Ben Usman

usmn@bu.edu
cs-people.bu.edu/usmn/

Research Interests

Unsupervised cross-domain image alignment and manipulation with generative models, such as flows and diffusion models. I also enjoy learning about ML for graphics, generalization and statistical learning theory, neural tangent kernels, overparameterization.

Education

2016 - now PhD in Computer Science
Boston University
Image and Video Computing Group
Advisor: Kate Saenko

2014 - 2015 MSc in Applied Mathematics and Computer Science
Moscow Institute of Physics and Technology
Skolkovo Institute of Science and Technology
Department of Control and Applied Mathematics

2010 - 2014 BSc in Applied Mathematics and Physics
Moscow Institute of Physics and Technology
Department of Innovation and Higher Technology (2012-2014)
Department of Problems of Physics and Energetics (2010-2012)

Fellowships and Awards

2012 Scholarships for Academic Excellence from Innovative Education Foundation

Work Experience

Research Assistant	now	Boston University
Research Intern	2018-21	Google Research, CA
	2017	Honda Research Institute, CA
	2014	ABBYY Language Services, Moscow
TA / TF / Grader	2017-19	BU CS542: Machine Learning
	2017	BU CS591: Deep Learning
Visiting Research Student	2016	UMass Lowell
	2015	MIT

Invited Talks

2021 **MetaPose: Fast 3D Pose from Multiple Views without 3D Supervision**, *Presentation at Google Perception Spotlight*
2019 **PuppetGAN: Cross-Domain Image Manipulation by Demonstration**, *Oral presentations at ICCV 2019, NECV 2019 and Google Perception Spotlight*
2019 **Deep Normalizing Flows for Density Estimation, Outlier Detection and Domain Adaptation**, *Seminar Talk at Hariri Institute for Computing*
2017 **Towards More Stable Domain Adaptation**, *Seminar Talk at Google CAM*

Publications

// 2021 *"Disentangled Unsupervised Image Translation via Restricted Information Flow" by B. Usman*, D. Bashkirova*, K. Saenko (in submission).*

We propose a new many-to-many image translation method that infers which attributes are domain-specific from data by constraining information flow through the network using translation honesty losses and a penalty on the capacity of the domain-specific embedding, and does not rely on hard-coded inductive architectural biases.

2021 *"MetaPose: Fast 3D Pose from Multiple Views without 3D Supervision" by B. Usman, A. Tagliasacchi, K. Saenko, A. Sud,* **CVPR 2022***. US Patent Application Submitted.*

We show that a small feed-forward network can quickly and accurately estimate 3D poses and camera parameters from multi-view imagery successfully resolving uncertainty of single-view pose predictions and providing additional regularization in poorly conditioned few-camera setup, and outperforming classical bundle adjustment.

2021 *"Evaluation of Correctness in Unsupervised Many-to-Many Image Translation" by D. Bashkirova, B. Usman, K. Saenko,* **WACV 2022***.*

Our evaluation protocol reveals that all existing unsupervised many-to-many translation models fail to infer which attributes are domain-specific and which are domain-invariant from data, and mostly rely on biases hard-coded into their architectures.

// 2020 *"Likelihood Ratio Minimizing Flows: Towards Robust and Quantifiable Neural Distribution Alignment" by B. Usman, A. Sud, N. Dufour, K. Saenko,* **NeurIPS 2020***.*

We show how to upper-bound an adversarial log-likelihood ratio domain alignment objective with a simple stable minimization objective, if the domain transformation is a normalizing flow, and show its relation to Jensen–Shannon divergence and GANs.

// 2019 *"Adversarial Self-Defense for Cycle-Consistent GANs" by D. Bashkirova, B. Usman, K. Saenko,* **NeurIPS 2019***.*

We show that the cycle-reconstruction loss causes embedding of low-amplitude structured noise into intermediate generated images. We propose an adversarial loss that prevents this kind of "cheating" and, as a result, improves translation accuracy.

2019 *"PuppetGAN: Cross-Domain Image Manipulation by Demonstration" by B. Usman, N. Dufour, K. Saenko, C. Bregler,* **ICCV 2019 Oral (4.3% acceptance rate)***.*

We propose a model that can precisely manipulate individual attributes of real images using only synthetic supervision for training, e.g. learn to realistically manipulate mouth expression or lighting on real human images using 3D rendered demonstrations.

// 2018 *"Unsupervised Video-to-Video Translation", D. Bashkirova, B. Usman, K. Saenko, arXiv.*

2018 *"VisDA: A Synthetic-to-Real Benchmark for Visual Domain Adaptation" by X. Peng, B. Usman, N. Kaushik, D. Wang, J. Hoffman, K. Saenko,* **CVPR Workshop 2018***.*

"Syn2Real: A New Benchmark for Synthetic-to-Real Visual Domain Adaptation" by X. Peng, B. Usman, K. Saito, N. Kaushik, J. Hoffman, K. Saenko, arXiv.

// 2017 *"Stable Distribution Alignment Using the Dual of the Adversarial Distance" by B. Usman, K. Saenko, B. Kulis,* **ICLR Workshop 2017***.*

Teaching

2020	CS542 Machine Learning	Final Project Supervisor and Guest Lecturer
2019	CS542 Machine Learning	Teaching Fellow (TF)
2017-18	CS542 Machine Learning	Homework Designer and Grader
2017	CS591 Deep Learning	Teaching Assistant (TA)

Service

Reviewed for NeurIPS17, CVPR17, ICRA17, CVPR18 (emergency reviewer), CVPR19, NeurIPS20, CVPR20 (outstanding reviewer, in top 3.8%), NeurIPS21.

Helped running VisDA challenges at ICCV17, ECCV18, ICCV19, and NeurIPS21 workshops.

Supervised a high-school student for the AI4ALL outreach research program during the summer of 2021, and a master student for their directed study during the summer of 2020.

Other interests

Outside machine learning, I enjoy learning about new ways of looking at math problems (mostly in statistics, linear algebra, geometry, optimization), paradigm-shifting features in programming languages (like Rust's ownership model), creative meta-programming (like narrowing python type annotations with z3), probabilistic programming languages, creative shader programming, better mental and programming models for hard problem domains like parallel computing or async user interaction, rules and patterns in art.