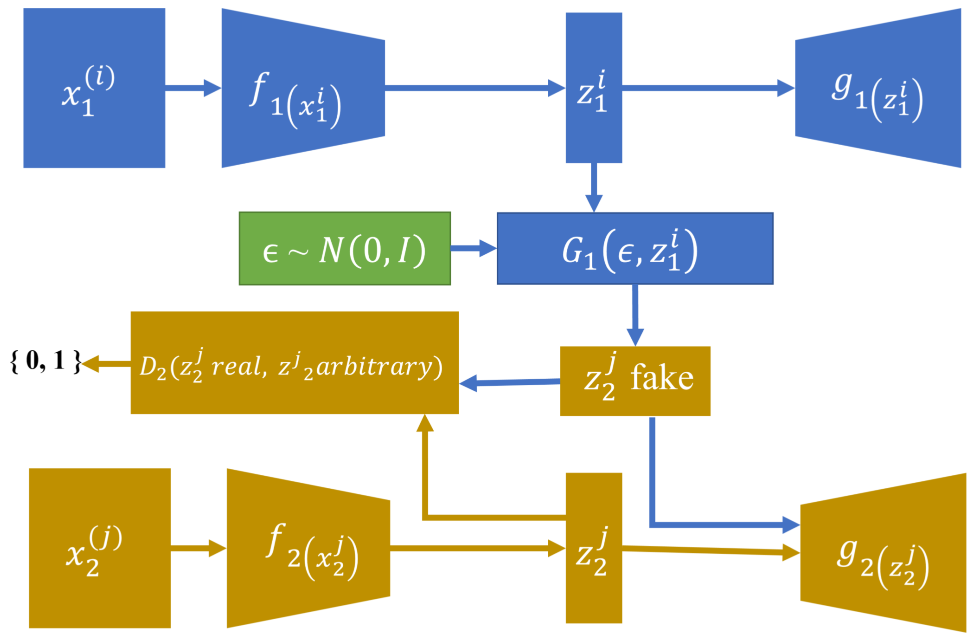

Figure 1. The training process of the generator that transforms latent $z_1^{(i)}$ in class i from domain 1 to $z_2^{(j)}$ in class j from domain 2. For clarity, we only illustrate learning from one direction that is from domain 1 to domain 2. G_1 is the generator that takes in latent embedding from domain 1 $z_1^{(i)}$ as conditional and ϵ sampled from an arbitrary prior to produce an sample $z_2^{(j)}$ $fake$ that maps to a possible probability space that can generate samples from class j in domain 2 . We in this case use simple Gaussian noise as prior. The discriminator for domain 2 $D_2(z_2^{(j)}\ real, z_2^{(j)}\ arbitrary)$ takes in real embedding from class j $z_2^{(j)}$ as conditional to distinguish whether the embedding $z_2^{(j)}$ $arbitrary$ is generated by true f_2 or from G_1. Note that the framework is bi-directional. To transform from domain 2 to domain 1 we just need a reverse set of generator and discriminator that transform encoding from domain 2 to 1.

specifies no other information other than the pair of image the model should learn to associate together. We believe this can approximate the actual learning process of humans when they learn to associate different images or seen objects together.

Without loss of generality, we define two VAEs V_1 and V_2 with encoding function $z_1 = f_1(x_1)$, $z_2 = f_2(x_2)$ as encoder functions for each VAE and $g_1(z_1)$ and $g_2(z_2)$ as the corresponding decoder network that are trained unconditionally on the two image domains respectively. We use the word unconditionally to indicate that the two VAEs are both trained with standard unsupervised fashion without any additional conditionals involved in training.Thus the total loss function we used in training our VAEs is:

$$\mathcal{L}(\theta, \phi, x^{(i)}) = \frac{1}{N} \sum_{i=1}^{N} \lambda_1 \mathbb{E}_q[\log p(x^{(i)}|z^{(i)})] - \lambda_2 KL_q \quad (5)$$

Here $E_q[\log p(x^{(i)}|z^{(i)})]$ denotes reconstruction cost which is implemented as pixel-wise squared difference and KL_q is a shorthand for $KL_q[p(\mathbf{x}, \mathbf{z})||q(\mathbf{z})]$, described above. $\{\lambda_1, \lambda_2\}$ are the two hyperparameters that are used to balance the reconstruction cost and the KL divergence. The two hyperparameters are shared among the two VAEs throughout our experiments. The effect of hyperparameters is illustrated in Fig. 3.

With the given cost functions for training we further assume that the two VAEs are well trained in that they can unconditionally reconstruct images with high fidelity given the images within their respective domain. Our training and implementation details are provided in the experiment section.

To learn a given domain conditional $\gamma_1^i \rightarrow \gamma_2^j$ we will sample images $x_1^{(i)}$ for arbitrary sample of class i in $\mathcal{D}_1$ and $x_2^{(j)}$ for arbitrary sample of class j from $\mathcal{D}_2$. We adopt the notion of generative adversarial network(GANs) and use a generator $G(\epsilon, z^{(i)})$ to transform embedding learned from one VAE to the corresponding embedding of another. The generator takes in a noise sampled from a simple prior distribution and the conditional embedding $z^{(i)}$. We will follow our presentation structure and use the setting of transforming from domain of V_1 to domain of V_2 form class i to class j. Thus the generator on domain 1 side generates encoding $z_2^{(j)}\ fake = G_1(\epsilon, z_1^{(i)})$. During training, the result from G_1 is passed to the discriminator on the domain 2 side $D_2(z_2^{(j)}, z_2\ arbitrary)$. The discriminator from domain 2 takes true encoding from $f_2(x_2^j)$ as conditional and to determine whether $z_2\ arbitrary$ is in the true embedded subspace of class j in domain 2. To make the discriminators stronger within this model we expand our loss function from the 'traditional' GAN loss. For clarity, we introduce $z^{'}$ for

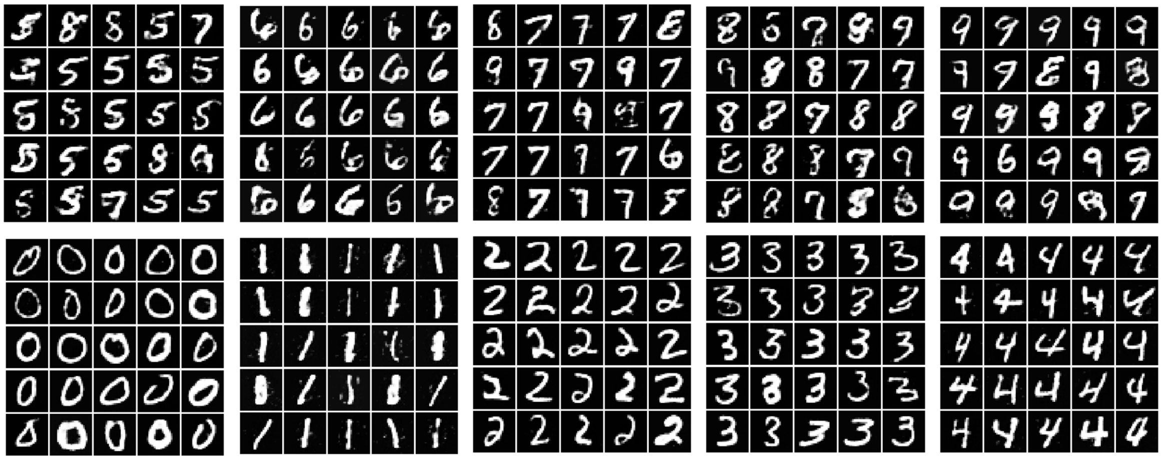

Figure 2. MNIST to MNIST generation with the conditional pairs described above. Bottom are the reconstructions from source encoding. Top are the images generated with transformed source encoding to the target domain.

the arbitrary embedding z input to the discriminator(can be either true embedding or false ones from generator output), and z for the true embedding generated by the encoder network. We shorthand notation $\mathcal{L}_{c=1}(z^{'}, z) \equiv -\log(D(z^{'}, z))$ for the loss of the true encodings under true conditionals, $\mathcal{L}_{c=1}(z, z^{'}) \equiv -(1-\log(D(z, z^{'})))$ as the discriminator loss for classifying a false embedding under true conditional, and finally an additional term $\mathcal{L}_{c=0}(z, \epsilon) \equiv -(1-\log(D(z, \epsilon))$ as the false classification for input a random noise from a simple distribution under true conditional. The third term here aims to strengthen the discriminator that under the conditional since the embedding generated from G should not be a random noise vector that seems to comply with the pattern of the distribution. The full loss function of the discriminator is then:

$$\begin{aligned}\mathcal{L}_D = \mathbb{E}_{z^{'} \sim q(z|x)}\left[\mathcal{L}_{c=1}(z, z^{'})\right] + \mathbb{E}_{z \sim G(z,\epsilon)}\left[\mathcal{L}_{c=0}(z, z^{'})\right] \\ + \mathbb{E}_{\epsilon \sim p(z)}\left[\mathcal{L}_{c=0}(z, \epsilon)\right]\end{aligned} \tag{6}$$

Opposite to the discriminator which restricts the generator from generating embedding simply from random noise, we introduce a regularization term that is inspired by the log regularization term proposed by [2] as $\frac{1}{n}||\epsilon - G(\epsilon, z)||_2^2$ here z represents the input embedding to the generator to be transformed. Intuitively, as the generator shifts the simple noise to the mapped distribution, it would usually maximize the distance between the embedding generated and the original noise. With this term added in minimizing the loss function will create a force that "pulls back" G from moving from ϵ too far, thus encourage variety of the embedding generated. Adding the regularization term resulting in our generator loss:

$$\mathcal{L}_G = \mathbb{E}_{z \sim p(\mathbf{z}), \epsilon \sim p(z)}\left[\mathcal{L}_{c=1} G(z, \epsilon) + \frac{\lambda_{reg}}{n}||\epsilon - G(\epsilon, z)||_2^2\right] \tag{7}$$

During training, we first train two VAEs until convergence, then we train two pairs of generator and discriminator alternatively until convergence. During sampling, we sample from desired class i in domain 1 and feed in V_1 then VAE will encode the sample to embedding $z_1^{(i)}$ and transformed by generator with a noise $G_1(\epsilon, z_1^{(i)})$, this new transformed embedding will be passed to decoder of V_2 to generate actual sample $g_2(G_1(\epsilon, z_1^{(i)}))$ or perform classification tasks that based on mapping.

IV. Experiments

A. Dataset and Setups

We performed our experiments mainly on MNIST[6] and SVHN[7] datasets. MNIST dataset contains 28×28 hand written digits from 0 to 9 in grey scale with approximately 60000 training samples and 10000 testing samples. SVHN dataset contains digit photos captured in street cropped to 32×32 in RGB. Training set contains more than 73000 images and testing set contains more than 26000 images for digits from 1 to 9. Fashion MNIST is a dataset introduced in [8] that contains different shoes and clothing in 10 different classes in 28×28 grey scale with more than 60000 in training set and more than 10000 in testing set.

For **MNIST** $\leftrightarrow$ **MNIST** mapping, we split the training set and testing set with number as their classes. To adhere to our assumption and avoid feature matching, we assign domain 1 contain digits $\{0, 1, 2, 3, 4\}$ and domain 2 contains digits $\{5, 6, 7, 8, 9\}$. The conditional generation law is defined to generate a certain class of digit given a class of digit. For conditional generation from domain 1 to domain 2 we define $\{0 \leftrightarrow 5, 1 \leftrightarrow 6, 2 \leftrightarrow 7, 3 \leftrightarrow 8, 4 \leftrightarrow 9\}$ going from left to right and conditional generation from domain 2 to domain 1 is the opposite direction.

For **MNIST** $\leftrightarrow$ **Fashion MNIST** mapping, we randomly select to use a transfer scheme of $\{0 \leftrightarrow T-shirt, 1 \leftrightarrow$

Table I
PRECISION OF IMAGE GENERATION AND CLASSIFICATION

Dataset	Size of training set	Precision(Image Generation)	Precision(Classification)
MNIST → MNIST	500	0.687	0.635
MNIST → MNIST	1000	0.832	0.713
MNIST → MNIST	2000	0.802	0.766
MNIST → MNIST	Full set	0.901	0.866
Fashion MNIST → MNIST	500	0.423	0.431
Fashion MNIST → MNIST	1000	0.378	0.398
Fashion MNIST → MNIST	2000	0.435	0.413
Fashion MNIST → MNIST	Full set	0.618	0.581
SVHN → MNIST	500	0.478	0.453
SVHN → MNIST	1000	0.505	0.478
SVHN → MNIST	2000	0.508	0.475
SVHN → MNIST	Full set	0.603	0.577

$Trouser, 2 \leftrightarrow Pullover, 3 \leftrightarrow Dress, 4 \leftrightarrow Coat, 5 \leftrightarrow Sandal, 6 \leftrightarrow Shirt, 7 \leftrightarrow Sneaker, 8 \leftrightarrow Bag, 9 \leftrightarrow AnkleBoot\}$ as our pair to test during the qualitative evaluation pair.

For **SVHN** $\leftrightarrow$ **MNIST** mapping we use similar scheme to regard digits as different classes. Since SVHN does not contain digit 0, we drop the pair 0-5 and only use $\{1 \leftrightarrow 6, 2 \leftrightarrow 7, 3 \leftrightarrow 8, 4 \leftrightarrow 9\}$ as our major experiment source and target class mapping.

We want to assess whether our model is able to successfully transform latent representation in one domain to the other. To evaluate the accuracy quantitatively, we evaluate our model in two folds. Firstly we train the two latent variable discriminators Ds as classifiers to evaluate whether the given latent variable transformed by our generator fits in the given class that the generator intends to transfer to and report the precision of classification in the table. In addition, since our model is based on a generative model, we also evaluate whether our model is able to transfer the encoding to make target domain be able to generate the correct class of image. To evaluate whether the resuting images belong to the target domain and correct class, we train a convolutional neural network on the full train set of MNIST that serve as a classifier for evaluating whether our image generated belongs to the right class if MNIST is the target domain. Our convolutional neural network achieves 99.2% classification accuracy on MNIST test set so that we are confident that our classifier would be a good fit to evaluate whether our model can generate samples in the right class in MNIST when MNIST is set as the target domain. Since we did not train a classifier that are accurate enough to evaluate the images generated using MNIST as the source and fashion MNIST or SVHN as the target(The best classifier we did on SVHN was 72.5% and 86.7% on Fashion MNIST), we limit our results reported in those accuracy using MNIST as the target domain. For the transfer from MNIST to fashion MNIST or SVHN, we still report qualitative results through generated images.

B. Quantitative Results

We trained our VAEs and the transfer generator and discriminator pairs using the training set of the above mentioned datasets. We run the above mentioned two evaluation methods on the test set to evaluate accuracy and to generate image. For each domain transfer setup (MNIST $\leftrightarrow$ MNIST, Fashion MNIST $\leftrightarrow$ MNIST, MNIST $\leftrightarrow$ SVHN) we tested the robustness of our model through utilizing different size of the training set. We utilize 500, 1000, 2000, and Full-set which is about 5%, 10%, 20% and 100% of the dataset for MNIST and of 2.2%, 4%, 8%, 100% for SVHN.

Detailed in the table, our model is able to achieve more than 90% class transfer accuracy on full set and still remain approximately 70% to 80% of accuracy given a portion of the dataset on MNIST-MNIST transfer tasks. For Fashion MNIST and the SVHN pairing we noticed a significant drawback in terms of accuracy. We noticed that this is potentially due to the fact that the features of classes in these two datasets are not as distinct as those in MNIST. For Fashion MNIST to MNIST transfer, we noticed that the results of shirt, T-shirt, coat, dress and top are heavily mixed together. This is because these cloth types are similar in terms of explicit features (rectangular middle cloth body, two sleeves) that make them being encoded into a overlapping latent space. This also applies to SVHN dataset. Many SVHN images are either obscure or contains multiple number characters. Those confusing features make them being encoded more closely and mixed with each other in the latent space. Despite those mixing we are still able to achieve around 50% accuracy in classification tasks. Additionally, with partial training set we do not experience a significant amount of performance degrade which illustrate that fact that our model is robust and can be trained within a small portion of dataset.

C. Qualitative Results

Observing the generated images manually, we see that images generated by transferred latent encoding mostly appear legible, indicating that the generator is able to successfully

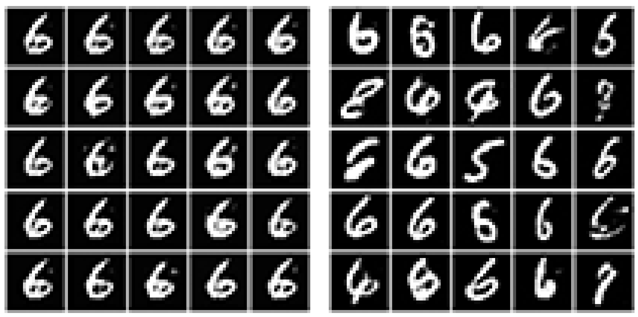

Figure 3. Result of MNIST to MNIST generated by transferred embedding with different KL coefficient λ_2. Fixing $\lambda_1 = 1$, we see that the right image with larger KL coefficient $\lambda_2 = 3$ does encourage the variety of image generation for more accurate representation of each image whereas images resulting generated resulting from lower KL coefficient $\lambda_2 = 0.3$ losses most of the varieties but has higher accuracy.

map the encoding from the source to the target domain in the right class. We do notice that, however, sometimes that the variability of the images generated appears in a relatively single form, meaning that there are a little bit of mode collapse happening since we decrease the variance of the encoding. For the fluency of the paper, we attach more samples generated at the end of the paper.

D. Implementation Details

We uses symmetric architecture for the two variational autoencoders. For the encoders, we use three layers convolutional neural networks followed by a linear layer of size 256. Z embedding is set to 50 for all experiments. For the decoder we use 1 linear layer of size 2048 followed by 4 layers of deconvolution layers. We apply RELU on all layers except the output uses tanh and batch normalization on convolutional and deconvolutional layers. To stabilize the representation further, we also experimented trade offs between having a larger KL coefficient λ_2 in VAE loss function for higher variety but lower accuracy and lower λ_2 for less variety of generation but higher representation accuracy. In Fig. 3, we show a comparison of MNIST to MNIST transfer with different λ_2 values.

For the generator and discriminator pair, we use a paired 4-layer fully connected network of size 512[9]. The conditional and the noise ϵ was simply concatenated and feed to the first layer. The generator outputs a transformed embedding as well as a gating factor resulting from sigmoid activated value of the transformed embedding. The final output transformed embedding is an interpolation between the input random noise ϵ and the transformed embedding with gating factor as the interpolation coefficient. This structure is also introduced in appendix of [2]. We adopt this architecture and believe that this interpolation can introduce more variety of output.

During experiments, we observe that increasing the λ_{reg} will encourage more diversity of image generation, which further prove that the regularization imposed is effective to encourage variety of latent space exploration, the image quality generated and classification accuracy will both be compromised if λreg is set too big (such as 0.5 or 1). We found in general 0.1 to 0.2 is the best interval for model tuning.

V. Related Works

GAN is well known for generating realistic samples through mapping latent manifolds with noise and conditionals got input into generator(s). Many prior works have illustrated successful attempts to disentangle how GAN encode features into latent space and conditional GANs(cGAN) is a particular class of GANs that are proven to control output contents through concatenating tractable conditionals with selected noise. In the conditional setting, a lot of previous work has been done to explore possibilities to control samples generated from GANs through encoded conditionals[5, 10, 11]. Here in particular, Zhu, Zhang and Pathak [12] proposed a way to use a encoder to encode latent representation of pictures in one domain as conditionals to generate images in another domain space. While achieving decent result, the model is heavily based on the assumption that images in two domains have pairwise commonalities in terms of visual features(e.g images in one domain is the segmentation of the other). On the other hand, BiGAN proposed by Donahue and Krahenbuhl.[13] does not assume pairwise commonalities between multiple domain. However, conpared to ours, their model uses weight sharing to learn to generate image from one domain to another and thus requires end to end training among all samples in two domains when people want to generate images in a new pair of domain.

Autoencoder(AE) provides a straightforward implementation of converting high dimensional data such as images or text sequence embedding into a lower dimensional latent space representation while guarantees quality of reconstruction. With this specialty, people are able to match or link data from different domains with another domain within latent space. In this way, as we mentioned, allows data to be matched with latent features without specifically engineered set of explicit features and thus allow more flexible matching. With the emergence of β-VAE [14] and its proceeding works[15, 16, 17], we are confident to assure that from empirical perspective, with more emphasis on the KL term in VAE loss function, we can 'force' VAE to explicitly decouple representations of different features in the latent space separately. Recent work done by Liu and peers.[18] has shown that VAE can perform well on matching the latent space between images from two domains. Deep matching autoencoder [19] shows its strength to match latent representation between images and text to do classification tasks. [20] bring this concept further by introducing VAE latent space inference on joint labels not seen in training data.

The emergence of GAN and related works spur the use of adversarial loss into autoencoder research to allow more flexible latent space matching objectives[21, 22, 23]. The common limitations among those models compared to ours is that they are mostly trained on a specific pair of domain and when mapping changes the model should be retrained on all samples from at least one domain.

VI. Discussion and Future Work

In this work we present a framework that can learn to map classes of images from two different domains through decoupling the latent space of the variational autoencoder and map latent domains through GAN. We show that our model can achieve satisfying results in commonly used datasets. We future show, especially in the Fashion MNIST to MNIST that our model still able to achieve decent result given that the two domains have little in common. Even though we only show the mapping between images of different domains, we believe that our model can be applied to arbitrary domains of data (such as audio, text) as long as they can be effectively encoded into latent representation. Future directions yet to explore consist of two folds: first, we observe that even though the classification accuracy is not affected by instability of training, the image generation part sometimes results in blurry and unidentifiable figures due to mapping to an untrained latent space. The conditionals we use are relatively inconsistent due to the noise of images. Finding a new way to maintain the conditionals more consistently is a direction to stabilize the training. Another direction is to explore a stronger way to disentangle the latent space so that the complex or mixed latent space of some class with similar latent representations could be separated more effectively.

Acknowledgment

The author would like to thank Sihao Yu and Siying Li for helpful discussion and encouragement during the early stage of this paper.

References

[1] Z. Luo, Y. Zou, J. Hoffman, and L. F. Fei-Fei, "Label efficient learning of transferable representations acrosss domains and tasks," in *Advances in Neural Information Processing Systems*, 2017, pp. 164–176.

[2] J. Engel, M. Hoffman, and A. Roberts, "Latent constraints: Learning to generate conditionally from unconditional generative models," *arXiv preprint arXiv:1711.05772*, 2017.

[3] D. P. Kingma and M. Welling, "Auto-encoding variational bayes," *arXiv preprint arXiv:1312.6114*, 2013.

[4] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Advances in neural information processing systems*, 2014, pp. 2672–2680.

[5] M. Mirza and S. Osindero, "Conditional generative adversarial nets," *arXiv preprint arXiv:1411.1784*, 2014.

[6] Y. LeCun and C. Cortes, "MNIST handwritten digit database," 2010. [Online]. Available: http://yann.lecun.com/exdb/mnist/

[7] Y. Netzer, T. Wang, A. Coates, A. Bissacco, B. Wu, and A. Y. Ng, "Reading digits in natural images with unsupervised feature learning," in *NIPS workshop on deep learning and unsupervised feature learning*, vol. 2011, no. 2, 2011, p. 5.

[8] H. Xiao, K. Rasul, and R. Vollgraf, "Fashion-mnist: a novel image dataset for benchmarking machine learning algorithms," *arXiv preprint arXiv:1708.07747*, 2017.

[9] I. Tolstikhin, O. Bousquet, S. Gelly, and B. Schoelkopf, "Wasserstein auto-encoders," *arXiv preprint arXiv:1711.01558*, 2017.

[10] S. Reed, Z. Akata, X. Yan, L. Logeswaran, B. Schiele, and H. Lee, "Generative adversarial text to image synthesis," *arXiv preprint arXiv:1605.05396*, 2016.

[11] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," *arXiv preprint*, 2017.

[12] J.-Y. Zhu, R. Zhang, D. Pathak, T. Darrell, A. A. Efros, O. Wang, and E. Shechtman, "Toward multimodal image-to-image translation," in *Advances in Neural Information Processing Systems 30*, I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, Eds. Curran Associates, Inc., 2017, pp. 465–476. [Online]. Available: http://papers.nips.cc/paper/6650-toward-multimodal-image-to-image-translation.pdf

[13] J. Donahue, P. Krähenbühl, and T. Darrell, "Adversarial feature learning," *arXiv preprint arXiv:1605.09782*, 2016.

[14] A. A. Alemi, I. Fischer, J. V. Dillon, and K. Murphy, "Deep variational information bottleneck," *arXiv preprint arXiv:1612.00410*, 2016.

[15] C. P. Burgess, I. Higgins, A. Pal, L. Matthey, N. Watters, G. Desjardins, and A. Lerchner, "Understanding disentangling in β-vae," *arXiv preprint arXiv:1804.03599*, 2018.

[16] H. Kim and A. Mnih, "Disentangling by factorising," *arXiv preprint arXiv:1802.05983*, 2018.

[17] T. Q. Chen, X. Li, R. Grosse, and D. Duvenaud, "Isolating sources of disentanglement in variational autoencoders," *arXiv preprint arXiv:1802.04942*, 2018.

[18] M.-Y. Liu, T. Breuel, and J. Kautz, "Unsupervised image-to-image translation networks," in *Advances in Neural Information Processing Systems*, 2017, pp. 700–708.

[19] T. Mukherjee, M. Yamada, and T. M. Hospedales, "Deep matching autoencoders," *arXiv preprint arXiv:1711.06047*, 2017.

[20] T. Inoue, S. Chaudhury, G. De Magistris, and S. Dasgupta, "Transfer learning from synthetic to real images using variational autoencoders for robotic applications," *arXiv preprint arXiv:1709.06762*, 2017.

[21] A. Makhzani, J. Shlens, N. Jaitly, I. Goodfellow, and B. Frey, "Adversarial autoencoders," *arXiv preprint arXiv:1511.05644*, 2015.

[22] V. Dumoulin, I. Belghazi, B. Poole, O. Mastropietro, A. Lamb, M. Arjovsky, and A. Courville, "Adversarially learned inference," *arXiv preprint arXiv:1606.00704*, 2016.

[23] L. Mescheder, S. Nowozin, and A. Geiger, "Adversarial variational bayes: Unifying variational autoencoders and generative adversarial networks," *arXiv preprint arXiv:1701.04722*, 2017.

Figure 4. Random sampling image generations. Row 2,4,6,8 are the reconstruction from the encodings z of the source domain. Row 1,3,5,7 are images generated from the z transformed from source z to the target domain

Northwestern | THE GRADUATE SCHOOL

Recommendation Form

The Graduate School Northwestern University Evanston, IL 60208-1113

Applicant Name: **Yingjing Lu**

Program: **Computer Science: MS**

Applicant Waived Rights*: **This applicant has waived the right to view their recommendation.**

Recommender Name: **Roger Dannenberg**

Organization Name: **Carnegie Mellon University**

Title: **Professor of Computer Science**

E-mail Address: **rbd@cs.cmu.edu**

Telephone Number: **4122683827**

Relationship to Applicant: **Research Advisor**

Certification (Date): **12-05-2018**

*"Public Law 93-380, Educational Amendments Act of 1974, grants students the right to have access to letters of recommendation in their placement files. By selecting the "Waive access" option you are waiving access to these letters."

Carnegie Mellon

Roger B. Dannenberg
Professor of Computer Science, Art & Music
Office Phone: 412-268-3827
E-mail: rbd@cs.cmu.edu
Fax: 412-268-5576
URL: http://www.cs.cmu.edu/~rbd/

25 November 2018

Dear Admissions Committee:

I am writing to recommend an outstanding undergraduate, **Yingjing Lu**, who has a general interest in applying machine learning algorithms to enhance feature recognition and to model reusability. In the past he explored different directions to utilize machine learning to analyze energy consumption, to predict links through text semantic analysis, and to optimize performance for generative image models. Yingjing had a workshop paper at ICDM that combines his interests in feature recognition and model reusability. He leverages the power of an existing feature extraction framework, Variational Autoencoder, and utilizes adversarial training to map different domain data with the extracted features. He designs the model with the idea that Autoencoders can be reused for different mappings. Yingjing conducted the research for the paper all by himself.

I am currently advising him on a new project on onset detection in musical signals. Music onset detection means finding the beginnings of notes, which is difficult because of reverberation, note overlap, and the variation in noise and onset times. In addition, training data is difficult to obtain. I have previously produced state-of-the-art results with hand-crafted features, neural nets, a single instrument (trumpet), and semi-supervised learning, but Yingjing is trying to achieve an *instrument-independent* onset detector using more modern deep convolutional networks without hand-crafted features.

Yingjing has worked very hard and independently. He takes ideas from meetings, implements and evaluates them and reports back, sometimes producing written reports with graphs when I cannot meet in person. He is really functioning as a very good graduate research assistant, even though he is taking a full course load as a senior undergraduate. He's the best undergraduate researcher I've seen at CMU in several years.

Yingjing is an exceptional student, as evidenced by his independent research that he designed and pursued on his own, outside of and in addition to his regular curriculum. His *independent* work was accepted for the HDM workshop panel at ICDM. He is making great progress on our current project, and I think he will make an excellent PhD student.

Sincerely yours,

Roger B. Dannenberg

Roger B. Dannenberg
Professor of Computer Science, Art & Music

Northwestern | THE GRADUATE SCHOOL

Recommendation Form

The Graduate School Northwestern University Evanston, IL 60208-1113

Applicant Name: **Yingjing Lu**

Program: **Computer Science: MS**

Applicant Waived Rights*: **This applicant has waived the right to view their recommendation.**

Recommender Name: **Kang Guo**

Organization Name: **Goldman Sachs**

Title: **Associate, Software Developer**

E-mail Address: **kangguoroo@gmail.com**

Telephone Number: **9173431776**

Relationship to Applicant: **Employer**

Certification (Date): **11-18-2018**

*"Public Law 93-380, Educational Amendments Act of 1974, grants students the right to have access to letters of recommendation in their placement files. By selecting the "Waive access" option you are waiving access to these letters."

Dear Admissions Officer,

I had the pleasure of supervising Yingjing Lu during his software engineering internship at Goldman Sachs this past summer. I was very impressed with his dedication and drive to succeed. His contribution to our internal integrated development environment had significant impact and our developers thank him for this. I highly recommend him to Northwestern University.

Yingjing showed a high aptitude for independent learning. Many of the software development pipelines within Goldman Sachs are proprietary. Taking this in mind, Yingjing took the initiative to research these technologies before beginning work on his project. He also approached me on his own to ask questions about code and functionality that were unclear. I appreciate Yingjing's decision to try to solve problems on his own before seeking help. That being said, he also understood when to seek help when necessary, so that progress on his project would not be impeded. I feel that this is a seldom seen quality of interns, and will be a strong facet of Yingjing's personality that will serve him well in his future academic career.

Perhaps the most apparent aspect of Yingjing's attitude was his dedication to work. His tasks were challenging in that he needed to understand how Goldman Sachs's IDE functioned and how developers used it. Most new analysts have several months to get a feel for our internal software development tooling, but as a summer intern, Yingjing only had 10 weeks to not only learn about the tooling, but implement improvements. He knew that there were a lot of things to learn, so we would often come in early to do independent research. When he was blocked by a problem, he would contact members of our team as well employees from other teams to a wholistic set of opinions.

Summer interns can sometimes struggle with communication in a corporate environment, but Yingjing was an exception. At the beginning of his project, he created an architectural overview for the team to review. Afterwards, he regularly updated the team on his progress throughout the 10 weeks he spent with us. Finally, he wrote detailed documentation outlining his design decisions. I felt that Yingjing managed his time well and did a lot of project planning all on his own. He had a good grasp on what needed to be done each week and definitely strived to reach goals that he set for himself.

Yingjing was an outstanding intern that I truly enjoyed having with us this summer. I give him my full recommendation to Northwestern University, and I wish him the best of luck. Should you require an additional information from me, feel free to contact me.

Sincerely,
Kang Guo
Associate Software Engineer
Goldman Sachs
kang.guo@gs.com