

# 2024/08/27

- Jiu

## ControlNeXt: Powerful and Efficient Control for Image and Video Generation

<https://arxiv.org/pdf/2408.06070>

ControlNetをはじめとするcontrollable generation methodsはプロンプトを入力とするbase model (e.g. stable diffusion)に対して、ほぼ同サイズのconditionを与えるcontrol branchを追加

- 計算コストが高い
- 学習に時間がかかる
- base modelのパラメータは固定

ControlNeXtではbase modelの一部の重みと軽量のcontrol branchを学習 → 学習効率の向上

また、Zero Convolutionの代わりにCross Normalizationを使用 → 学習の安定と高速な収束

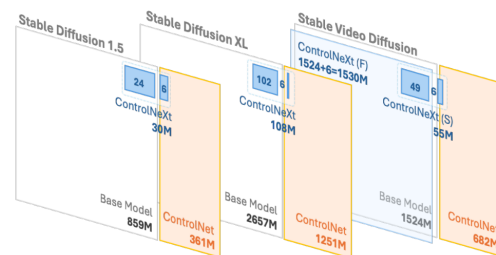


Figure 4. Parameter efficiency of ControlNeXt. We present the number of learnable parameters with various base models.

- haoyuan

## YOLOv10: Real-Time End-to-End Object Detection

<https://arxiv.org/pdf/2405.14458>

(NeurIPS 2024 多分通る)

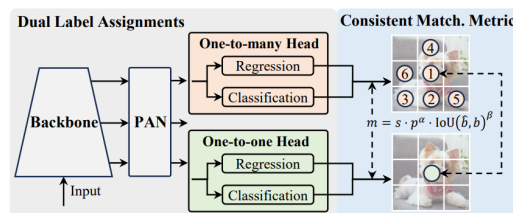
補足: <https://arxiv.org/pdf/2012.03544>

物体検出手法のYOLOシリーズは後処理のNMSによりend-to-endでの学習ができない

YOLOv10ではone-to-many headとone-to-one headを使用し、NMSを排除

学習時は二つのheadを学習、推論時はone-to-one headのみ使用

Consistent Matching Metricにより二つのheadの整合性を取る



- Jin

## FreeU: Free Lunch in Diffusion U-Net

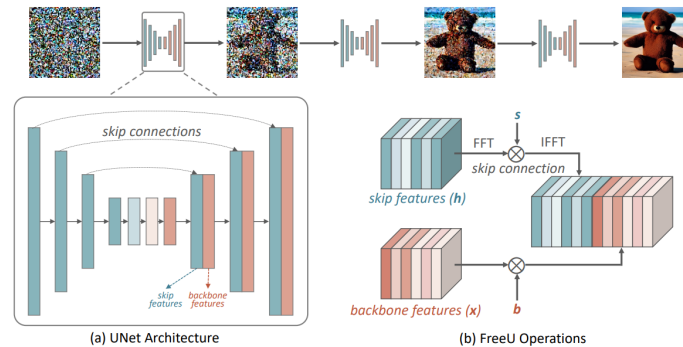
<https://arxiv.org/pdf/2309.11497>

CVPR 2024

denoising processにおいて画像のhigh-frequencyとlow-frequencyに分類すると、LFの変化は緩やかな一方、HFの変化は速い

U-Netはbackboneとskip featuresからなり、backbone featuresがlow-frequency、skip featuresがhigh-frequencyの情報を持つ

backboneとskip featuresにそれぞれ係数を導入することで生成画像の品質を改善



- fujinami

## Adding Conditional Control to Text-to-Image Diffusion Models

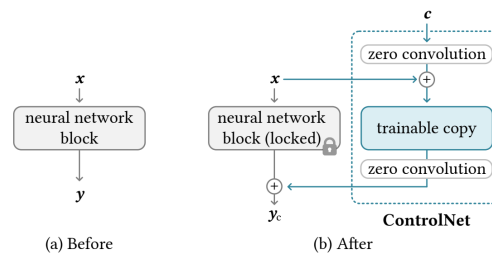
<https://arxiv.org/pdf/2302.05543>

[https://openaccess.thecvf.com/content/ICCV2023/papers/Zhang\\_Adding\\_Conditional\\_Control\\_to\\_Text-to-Image\\_Diffusion\\_Models\\_ICCV\\_2023\\_paper.pdf](https://openaccess.thecvf.com/content/ICCV2023/papers/Zhang_Adding_Conditional_Control_to_Text-to-Image_Diffusion_Models_ICCV_2023_paper.pdf)

ICCV 2023

text-to-imageのbase modelにcondition (sketch, depth map, edge...) を加えることで生成画像を制御するControlNetを提案  
base modelのパラメタは固定

condition branchではzero convolutionを挿入することで学習が安定



- nutaba

## FedAS: Bridging Inconsistency in Personalized Federated Learning

[https://openaccess.thecvf.com/content/CVPR2024/papers/Yang\\_FedAS\\_Bridging\\_Inconsistency\\_in\\_Personalized\\_Federated\\_Learning\\_CVPR\\_2024](https://openaccess.thecvf.com/content/CVPR2024/papers/Yang_FedAS_Bridging_Inconsistency_in_Personalized_Federated_Learning_CVPR_2024)

Federated Learning: 分散学習の一種、データではなくモデルのパラメタをやり取り (プライバシー配慮?)

FedAvg: Federated Learningの代表的なアルゴリズム

- global modelとlocal modelの一貫性の欠如 (1)
- 学習への参加率が低いクライアントによる学習の阻害 (2)

Parameter Alignmentにより (1) を解消

- パラメタ更新時にlocalにglobalを取り入れ整合性をとる

Client Synchronizationにより (2) を解消

- 各クライアントに重みづけ、参加率の低いクライアントの重みを小さく

