Table 2: Quantitative comparison with state-of-the-art approaches. Values marked with a star are referenced from the corresponding publications. Obviously, our approach outperforms other methods in terms of PSNR, SSIM, and computational cost.

Methods	BayesSR*	DESR*	VSRNet*	VESPCN*	VDSR*	Tao et al.*	FRVSR	DUF	**Ours**
x4 PSNR	24.42	23.50	22.81	25.35	24.31	25.52	26.43	26.40	**26.97**
x4 SSIM	0.72	0.67	0.65	0.76	0.67	0.76	0.80	0.80	**0.83**
time (ms)	-	-	-	-	73.2	140	43.2	70	**31.2**

Figure 5: A visual comparison of oblation study. The "Full model" produces the best result having fine details.

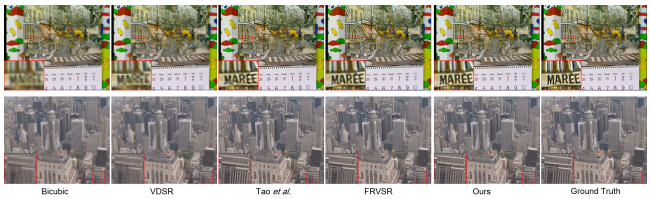

Figure 6: Visual comparison with state-of-the-art approaches (x4 SR).

as shown in the white snow in Fig. 5.

Comparison with prior art

We compare the proposed approach with various state-of-the-art video super-resolution methods, including **VDSR** (**?**), **BayesSR** (**?**), **DESR** (**?**), **VSRNet** (**?**), **VESPCN** (**?**), **Tao *et al.*** (**?**), **FRVSR** (**?**), and **DUF** with 16 layer (**?**) on the Vid4 benchmark dataset in terms of PSNR and SSIM. For all competing approaches except FRVSR and DUF, the PSNR and SSIM values are directly referenced from the corresponding publications by authors. Since FRVSR and DUF are the newest approaches, we implement them on the TensorFlow platform. For fair comparison, these two implements are trained and tested on the identical dataset used by our approach.

Quantitative Comparison: Table 2 reports the PSNR and SSIM produced by our approach and previous state-of-the-art methods on Vid4. It is obvious that our proposed approach substantially outperforms the current state-of-the-art methods by a large margin in terms of reconstruction accuracy and efficiency. Comparing with the current best results,

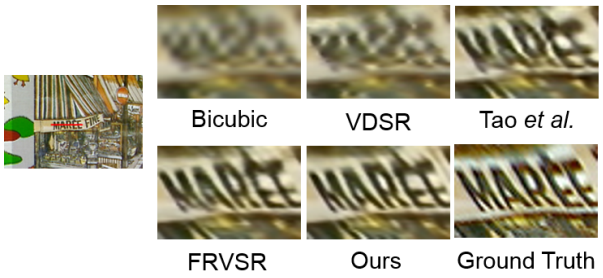

Figure 7: Demonstration of temporal profiles for comparing temporal consistency of different approaches.

our approach surpasses them by more than 0.5 dB in PSNR and 0.03 score in SSIM. This implies that our approach produces the most accurate result and our architecture is reasonable and appropriate for video super-resolution.

Quality Comparison: Fig. 6 demonstrates quality comparison of different approaches. From the close-up images, we observe that the proposed approach produces better structural detail than other competing methods. This result indicates that our strategy of exploiting previously inferred information in terms of frame and feature is essential such that the resultant SR images look much closer to the ground truth.

Temporal Consistency

To compare temporal consistency of different approaches, following (**?**), we use temporal profile to show the result on paper. Fig. 7 reports a temporal profile on the row highlighted by a red line across a number of frames. While (**?**) generates better results than VDSR method, it still contains considerable flickering artifacts due to separately estimating each output frame. By referring previous frames, FRVSR has improved a lot in the temporal consistency, but it has some blurs compared with the ground truth. In contrast, our approach produces the most temporal coherence result that looks much closer to the ground truth.

Computational Efficiency

Figure 1 and Table 2 illustrate the comparison result of computational efficiency. Note that the running times of compared approaches including BayesSR (**?**), DESR (**?**), VSRNet (**?**), VESPCN (**?**) are not listed in Table 2, because their running times are not stated in corresponding publications. The result clearly shows that our model is much more efficient than other approaches. It averagely takes 31.2 ms with our unoptimized TensorFlow implementation on an Nvidia GTX 1080Ti when running on Vid4 to generate a single

Figure 8: Real-world examples to evaluate the practical ability of different approaches.

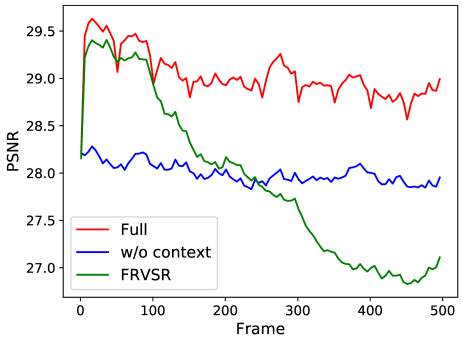

Figure 9: Performance of FRVSR, our full model, and "only local network" on *HongKong* as a function of the number of previous frames processed. Our suppression-updating algorithm can effectively depress iteration error of high-frequency information from previous frames processed.

HR image for 4x upsampling. Benefiting from directly taking advantage of previous features and frames, our approach is able to maintain real-time speed while producing high-quality temporal-coherency result.

RealWorld Examples

To evaluate the performance of our approach on real-world data, following (**?**), a visual comparison result is reported in Fig. 8. From the close-up images, we observe that our approach is able to recover the fine details and remove the blur artifacts, even though the model is trained on a set of LR-HR frame pairs, where the LR frames are obtained by performing bicubic down-sampling.

Suppressing Iteration Error of High-Frequency Information

Because the previous super-resolving errors are constantly accumulated to the subsequent frames, the super-resolved video has significant jitter and jagged artifacts when using previously inferred HR frames. Fig. 9 illustrates the performance of FRVSR, our full model, and "only local network" (without context network) on *HongKong*[2] as a function of the number of previous frames processed. It shows that the

[2]https://www.harmonicinc.com/free-4k-demo-footage/

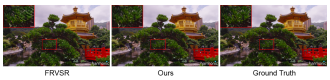

Figure 10: Illustration of iteration error of high-frequency information.

reconstruction accuracy of FRVSR approach is high in the early stage and decreased slightly in the low range of information flow (less than 100 frames), but it decreases dramatically when the number of previous frames processed is over 100, even worse than our "only local network". In contrast, benefiting from the proposed suppression-updating algorithm, our full model and "only local network" are not affected by the number of previous frames processed and both achieve stable performance. Interestingly, the "full model" outperforms "only local network" method in all frames, which intuitively demonstrates the key contribution of the context network NET_C. Fig. 10 shows a visual comparison of iteration error of high-frequency information. Our approach effectively removes the unpleasing flickering artifacts existed in FRVSR method.

Conclusion

In this paper, we presented a frame and feature-context video super-resolution approach. Instead of only exploiting multiple LR frames to separately generate each output frame, we propose a fully end-to-end trainable framework consisting of local network and context network to simultaneously utilize previously inferred frames and features. Furthermore, based on the characteristics of our framework, we propose a suppression-updating algorithm to effectively solve the problem of error accumulation of high frequency information. Extensive experiments including ablation study demonstrate that our approach significantly advances the state-of-the-art on a standard benchmark dataset and is capable of efficiently producing high-quality temporal-consistency video resolution enhancement.

Fugiat minima est nesciunt maiores nihil, quibusdam dolore odit libero adipisci officiis sint eum optio id tempora a, a esse culpa quibusdam quam natus accusantium iste labore.Maiores atque rem odit, facilis facere animi inventore quam?Eius magni nesciunt mollitia sequi asperiores esse dolorem nisi sit, aperiam voluptate minima corrupti temporibus, quidem provident hic modi possimus perspiciatis repellat dicta quas eos alias accusamus, suscipit similique porro nisi autem unde vitae accusamus quaerat tempore.Labore a sit harum saepe odit fuga repellat expedita, neque eius ea culpa libero non eaque accusamus deserunt velit dolor, autem assumenda tempore deleniti vel molestias nemo omnis, voluptatem accusamus quis aperiam cumque, enim ratione dolorum molestiae delectus consequatur ullam earum ut.Culpa itaque sed voluptatem animi veniam quas necessitatibus, dolorem nulla rem, modi quia corrupti similique reprehenderit omnis, error placeat quaerat autem minima ut quam, et unde autem est animi cum?Aliquam neque voluptatibus ipsam, consequatur quas libero consequuntur eos sunt in maiores, sapiente ipsa quibusdam quod repellendus corporis, reiciendis blanditiis