Event-Based Deep Learning Model for Advanced Image and Video Deblurring

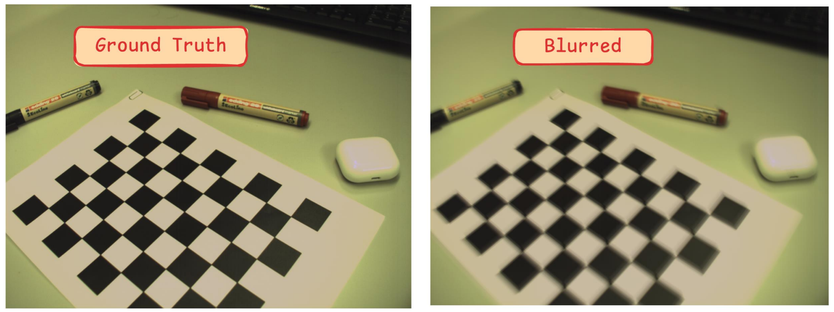
Yashu GautamKumar Patel, Computer Science (B.S)

Mentor: Dr. Bharatesh Chakravarthi, Assistant Teaching Professor
School of Computing and Augmented Intelligence

QR CODE

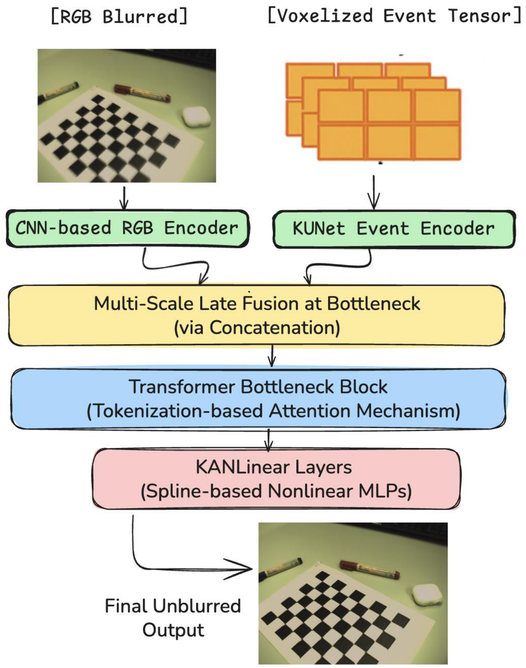
Introduction and Background

- **Problem :** Conventional cameras produce blurry images when capturing fast-moving scenes, especially in night-time or low-light conditions. This causes critical details to be lost, impacting real-world tasks like autonomous navigation, surveillance, or robotics.



- **Impact :** Blur limits both human understanding and AI decision-making. A drone can miss a street sign, or a robot can misread a hazard. Fixing blur means better safety and more reliable vision systems.
- **Our Goal :** Create a lightweight yet powerful deep learning model that can remove motion blur using both RGB and event camera inputs, making it feasible to run on real-time devices like drones and rescue bots—even at night.

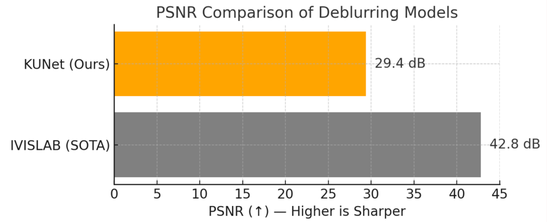
KUNet Model Architecture



How KUNet works :

UNet takes a blurry RGB frame and a time-synced event stream, voxelized into a 6-channel tensor. Each stream is processed through a dedicated encoder - CNN layers for RGB and a multi-scale KUNet Event Encoder for events. These feature streams are fused at the bottleneck using late multi-scale fusion, which preserves both fine motion (short events) and broader context (mid and RGB features). The fused representation is refined through a Transformer block and spline-based KANLinear layers, resulting in a clean, deblurred output image.

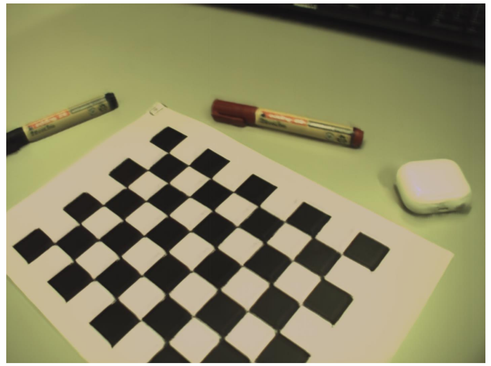
Performance :



Results & Outcomes

Blurred Image :

Deblurred Image (By KUNet) :



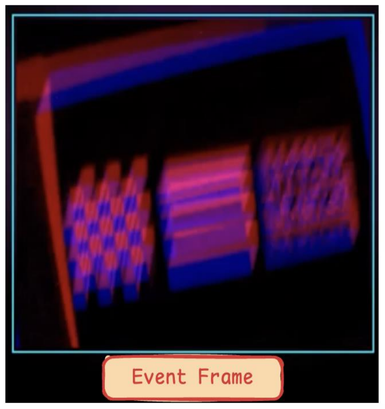
Key Takeaways :

- KUNet successfully restores major image details using both RGB and event data — even in dark or high-motion scenes.
- Although its PSNR **(29.4**) is below the SOTA model **(42.8),** it runs **10× faster**.

KUNet Inference Time: **~1–2 sec/image**
IVISLAB (SOTA): **~15–25 sec/image**

Challenges in Event-Based Deblurring

Dual-Modality Fusion	RGB and Event data look very different, combining them without losing details is tricky.
Non-Uniform Blur	Some parts blur more than others (e.g., roads vs. moving cars), no one-size-fits-all fix.
Night-Time Noise	Event cameras can trigger from tiny flickers, this causes false signals in the dark.
Real-Time Limits	The model must be small enough to run fast on devices like Jetson Orin.
Hard-to-Get (Ground Truth)	Training data is scarce, especially sharp, aligned images from real moving scenes.



What's Next ?

- **Make it real-time:** Build a live deblurring pipeline that works with continuous RGB + Event camera feeds, like a video, not just images.
- **Run it on drones:** Optimize KUNet for Jetson Orin and other edge devices so it can fly with drones or robots and run on-the-go.
- **Handle the dark better:** Train with noisy, low-light scenes (like night parking lots) to make the model more reliable in harsh conditions.

FURI

KEY TERMS

PSNR (Peak Signal-to-Noise Ratio) – A score that measures image clarity. Higher PSNR = sharper, more accurate output.
Event Frame – A sensor-based representation of motion captured as light flickers over time, not regular RGB pixels.
Voxelization – Turns events into 3D pixel blocks (x, y, t) to align them with blurred images.

REFERENCES:

[1] **Sun, L., Alfarano, A., Duan, P., et al.** "NTIRE 2025 Challenge on Event-Based Image Deblurring: Methods and Results." *arXiv preprint arXiv:2504.12401*, Apr. 2025. https://www.themoonlight.io/file?url=https%3A%2F%2Farxiv.org%2Fpdf%2F2504.12401

Ira A. Fulton Schools of Engineering
Arizona State University