SUPER RESOLUTION IMAGE

Team Members:

Computer Vision – CS512 – F24
Professor: Agam Gady

Nandini Devaraj A20593534
Rithika Kavitha Suresh A20564346

Problem Statement and Background Material

Objective of Image Super-Resolution (SR):

The goal of SR is to reconstruct high-resolution images from low-resolution inputs, which is crucial for various applications like medical imaging, surveillance, and image enhancement.

Challenges:

- Feature Utilization: Existing methods often fail to fully exploit the hierarchical features present in low-resolution images, leading to suboptimal performance.
- Network Convergence: As networks deepen, convergence speed can degrade, complicating training processes.
- Depth Utilization: Effectively leveraging features from different depths of the network remains a challenge, often leading to inefficient learning.
- Computational Efficiency: Balancing high-quality reconstruction with manageable computational costs is essential, especially in real-time applications.

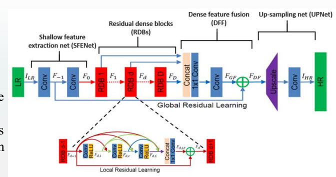
Residual Dense Network (RDN)

- Shallow Feature Extraction Network (SFENet):
Extracts initial features from low-resolution images.

- Residual Dense Blocks (RDBs):
These blocks utilize dense connections and local feature fusion to enhance feature extraction while maintaining computational efficiency.
Each RDB allows direct connections from previous layers, facilitating a contiguous memory mechanism that improves information flow.

- Dense Feature Fusion (DFF):
Combines features from all preceding layers to create a comprehensive representation of the input data.

- Up-sampling Network (UPNet):
Upscales the processed features to generate high-resolution outputs.



Key Innovations:

- The use of local residual learning and global feature fusion enhances both local and global feature extraction capabilities, significantly improving SR performance compared to traditional methods.

Implementation Details

Input: Low-resolution image ILR

Output: Super-resolved image ISR

// Shallow Feature Extraction

F-1 = Conv(ILR)

F0 = Conv(F-1)

```
// Residual Dense Blocks
for d = 1 to D:
    Fd,0 = Fd-1
    for c = 1 to C:
        Fd,c = ReLU(Conv([Fd,0, Fd,1, ..., Fd,c-1]))
    Fd,LF = LFF([Fd,0, Fd,1, ..., Fd,C])
    Fd = Fd-1 + Fd,LF

// Dense Feature Fusion
FGF = GlobalFeatureFusion([F1, F2, ..., FD])
FDF = F-1 + FGF

// Up-sampling
FUP = SubPixelConv(FDF)
ISR = Conv(FUP)

return ISR
```

Implementation Details…contd

Training Details

Datasets:
DIV2K dataset: 800 training images, 100 validation images
Flickr2K dataset: 2650 high-quality images

Degradation Models:
Bicubic downsampling (BI)
Blur-downsample (BD)

Loss Function:
L1 loss between reconstructed SR image and ground truth HR image
Optimization:
ADAM optimizer

VALIDATION DETAILS

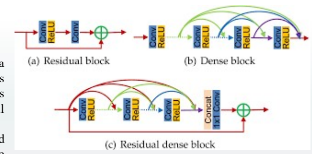
Benchmark Datasets:
1.Set5
2.Set14
3.B100
4.Urban100
5.Manga109

Evaluation Metrics:
1.PSNR (Peak Signal-to-Noise Ratio)
2.SSIM (Structural Similarity Index)
Scaling Factors:
x2, x3, x4 upscaling

Comparison Methods:
SRCNN, VDSR, DRCN, LapSRN, DRRN, MemNet, EDSR, SRMDNF

Model Architecture:

Residual Dense Block (RDB):
- The Residual Dense Block (RDB) is implemented as a PyTorch class (ResidualDenseBlock). Each block consists of four convolutional layers with dense connections between them, where each layer receives input from all previous layers within the block.
- Each convolutional layer has a kernel size of 3×33×3 and padding of 1, ensuring that the spatial dimensions are preserved throughout the block.
- After each convolution, a ReLU activation function is applied to introduce non-linearity.
- A final 1×11×1 convolution is used to reduce the number of channels after concatenating all outputs from previous layers. This operation helps in controlling the computational complexity while still maintaining rich feature information.

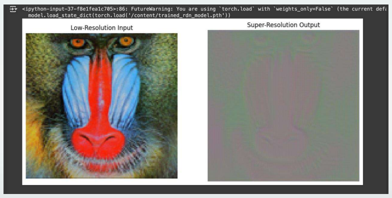


(a) Residual block (b) Dense block

(c) Residual dense block

Residual Dense Network (RDN):
- The RDN is built by stacking multiple RDBs. In this implementation, six RDBs (num_blocks=6) are used to form the core of the network.
- After passing through all RDBs, the output is passed through a final 3×33×3 convolutional layer to generate the super-resolved image.
- The forward pass concatenates outputs from each RDB and passes them through subsequent blocks to ensure efficient feature reuse and better gradient flow.

In the process..



Results and Performance

•RDN outperformed state-of-the-art methods on benchmark datasets (Set5, Set14, B100, Urban100, Manga109) for different scaling factors (x2, x3, x4)

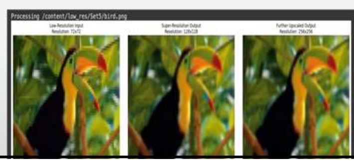
•RDN achieved higher PSNR and SSIM values compared to other methods like SRCNN, VDSR, Bicubic.

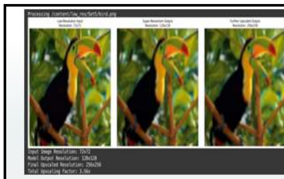
NO.	Evaluation index	Bicubic	SRCNN	VDSR	RDN
1	PSNR	30.98	32.44	33.62	33.78
	SSIM	0.711	0.737	0.754	0.749
2	PSNR	33.85	33.97	34.21	34.35
	SSIM	0.728	0.741	0.759	0.761
3	PSNR	36.85	37.93	38.18	38.45
	SSIM	0.684	0.712	0.723	0.732
4	PSNR	30.65	31.12	31.36	32.45
	SSIM	0.744	0.773	0.786	0.786
5	PSNR	30.95	32.06	32.13	32.17
	SSIM	0.779	0.809	0.827	0.831

Performance Across Different Images:

- Visual comparisons showed that RDN produced sharper edges and more accurate textures compared to other methods.
- RDN demonstrated superior performance in recovering fine details and complex structures in images

RDN achieved better performance with fewer parameters compared to some larger models like EDSR





Model Efficiency:

- The training process ran for 100 epochs, as shown in one of the screenshots, using an Adam optimizer. The model was trained on either a GPU or CPU depending on availability.
- Despite being trained for only 100 epochs, the model demonstrates strong performance across various test images, suggesting that RDN is efficient in learning high-quality representations from low-resolution inputs.
- RDN shows good performance across different scaling factors (x2, x3, x4), demonstrating its efficiency in handling various super-resolution tasks

Potential Limitations:

Dependency on Large Datasets:

The model is trained on a combination of DIV2K and Flickr2K datasets, totaling 3,450 high-quality images1

This reliance on large, high-quality datasets could be a limitation in domains where such extensive data is not available.

Limited Scaling Factors:

The paper primarily discusses results for x2, x3, and x4 upscaling. It's not clear how well the model would perform for more extreme scaling factors.

Future Improvements:

- To further improve performance, additional training epochs could be considered to allow for more fine-tuning of weights.
- Incorporating more advanced loss functions such as perceptual loss or GAN-based approaches could help enhance texture details and reduce blurring in high-resolution outputs.
- Experimenting with deeper network architectures or adding more Residual Dense Blocks (RDBs) may also improve performance for larger scaling factors.

Conclusion

The proposed Residual Dense Network (RDN) demonstrates superior performance for image super-resolution by effectively utilizing hierarchical features from the original low-resolution image.

Key innovations of RDN include:

- 1. Residual Dense Blocks (RDBs) that allow for dense local feature extraction and efficient information flow through contiguous memory mechanisms.
- 2. Local Feature Fusion (LFF) within RDBs to adaptively learn and preserve important local features.
- 3. Global Feature Fusion (GFF) to combine hierarchical features from all RDBs in a holistic manner.
- 4. Global residual learning to integrate shallow and deep features.
- RDN achieves state-of-the-art performance on benchmark datasets across different degradation models (BI, BD, DN) and scaling factors (x2, x3, x4). The network's ability to fully exploit hierarchical features leads to superior reconstruction quality, especially for complex textures and structures.
- The proposed architecture demonstrates good efficiency in terms of parameter usage and computational complexity relative to its performance gains. RDN's effectiveness in extracting and utilizing multi-level features from the original low-resolution input makes it a promising approach for high-quality image super-resolution tasks.
- Overall, RDN represents a significant advancement in deep learning-based image super-resolution, offering a novel and effective framework for exploiting hierarchical features in the low-resolution space.

THANK YOU!

References

- Zhang, Y., Li, K., Li, K., Wang, L., Zhong, B. and Fu, Y., 2018. Image super-resolution using very deep residual channel attention networks. In *Proceedings of the European conference on computer vision (ECCV)* (pp. 286-301).
- Zhang, Z., Tang, Z., Wang, Y., Zhang, Z., Zhan, C., Zha, Z. and Wang, M., 2021. Dense residual network: enhancing global dense feature flow for character recognition. *Neural networks*, *139*, pp.77-85.
- Liu, C., Sun, X., Chen, C., Rosin, P.L., Yan, Y., Jin, L. and Peng, X., 2019. Multi-scale residual hierarchical dense networks for single image super-resolution. *IEEE Access*, *7*, pp.60572-60583.
- An, Z., Zhang, J., Sheng, Z., Er, X. and Lv, J., 2021. RBDN: Residual bottleneck dense network for image super-resolution. *IEEE Access*, *9*, pp.103440-103451.