

# Multimediaゼミ 2024年10月22日

- hayami
  - Boosting Neural Representations for Videos with a Conditional Decoder
    - 従来のINRの映像表現能力を高める
    - TAT Block
      - 従来の特徴量と時間埋め込みの融合にはGANに基づくAdaINモジュールを使用
      - INRのoverfittingにおいて正規化は不要
      - GELU → SINE
      - 元のモデル(NeRV)などを参考にしたブロック
      - HNeRVと比べると時刻情報が加わるので精度向上
      - NeRVと比べるとブロックの影響が大きいから精度向上したのでは
- koizumi
  - PaDiM: a Patch Distribution Modeling Framework for Anomaly Detection and Localization
    - memory bankの論文
    - 各パッチごとに対応する特徴量を各層から抽出して、チャンネルごとに連結し、それをN枚ごとにグループ化
    - グループごとにガウスパラメータを取得
    - 学習時に定めたガウスパラメータを用いて、テスト画像がN枚の正常画像が作る分布とどれだけ離れているかマハラノビス距離で計算
      - マハラノビス距離：分布が円になるようにして計算する
    - PCAよりも次元をランダムに減らしたほうがいい
    - データセットの影響も大きい

- tatsumi
  - Idempotent Learned Image Compression with Right-Inverse
  - Idempotent : 冪等性  $f(x) = f(f(x))$
  - 従来手法
    - 可逆変換 ( $D = E^{-1}$ )
    - $x_n = D \circ Q \circ E(x_0)$
  - 提案手法
    - 右可逆 :  $E \circ D(y_i) = y_i$
    - 並列処理ができるようにnon-overlappingにした
    - Blocked convolutionに対して、coupling enhancementとnull-space enhancementを導入
    - 非線形の部分に対して可逆に直す
    - 最初のconvolution layerを右可逆にしないような nearly idempotentを作成できた
- tanaka
  - Implicit Neural Representation Image Codec with Mixed Context for Fast Decoding
    - デコード速度の改善
    - 低解像度の潜在変数にARU、高解像度にARM
    - 高効率のARUブロックと品質・速度バランスのためのARMを含むMARMを提案
      - ARUはチャンネルごとの自己回帰モデルで、低解像度の潜在変数を使用して高解像度の潜在変数のデコードパラメータを予測する
      - ARMは従来手法と同じコンテキストモデル
      - 2つのブロックの比率はMで制御
      - ARU pass1でアンカーを生成、ARU pass2で非アンカーを生成
    - 実験における提案手法2つ

- std : ARMとARUのバランスをとっている
- fast : 速度を重視しており、ARUブロックの比率を増やし、ARMブロックを減らしている
- ARUブロックが多いほどデコード時間が短い