On-device Super-Resolution Tuned for Foveated Displays

Azlarkhon Khasankhujaev
Muhammadsodiq Qobiljonov
Mirsolih Mirsagatov

1 Title & Team

Project Title: On-device Super-Resolution Tuned for Foveated Displays
Team Members: Azlarkhon Khasankhujaev, Muhammadsodiq Qobiljonov, Mirsolih Mirsagatov
GitHub Repository: https://github.com/Azlarkhon/On-device-super-resolution-tuned-for-foveating-displays
Coordinator: Azlarkhon Khasankhujaev

2 Abstract

We propose an on-device, gaze-aware super-resolution (SR) framework tailored for foveated displays—such as eye-tracking head-mounted systems—where the user's gaze is used to concentrate computational effort on the foveal region while reducing resources in the periphery. By leveraging eye-tracking to generate a foveation map, we partition the frame into high-fidelity (fovea) and lower-fidelity (periphery) zones, apply a modular SR network with latency-aware scheduling, and optimize using perceptual losses (LPIPS, SSIM) that reflect human visual sensitivity. We target real-time interactive performance on mobile/embedded GPU hardware. Evaluation will measure end-to-end latency and FPS, quantitative image metrics (LPIPS, SSIM) both regionally and globally, and a user study capturing MOS (Mean Opinion Score) for perceived sharpness. The expected outcome is a deployable prototype on device achieving $\geq$60 FPS with latency below $\sim$20 ms, a measurable improvement in perceptual quality for the foveal region compared to a uniform SR baseline, and a documented increase in user-perceived sharpness.

3 Problem & Motivation

High-resolution near-eye displays (NEDs) and head-mounted AR/VR systems are increasingly demanding in terms of rendering and compute resources. Simultaneously, the human visual system (HVS) exhibits a steep decline in visual acuity from the fovea (central few degrees) to the periphery. Despite this, many super-resolution (SR) approaches treat the full image uniformly, allocating compute to regions of the frame the user may not closely scrutinize. For on-device scenarios (mobile/embedded GPU) where bandwidth, memory, power and latency are constrained, this inefficiency poses a real bottleneck.

By designing a gaze-aware SR pipeline that spatially and temporally allocates compute—focusing on the foveal zone and reducing resources in the periphery—we can reduce computational overhead, decrease latency, maintain high perceptual quality where it matters, and extend battery life on portable devices. Stakeholders include device manufacturers (reduced hardware demands), AR/VR application developers (better performance/responsiveness), and end-users (sharper perceived imagery with lower lag). Measurable goals for this project include: achieving $\geq$60 FPS interactive performance, end-to-end latency under $\sim$20 ms, $\sim$10% improvement in LPIPS in the foveal region, and increasing user MOS for sharpness by $\geq$0.5 points on a 5-point scale.

4 Related Work

We summarize key prior efforts in foveated rendering, super-resolution, and gaze-aware compute scheduling:

Year	Paper	Key idea	Relation to our work
2021	Wang et al., "FOCAS: Practical Video Super Resolution using Foveated Rendering"	SR using more blocks in fovea, fewer in periphery; latency budget modeled.	Closest baseline; we extend to on-device latency scheduling.
2022	Nam, Kang & Cho, "Foveated Super Resolution Network for VR HMDs"	Integrated foveated SR algorithm for VR head-mounted displays.	Similar domain; less focus on latency scheduling.
2023	Lee et al., "Cross-Resolution Flow Propagation for Foveated Video SR"	Video SR fused with foveated region, uses gaze region as HR context.	Video-specific; we target real-time adaptation.
2023	Ye et al., "Neural Foveated Super-Resolution for Real-time VR Rendering"	Neural accumulator and partition-assemble for real-time foveated SR.	Advanced method; we adapt for embedded devices.
2020	Malkin et al., "CUDA-Optimized Real-time Rendering of a Foveated Visual System"	GPU-based foveated rendering; discusses acuity-based models.	Foundational for gaze-aware compute.

Table 1: Key related work summary

In summary, existing literature explores foveated SR in high-performance setups. Our novelty lies in making it efficient and deployable **on-device**, combining perceptual loss optimization with latency-aware scheduling and a user-centered evaluation of perceptual sharpness.

5 Data & Resources

Datasets: We will use DIV2K and Vimeo90K (for video SR), both research-licensed and widely used for image/video super-resolution. For gaze tracking, we may collect a small consent-based dataset using an eye-tracking headset to fine-tune foveation behavior.

Compute/Hardware: Training on an NVIDIA RTX 3080/4090; deployment on a Jetson Orin or equivalent mobile GPU. Framework: PyTorch for training, ONNX/TensorRT for inference.

Ethics/Privacy: All human data collection (if any) will involve informed consent and anonymization. Gaze data will not contain personal identifiers, and all dataset licenses will be respected.

6 Method

Baselines to Reproduce

- Uniform SR network (EDSR or RCAN) applied across full frame.
- Foveated SR method from FOCAS and Nam & Kang: higher SR capacity for fovea, reduced in periphery.

Proposed Method

1. Real-time eye-tracking yields gaze coordinates each frame.
2. Generate a foveation map based on visual acuity distribution.
3. Partition the frame into foveal and peripheral regions.
4. Apply a high-capacity SR model on the fovea and a lightweight SR on the periphery.
5. A latency-aware scheduler dynamically adjusts peripheral SR quality to sustain FPS.
6. Use region-weighted perceptual losses (LPIPS, SSIM) with higher weight for the fovea.
7. Conduct ablations: (a) with/without scheduler, (b) gaze-aware vs uniform SR, (c) different foveal radii, (d) varying region-weight weights.

7 Experiments & Metrics

Measurements:

- End-to-end latency (gaze-to-frame, ms).
- FPS on device.
- LPIPS, SSIM (global and foveal).
- User study: MOS (5-point scale) for perceived sharpness.

Success thresholds:

- Latency < 20 ms, FPS ≥ 60.
- Foveal LPIPS improvement $\geq$10%.
- SSIM > 0.90 in foveal region.
- MOS $\geq$0.5 improvement vs baseline.

We will analyze trade-offs between latency and perceptual gains, and perform ablations to show scheduler and gaze-awareness effectiveness.

8 Risks & Mitigations

- **Compute bottleneck:** May exceed latency budget. *Mitigation:* Quantize/prune models; lower periphery SR load.
- **Eye-tracking error:** Gaze misalignment causes blur. *Mitigation:* Predictive gaze smoothing and tolerance margins.
- **Domain shift:** SR datasets differ from on-device data. *Mitigation:* Fine-tune on small device-specific dataset.
- **User study noise:** MOS variance high. *Mitigation:* Pilot test and balanced design ($N \geq 20$).

9 Timeline & Roles

Week-by-Week Snapshot

Given that we start on 27 October and end on ~27 December (approx. 9 weeks), we propose the following schedule:

Week	Dates	Milestone	Owner
W1	27 Oct–2 Nov	Repository setup + ROADMAP.md	Azlarkhon
W2	3–9 Nov	Baseline SR implementation	Muhammadsodiq
W3	10–16 Nov	Eye-tracker integration + foveation map	Mirsolih
W4	17–23 Nov	Proposed SR architecture + losses	Azlarkhon
W5	24–30 Nov	Desktop prototype + latency profiling	Muhammadsodiq
W6	1–7 Dec	On-device deployment + tuning	Mirsolih
W7	8–14 Dec	User-study design + pilot test	Azlarkhon
W8	15–21 Dec	Ablations + data collection	Muhammadsodiq & Mirsolih
W9	22–27 Dec	Final evaluation + report + demo	All

Table 2: Project timeline and milestones

Roles (RACI Style)

Task	Responsible	Accountable	Consulted	Informed
Baseline SR code	Muhammadsodiq	Azlarkhon	Mirsolih	All
Eye-tracker integration	Mirsolih	Azlarkhon	Muhammadsodiq	All
On-device optimization	Mirsolih	Azlarkhon	Muhammadsodiq	All
User-study design	Azlarkhon	Azlarkhon	Mirsolih, Muhammadsodiq	All
Final write-up	All team	Azlarkhon	—	All

Table 3: Team roles and responsibilities

Repository Roadmap: https://github.com/Azlarkhon/On-device-super-resolution-tuned-for-foveating-displays/blob/main/ROADMAP.md

10 Expected Outcomes

Deliverables:

- Full training and deployment code for gaze-aware SR.
- Detailed report with experiments and ablations.
- Poster and video demo of real-time results.

Stretch goals:

- 90 FPS real-time capability.
- Adaptive upscale factor (e.g., 4× in fovea, 2× periphery).
- Power/energy reduction analysis.

11 Ethics & Compliance

If human gaze data is collected, IRB or equivalent approval will be sought. Informed consent and anonymization will be ensured. Dataset licenses will be observed. The user study will include diverse participants and check for fairness (no disadvantage to low-vision or demographic groups). Data will be securely stored and destroyed after use.

References

1. L. Wang et al., "FOCAS: Practical Video Super Resolution using Foveated Rendering," *Proc. ACM Multimedia*, 2021.
2. H. Nam, H. Kang, and H. Cho, "Foveated Super Resolution Network for Virtual Reality Head-Mounted Displays," *SID Symposium Digest of Technical Papers*, vol. 53, no. 1, pp. 869–872, 2022, doi:10.1002/sdtp.15631.
3. E. Lee, L.-F. Hsu, E. Chen, and C.-Y. Lee, "Cross-Resolution Flow Propagation for Foveated Video Super-Resolution," *WACV*, 2023.
4. J. Ye et al., "Neural Foveated Super-Resolution for Real-time VR Rendering," *arXiv preprint arXiv:2304.04590*, 2023.
5. E. Malkin, A. Deza, and T. Poggio, "CUDA-Optimized Real-time Rendering of a Foveated Visual System," *arXiv preprint arXiv:2008.10778*, 2020.