Akshaya Athwale

akshaya.athwale.1@ulaval.ca | Website | +1 581 922-6848

Overview

I specialize in geometry-informed deep learning, developing novel architectures that integrate prior domain knowledge to reduce the need for large datasets. I have extensive experience with the geometry of wide-angle and fisheye images.

PUBLICATIONS

- **A. Athwale, O. Ahmad, J.-F. Lalonde:**
  - "Network-agnostic distortion-invariant projections for wide-angle image understanding". **Under review**.
- **A. Athwale, I. Shili, Émile Bergeron, O. Ahmad, J.-F. Lalonde:**
  - "DarSwin-Unet: Distortion Aware Encoder-Decoder Architecture". **IEEE Winter Conference on Applications of Computer Vision (WACV) 2025**.
- **A. Athwale, A. Afrasiyabi, J. Lague, I. Shili, O. Ahmad, J.-F. Lalonde:**
  - "DarSwin: Distortion Aware Radial Swin Transformer". In **IEEE International Conference on Computer Vision, ICCV 2023**.
- **S. Sengupta, A. Athwale, T. Gulati, V. Lakshminarayanan:**
  - "FunSyn-Net: Enhanced Residual Variational Auto-encoder and Image-to-Image Translation Network for Fundus Image Synthesis". in **SPIE Medical Imaging Conference 2020**.
- **Y. Hold-Geoffroy, A. Athwale, J.-F. Lalonde:**
  - "Deep Sky Modeling for Single Image Outdoor Lighting Estimation". In **IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2019**.

EDUCATION

- **Université Laval** — **Quebec, Canada**
  Doctorate Candidate — **2021 - Present**
- **Aalto University** — **Espoo, Finland**
  Research Assistant/Project Employee — **2020 - 2021**
- **Indian Institute of Technology (IIT-ISM) Dhanbad** — **Dhanbad, India**
  Integrated Master of Technology in Mathematics and Computing GPA: 8.85/10.0 — **2015 - 2020**

RESEARCH EXPERIENCE

- **Graduate Research Assistant, Université Laval** — **Quebec City, Canada**
  Supervised By : Prof. Jean-Francois Lalonde and Ola Ahmad, August'21 - Present
  - **Project 1: Geometrically Informed Transformer for Wide-Angle Images**
    - **Objective:** Developed a geometrically informed model capable of understanding the underlying domain of wide-angle images and adapting to lens distortion.
    - Due to the scarcity and complexity of fisheye image datasets with multiple distortions, we designed a novel **Radial Transformer Network** that explicitly incorporates fisheye lens geometry. This enables the network to adapt to any fisheye image and handle unseen lens distortions at test time without requiring large amounts of training data or fine-tuning for unseen lens distortion at test time.
    - The geometric transformer encoder leverages the intrinsic properties of lens distortion to project the fisheye image onto a distortion-invariant polar representation and we propose to injects the distortion parameters into various modules of network like novel angular positional encoding to process this representation.
    - We validated the model on a classification task involving novel lens distortions unseen during training. The model successfully generalized to these distortions, outperforming conventional approaches lacking geometric awareness.
    - Our findings were published in a research paper presented at **ICCV 2023**, demonstrating the potential of geometry-informed deep learning for wide-angle imagery.
  - **Project 2: Distortion-invariant, model-agnostic wide-angle image projection**
    - **Objective:** Develop a distortion-invariant projection method for wide-angle images that generalizes across different network architectures, enabling seamless integration into production systems.
    - While distortion-aware models, including our radial transformer network, perform well, they remain highly network-specific. Building on insights from our previous work, we revisited the undistortion strategy to develop a projection method that is both **model-agnostic** and distortion-invariant, ensuring compatibility with a wide range of architectures.

- We proposed a novel projection method based on square-to-disc functions commonly used in computer graphics, which projects the input fisheye image into a consistent distoriton space with fewer artifacts compared to widely used projection methods such as perspective projection. This allows the projected images to be processed by any network, making our solution easy to deploy and integrate into real-world applications.
- The project has been submitted and is currently **under review**.

- **Research Intern, Thales Digital Solutions** **Montréal, Canada**

Supervised By: Ola Ahmad and Prof. Jean-François Lalonde, June 2023 – October 2023

  - **Objective:** Extend the geometric model from our previous work to handle more realistic pixel-level tasks and propose a new distortion-aware projection function that improves sparsity in sampling between different distortion.
  - Proposed a novel encoder-decoder transformer architecture for depth estimation, integrating geometric distortion awareness from wide-angle images to improve real-world applicability.
  - Developed an improved sampling strategy for the distortion-aware polar projection, minimizing spatial errors across projections from images with varying distortion levels.
  - Demonstrated the effectiveness of the approach on depth estimation tasks, leading to a research paper presented at **WACV 2025**.

- **Research Assistant/Project Employee, Aalto University** **Espoo, Finland**

Supervised By : Prof. Alexander Ilin, May'20 - August'21

  - **Objective:** Develop a predictive model for catalysts using physics-informed neural networks (PINNs) based on Lagrangian and Hamiltonian mechanics to improve accuracy and physical consistency.
  - Proposed a novel method for predicting catalysts using physics-informed neural networks (PINNs) based on Lagrangian and Hamiltonian mechanics.
  - Developed a PINN framework that embeds the physical laws governing catalytic reactions directly into the learning process by modeling system dynamics using Lagrangian-based neural networks, minimizing the action integral to ensure the solution adheres to physical constraints.
  - Captured the energy-conserving nature of the catalytic system with Hamiltonian-based neural networks, improving both accuracy and physical consistency.

- **Mitacs Globalink Research Internship, University Laval** **Quebec City, Canada**

Guide : Prof. Jean-Francois Lalonde, May'18 - December'18

  - **Objective:** Deep Sky Modeling and Illumination Estimation from a Single Low Dynamic Range (LDR) Image.
  - Due to the scarcity of datasets with images and their corresponding ground truth illumination, we proposed a three-step architecture to generate physically plausible high dynamic range (HDR) sky environment maps in latitude-longitude format from a single LDR image. Our approach leverages complementary datasets: Laval HDR, which provides radiometrically calibrated and physically accurate HDR panoramas, and SUN360, which contains diverse LDR panoramas under various illumination conditions.
  - First, we learn a latent space representation of HDR skies using the Laval HDR dataset. We train the encoder to be robust to distortions caused by variations in white balance, exposure, and occlusions. The learned HDR representations are used as supervision for single LDR images cropped from corresponding panoramas to model the sky and estimate illumination.
  - Second, we convert the LDR panoramas from SUN360 to HDR using an existing network and extract representations of these panoramas using the autoencoder.
  - Finally, we train an encoder using LDR images cropped from SUN360 panoramas, with supervision from their corresponding HDR illumination representations.
  - This work was published as a conference paper at **CVPR 2019**.

- **Research Internship, University of Waterloo** **Waterloo, Canada**

Guide : Prof. Vasudevan Lakshminarayanan, May'19 - July'19

  - Objective : Generating synthetic retinal images.
  - We proposed a pipeline of two models for generating of retinal images and corresponding blood vessel.
  - First model outputs the blood vessel using enhanced Residual Variational Auto-Encoder using a random normal vector as input. Second model is image-to-image translation network which takes the generated blood vessel as input to outputs retinal fundus image.
  - Results were clinically relevant and our research paper was published in **SPIE Medical Imaging Conference 2020**.

REFERENCES

- **Prof. J. F. Lalonde**: Professor, Department of Electrical and Computer Engineering, University Laval, jflalonde@gel.ulaval.ca
- **Ola Ahmad**: Chief AI Scientist at Thales Canada, ola.ahmad@thalesdigital.io
- **Prof. Alexander Ilin**: Professor, Computer Engineering, Aalto University, +358415014497, alexander.ilin@aalto.fi
- **Prof. Vasudevan Lakshminarayanan**: Professor, Electrical and Computer Engineering, University of Waterloo, vengulak@uwaterloo.ca
- **Deekshith Marla**: Founder & CTO (Featured in Forbes 30 Under 30), Lithasa Technology, Arya.ai, deekshith@arya.ai