



Uncertainty Aware Retransmission DiffCom

慶應義塾大学 理工学部

大槻研究室 4 年

学籍番号 62204739

名前 小野 拓馬



- セマンティック通信

- ▶ データが持つ**本質的な意味**をエンコーダで抽出して伝送
- ▶ データ量や通信帯域のリソースの最適化が可能

- Deep JSCC (深層共同ソースチャネル符号化) [1]

- ▶ 画像ピクセルを複素チャネルシンボルに直接写像し、**符号化と変調を深層学習で一括最適化**する手法
- ▶ エンコーダとして採用

JSCC : Joint Source-Channel Coding
SNR : Signal-to-Noise Ratio

- 画像伝送と拡散モデルの組み合わせ

- ▶ 拡散モデル : 画像生成AI
 - 順方向過程 : 画像にガウスノイズを付加
 - 逆拡散過程 : ノイズ除去し, 画像を生成
- ▶ 強力な生成能力を持つ**拡散モデルをチャネルノイズの除去に用いる**ことで低SNR 環境下でも写実的な復元が可能

[1] E. Bourtsoulatze et al. *IEEE Transactions on Communications and Networking* 2019

関連研究 (1/2)



- DiffCom : 受信信号をガイダンスとして拡散モデルで画像生成 [2]
- 送信信号をJSCCエンコーダで圧縮し送信
- 受信信号をデコーダに通し、拡散モデルで受信信号でガイダンスし元画像を復元

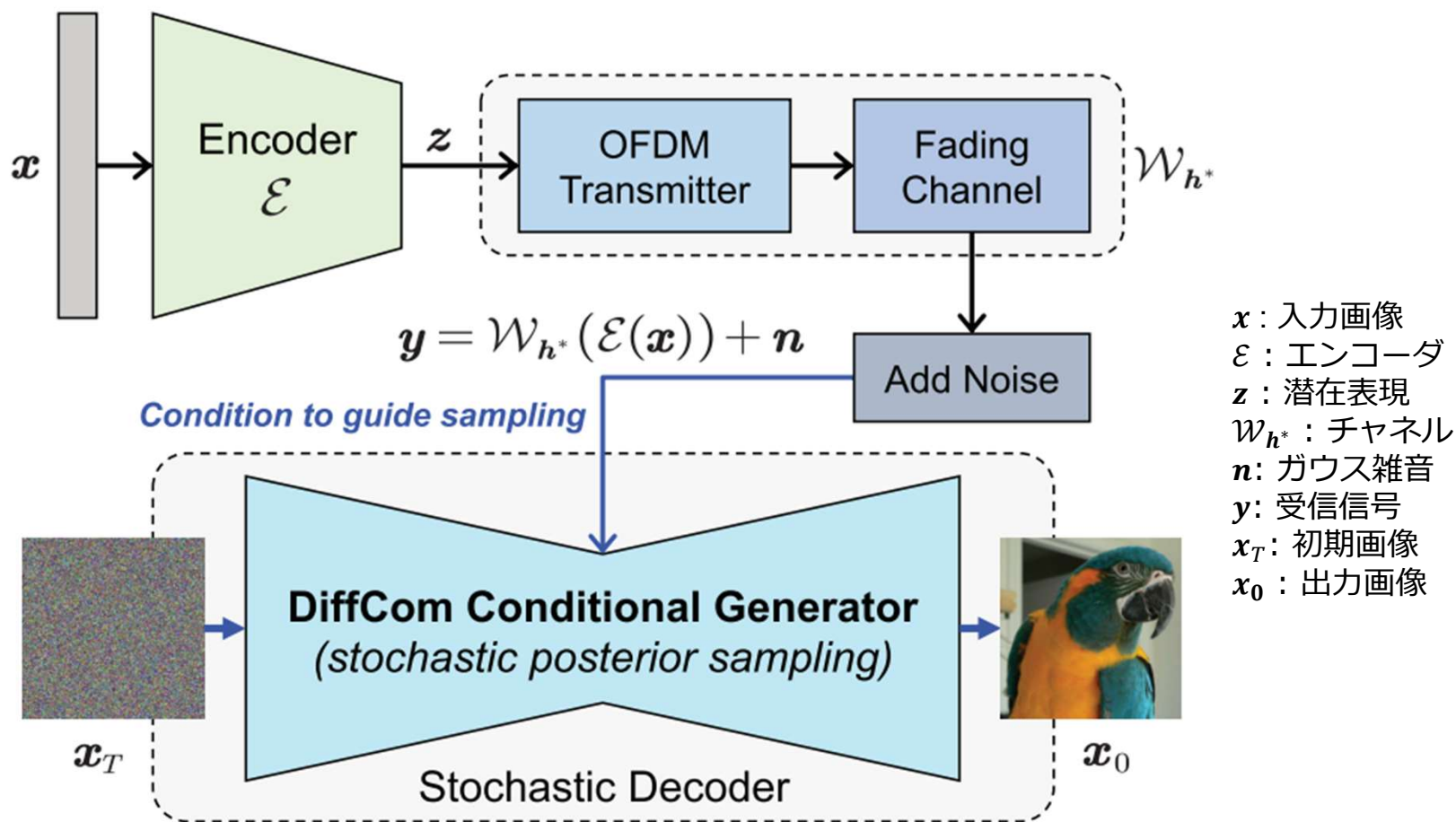


図 1. DiffCom のシステムモデル [2] [2] S. Wang et al, *IEEE Journal on Selected Areas in Communications* 2025



- 拡散モデルの不確実性 [3]

- ▶ 不確実性：初期ノイズや逆拡散過程に加えるノイズによって同じ入力やプロンプトでも出力される画像が異なる
- ▶ この関連研究では不確実性をピクセル単位で評価

- 不確実性の推定メカニズム

- ▶ 摂動：逆拡散過程の途中状態から推定されたクリーン画像に対し、再度ノイズを加え、サンプルを生成
- ▶ 摂動を複数回行い、作成したサンプルに対する拡散モデルの出力の分散を計算。この分散を不確実性と定義
- ▶ ピクセル単位で不確実性を測るので画像の構造重視で送信画像を復元するという目的には不向き

[3] M. De Vita et al, *IEEE/CVF WACV* 2025



- 受信側でDiffComを用いた関連研究の課題

- ▶ 拡散モデルが生み出す**幻覚**

- 受信側の拡散モデルが周囲との整合性を重視し、元の送信画像にない物体や文字を創造

- 不確実性を用いた関連研究の課題

- ▶ 画像生成の拡散モデルに使われており、ピクセル単位での測定
 - ▶ 不確実性が通信と拡散モデルの文脈で使われていない

- 研究の方針

- ▶ DiffCom のシステムモデルに**再送の仕組み**を追加
 - ▶ 不確実性を用いて拡散モデルが幻覚を生み出す箇所を構造単位で特定
 - ▶ 意味的に重要な（不確実性の高い）箇所のみを選択的に再送



- 不確実性の推定メカニズム

- ▶ 摂動 : 現在のステップから推定されたクリーン画像に対し、再度ノイズを加え、サンプルを生成
- ▶ 摂動を複数回行い、作成したサンプルに対する拡散モデルの出力の分散を計算。この分散を不確実性と定義

- 画像の構造的な不確実性の定義

- ▶ 画像の構造的な不確実性を捉える
 - 平均プーリングでダウンサンプリング
 - 線形補間でアップサンプリング
- ▶ ピクセル単位での誤差に注目するのではなく、画像の構造重視 (物体の位置、文字)



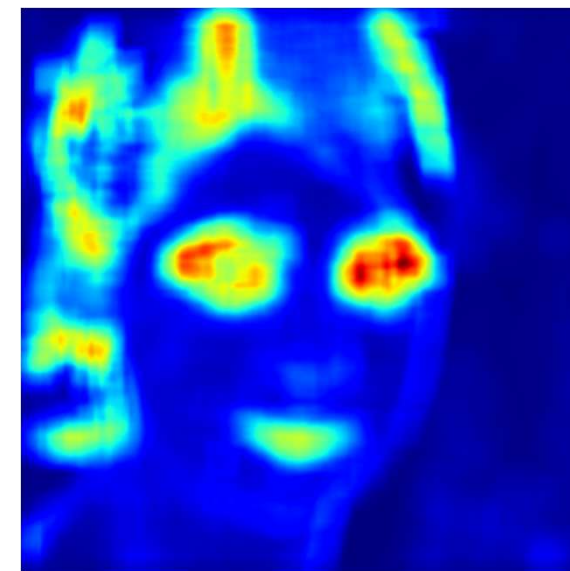
- 不確実性が高い潜在表現の箇所を再送要求
 - ▶ 受信側で不確実性マップ図2 (c) を作成
 - ▶ 不確実性が高い領域(マスク)を送信側にフィードバック
 - ▶ 送信側は該当する特徴量のみを再送
 - ▶ 幻覚の部分を効果的に修復可能



(a) Ground Truth



(b) Reconstructed Image



(c) Uncertainty Heatmap

図 2. 不確実性測定の例



- 再送データの効率的な伝送

- ▶ 不確実性の高いマスクに対応する潜在表現の部分的な送信
 - 送信データの限定による通信量の大幅な抑制
 - データ量の少なさを活かした、十分な電力割り当て等による理想的(ノイズレス)な伝送過程

- 信号の置換と再サンプリング

- ▶ 初回受信信号 y の内、マスクされた高不確実領域の再送データへの置換

$$y_{final} = y \odot (1 - mask) + y_{retrans} \odot mask$$

- ▶ 修正後の信号をガイダンスとした逆拡散過程による画像の再構成
- ▶ 正解信号に近いガイダンスの付与による修正箇所以外の領域も含めた画像全体の復元精度の向上

y : 初回の受信信号
 $y_{retrans}$: 再送受信信号
 $mask$: 不確実性に基づくマスク



- データセット：人間の顔画像を集めた FFHQ [4]

- ▶ このうち100 枚の画像を本システムの評価に使用

- ▶ 拡散モデル

- FFHQ データセットで事前学習された DDPM [5]

- ▶ エンコーダ

- Deep JSCC : RGB画像 $3 \times 256 \times 256$ を $2 \times 64 \times 64$ に圧縮 [1]

- FFHQ で訓練

FFHQ : Flickr-Faces-HQ

DDPM : Denoising Diffusion Probabilistic Models

JSCC : Joint Source-Channel Coding

AWGN : Additive White Gaussian Noise

SNR : Signal-to-Noise Ratio

PSNR : Peak Signal-to-Noise Ratio

LPIPS : Learned Perceptual Image Patch Similarity

[1] E. Boursoulatz et al. *IEEE Transactions on Communications and Networking* 2019

[4] T.Karras et al, *IEEE/CVF CVPR* 2019

[5] J. Choi et al, *IEEE/CVF ICCV* 2021

- チャネル設定

- ▶ AWGN チャネル : SNR 0 dB

- 評価指標

- ▶ PSNR : 画像のピクセル単位での評価

- ▶ LPIPS : 人間の知覚に基づく深層学習ベースの評価

実験結果：不確実性上位10%を再送



- 再送前後での評価指標の変化

- ▶ PSNR +0.1301 dB
- ▶ LPIPS -0.0001
- ▶ 画像全体の評価指標の変化は微小

- 局所的には改善

- ▶ 帽子の文字が明確化
- ▶ 服の色の変化が明確化
- ▶ 肌の滑らかさが上昇



図 3. 送信画像



図 4. 再送前の再構成画像

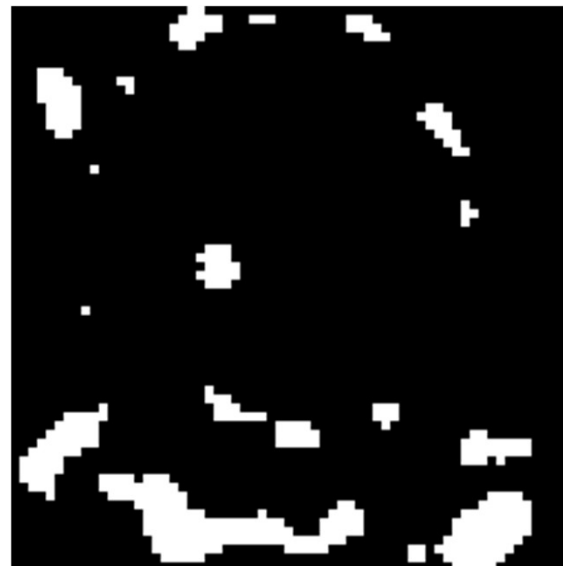


図 5. 不確実性上位10%



図 6. 再送後の再構成画像

PSNR : Peak Signal-to-Noise Ratio

LPIPS : Learned Perceptual Image Patch Similarity



- 関連研究の問題点

- ▶ 受信側で DiffCom を用いた関連研究の課題
 - 受信側の拡散モデルが周囲との整合性を重視し、幻覚を生成
- ▶ 不確実性を用いた関連研究の課題
 - 画像生成の拡散モデルに使われており、ピクセル単位での測定

PSNR : Peak Signal-to-Noise Ratio

LPIPS : Learned Perceptual Image Patch Similarity

- 研究の方針

- ▶ 不確実性を用いて拡散モデルが幻覚を生み出す箇所を特定
- ▶ 送信側にフィードバックし再送することで幻覚を防止

- 今後の課題

- ▶ 現在は再送の信号が理想的にノイズなしで送信可能と仮定
 - 同一ノイズ環境下での再送シミュレーション
- ▶ 画像の PSNR や LPIPS 改善が微小 (PSNR +0.1301 dB, LPIPS -0.0001)
 - 不確性の定義やガイダンスの強度を変更して画像全体の復元精度の上昇



- [1] E. Bourtsoulatze, D. Burth Kurka and D. Gündüz, "Deep Joint Source-Channel Coding for Wireless Image Transmission," in IEEE Transactions on Cognitive Communications and Networking, vol. 5, no. 3, pp. 567-579, Sept. 2019, doi: 10.1109/TCCN.2019.2919300.
- [2] S. Wang, J. Dai, K. Tan, X. Qin, K. Niu and P. Zhang, "DiffCom: Channel Received Signal Is a Natural Condition to Guide Diffusion Posterior Sampling," in IEEE Journal on Selected Areas in Communications, vol. 43, no. 7, pp. 2651-2666, July 2025, doi: 10.1109/JSAC.2025.3559158.
- [3] M. De Vita and V. Belagiannis, "Diffusion Model Guided Sampling with Pixel-Wise Aleatoric Uncertainty Estimation," 2025 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), Tucson, AZ, USA, 2025, pp. 3844-3854, doi: 10.1109/WACV61041.2025.00378.



- [4] T. Karras, S. Laine, and T. Aila, “A style-based generator architecture for generative adversarial networks,” in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2019, pp. 4401–4410.
- [5] J. Choi, S. Kim, Y. Jeong, Y. Gwon, and S. Yoon, “ILVR: Conditioning method for denoising diffusion probabilistic models,” in Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV), Oct. 2021, pp. 14347–14356.