# 画像超解像 (SR) モデル技術マップ

大谷 豪 (慶應義塾大学), 児玉 憲武 (横浜国立大学), 松尾 雄斗 (東京電機大学)

**PSNR-oriented**
ピクセル単位での整合性を重視

**Perception-oriented**
視覚的観点からの画像品質を重視

## CNN

### CNNを導入
- **SRCNN** (ECCV' 14)
  3つの畳み込み層で構成

### アップスケーリング
- **FSRCNN** (ECCV' 16)
  deconvolutionを導入
- **ESPCN** (CVPR' 16)
  pixel shufflerを提案

### 効率化
- **CARN** (ECCVW' 18)
  抽出した特徴をアンカーで分類後に回帰
- **LatticeNet** (ECCV' 20)
  ラティスブロックを使用して残差ブロックの出力を効率的に結合
- **LAPAR** (NeuIPS' 20)
  事前定義された複数のフィルタ基底の辞書を使用

### 深層化
- **VDSR** (CVPR' 16)
  VGG-netを応用SRへの対応のため残差接続を追加
- **SRResNet** (CVPR' 17)
  ResNetを応用残差ブロックを導入
- **EDSR** (CVPRW' 17)
  BN層を削除残差スケーリングを追加
- **RCAN** (ECCV' 18)
  チャネル注意機構を追加RIR構造を採用

RIR構造を踏襲

### モジュール追加
- **MemNet** (ICCV' 17)
  Gate Unitの追加情報の長期記憶を実現
- **SRDenseNet** (ICCV' 17)
  DenseNetを応用NNの性能向上を促進
- **RDN** (CVPR' 18)
  residual dense blockを提案

## Transformer

### Encoder-Decoder型
- **IPT** (CVPR' 21)
  **Transformerを導入**
  複数のlow-level visionタスクを同時に事前学習
- **EDT** (IJCAI' 23)
  単一のlow-level visionタスク内で事前学習

### 効率化
- **ELAN** (ECCV' 22)
  シフト畳み込みとGMSAモジュールによるELABを提案
- **EMT** (Engineering Applications of Artificial Intelligence' 24 )
  自己注意層をPixel Mixer層にいくつか置き換えたMTBにより局所情報の処理を強化
- **SRFormer** (ICCV' 23)
  チャネル数を削減しつつウィンドウサイズを拡大可能なPSAを提案

### モジュール改善
- **SwinIR** (ICCVW' 21)
  Swin Transformerに基づいて設計されたRSTBを提案
- **HAT** (CVPR' 23)
  チャネル注意機構とシフトウィンドウベースの自己注意機構を追加したHABを提案
- **DRCT** (CVPRW' 24)
  RRDBを応用したSDRCBを提案
- **MambaSR** (CVPRW' 24)
  MambaとTransformerを組み合わせたSSFormer blockを提案

### Mamba

- **RCAN-it** (arXiv' 22)
  学習設定をアップデート
- **ARFFT** (CVPR' 23)
  受容野を画像全体に拡張
- **DAT** (ICCV' 23)
  空間とチャネルの次元を横断してAttentionを集約
- **DVMSR** (CVPRW' 24)
  ViMと知識蒸留により高速高品質のSRを実現

## GAN

### SRGAN系統
- **SRGAN** (CVPR' 17 )
  **GANを導入**
  生成器としてSRResNetを用いて敵対的学習
- **ESRGAN** (ECCVW' 18 )
  **RRDB**の提案および識別器の変更によりSRGANを強化
- **Real-ESRGAN** (ICCVW' 21)
  実画像の劣化プロセスを模擬したhigh-order degradatcionを導入してESRGANの汎用性を向上

- **EnhanceNet** (ICCV' 17)
  テクスチャ損失と敵対的学習の組み合わせ
- **BSRGAN** (ICCV' 21)
  実画像の劣化プロセスを模擬したshuffle strategyを導入
- **SRFeat** (ECCV' 18)
  画像識別器と特徴識別器を採用

### 高倍率
- **PULSE** (CVPR' 20)
  StyleGANを用いて高倍率でのSRを実現
- **GLEAN** (CVPR' 21)
  StyleGANの利用方法を再考し汎用性向上とコスト削減を達成

### Cycle GANを使用
LR: 低解像度画像
HR: 高解像度画像

- **CinCGAN** (CVPRW' 18)
  ペアデータなしでノイズの多いLRからHRを生成
- **DRN** (CVPR' 20)
  LR→HRの変換とHR→LRへの変換を同時に学習

- **KernelGAN** (NeurIPS' 19 )
  画像固有の縮小カーネルを推定することで教師なし学習で実世界の劣化に対応

## Diffusion

### Diffusionを導入
- **SR3** (TPAMI' 22)
  GANやFlowベースの既存手法を上回る性能を達成

### Stable Diffusion系統
- **LDM** (CVPR' 22)
  Stable Diffusionの元論文拡散過程を潜在空間で行うことで計算効率を大幅に向上
- **StableSR** (arXiv' 23)
  Stable Diffusionにモジュールを追加してSR性能を向上

### 効率化
- **SRDiff** (Neurocomputing' 22)
  残差予測により学習を高速・安定化
- **DiffIR** (ICCV' 23)
  画像復元に必要な情報を集約したIPRを用いて計算量を大幅に削減
- **ResShift** (NeuIPS' 23)
  拡散ステップ数を削減
- **DiffBIR** (arXiv' 23)
  画像復元のためのIRControlNetを提案

- **FeMaSR** (ACM Mlutimedia' 22)
  VQGANで事前学習された特徴コードブックとLRをマッチング

### ControlNetを使用
- **SeeSR** (CVPR' 24)
  出力画像のピクセル単位での整合性を強化
- **PASD** (arXiv' 23)
  テキストプロンプトを学習に利用
- **CoSeR** (CVPR' 24)
  入力画像からStable Diffusionで生成した参照画像を利用

## Flow
- **SRFlow** (ECCV' 20)
  **Flowを導入**
- **HCFlow** (ICCV' 21)
  可逆的なFlowベースのモデル
- **AdFlow** (WACV' 22)
  L1lossの一般化としてFlowを使用

2014 2018 2021 2024