**Motivation**
- Various complex temporal modeling techniques in existing VSR methods hinder the deployment of VSR networks on resource-constrained platforms, e.g., smartphones and wearable devices.
- To reduce the required computational cost and memory consumption, we argue that, designing efficient alignment and multi-frame fusion/refinement schemes are the keys to lightweight VSR networks.

**Contribution**
- A novel **motion propagation scheme** is proposed for efficient feature alignment, which estimates preliminary motion fields called "difference flow" and modifies them per frame pair adaptively. A dense alternative of RFDB is designed to distill and refine warped features **without deep stacked structures**, which maintains the efficiency and effectiveness of the network.
- The proposed network achieves comparable performance with state-of-the-art VSR methods on two benchmark datasets Vid4 and Vimeo90K-T while **enjoying a clear advantage in model size and computational efficiency**.

**Efficiency**
- Tradeoff between GFLOPs and PSNR
- Insensitive GFLOPs grows when input frames increases

**Methods**
- Bidirectional recurrent arch to propagate difference flows

**Quantitative results**

**Visual results**

**Ablation studies**

**Feature refinement**

GFLOPs are computed for generating an $512 \times 1024$ output frame.