Paper summary

AI VISION Lab

1. 공부한 논문의 제목, 게재된 학회 혹은 저널 등 논문 기본 정보를 적으세요.
   - **A. 이름: DenoiseRep: Denoising Model for Representation Learning**
   - **B. 저널: NeurIPS**
   - **C. 도메인: Representative**
   - **D. 출판연도: 2024**
   - **E. 저자: Zhengrui Xu, Guan'an Wang, Xiaowen Huang, Jitao Sang**

2. 논문에서 제안한 알고리즘 및 프레임워크에 대해 본인이 이해한대로 다이어그램을 그려보세요. **논문 Figure를 그대로 따라 그리면 안됩니다.**
   A. **Feature을 잘 추출하는 것은 Discriminative Model에게 있어서 매우 중요한 업무라고 할 수 있음.**

   B. 이를 효과적으로 추출하기 위해 우수한 성능의 모델인 Denoise Model의 구조 (Diffusion)를 적용하고자 함.

   C. 일반적으로 사용되는 Feature Extract backbone (예: Res50, ViT, CNN)에 우수한 성능을 보인 Diffusion과 같이 Noise를 주입하고 이를 제거하는 과정을 거침으로써, Robust 한 Feature를 추출할 수 있다는 판단을 함. (가설)

   D. 그러나, 매번 추출 후, Forward – Reverse과정을 거치는 과정이 필요하고 이는 모델을 여러 번 호출하게 만드는 Latency문제를 야기함.

   E. 따라서, Embedding (Backbone)과정과 Denoising (Diffusion) 과정을 한 step만에 해결할 수 있는 아키텍처를 제안함. (이때, 연산량 + Parameter의 증가를 해결할 수 있는 구조를 제안)

   F. 일단, Backbone에 해당하는 부분의 경우 Train하지 않고 본인의 업무를 하도록 둠.

   G. Denoise 부분에 대해서만 Train하도록 함 (L2 Loss, Embed Layer의 Output을 Input으로 사용하며, Input에 대해 Forward Process – Reverse Process, 이

때 Noise를 Predict하는 식으로 학습. 해당 구조는 Linear한 구조로, 한 번에 1개의 Denoise step이 아닌, 2-n개씩 처리가능한 확장이 가능한 구조임. 이 때, Noise는 20회 정도 주입한다고 되어있음.)

3. 본인이 생각하는 이 논문의 장점이 무엇이라고 생각하나요? **논문 Contribution bullet을 그대로 따라 적으면 안됩니다.**
   A. 해당 부분을 적용하는데 난이도가 상당히 쉽다는 것이 가장 큰 장점

   B. 앞서 보인 아키텍처를 통해 얻은 Robust Feature을 통해 같은 Task를 진행했을 때 성능의 향상을 보였다는 점은 고무적임.

4. 이 논문을 읽으면서 느낀 점, 혹은 배운 점이 있으면 적어보세요.
   A. 이 논문을 계기로 Representation Generation Method의 다양한 논문을 Review하게 될 예정임.

5. 이 논문의 한계점이 있다면 무엇이라고 생각하나요?
   A. 다만, Diffusion을 feature에도 적용한다? 연산이 추가적으로 필요한 구조라는 단점을 가지고 있음

6. 본인의 연구에 접목시켜볼 점이 있을지 생각하고 적어보세요.
   A. 다양한 Modality를 사용하는 상황에서 이를 Ablation 실험에 시도하는 것은 좋다고 생각됨. 그러나 더 많이 고려해볼 필요가 있다고 생각함.

7. 본 Summary를 작성하는 과정에서 생성형AI를 사용했나요?
   A. 아니요

날짜: 2025-08-17

이름: 신준원

DeniseRep : Denoising model for Representation Learning

Introduction

denoising model ⇒ discriminative task에도 효과적임.

↳ Representation learning 에 활용.

⇒ 왜 중요한 것인가?

⇒ data의 representation을 잘 학습한 경우 → 정보 추출을 효과적으로 할 수 있음.

↳ classifier or other predictor 등

따라서, 본 논문에서는 denoising을 활용한 효과적인 representation task 학습 방법을 제안함.

↳ feature discrimination

① joint feature extraction & denoising

Generative model (with denoising)

DDPM : $P(X, Y)$

Sample Condition

Generative vs Discriminative

Denoising 학습 ⇒ Generative model > Discriminative model

Discriminative model : $P(Y|X)$

↓

label, etc

↳ 위한 diffusion model : ① DiffusionDet : Detfusion + Object detection (Stable Diffusion (Pretrained))

② DiffSeg : Diffusion + Image segmentation

학습과정 ⇒ ex) noise box, noise segmentation

⇒ learning representation ⇒ 학습한 정보추출에 활용

↳ task에 활용 예정 (with Denoising model)

task ⇒ ReID → 위 feature 활용.

↳ Cloud 도움

Denoise Rep

→ Representation learning. → 모델이 특징을 잘 학습하는가가 중요.

① discriminative model's backbone (Vision) : ResNet, ViT

→ feature extraction → Denoising 적용 → Introduced.

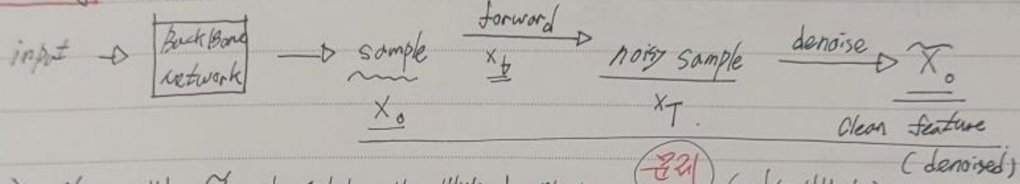
⇓ 성능 우수한 feature extraction 가능

② specific head : MLP for classification

RCNN for object detection

FCN for segmentation

Denoising (from DDPM)



⇒ x_0 대비 $\tilde{x}_0$ 의 특징에 대한 변별력이 뛰어남. (문제) (다수 반복 필요)

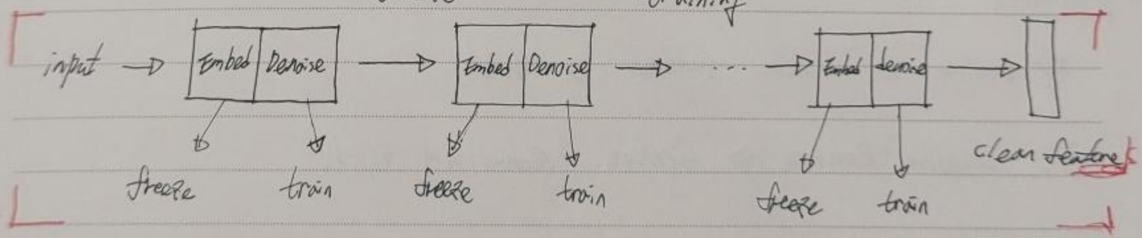
Denoise T step

⇒ denoise step 다수 필요.

결론, Backbone + denoising layer fusion.

⇒ 2Branch 구성 (original embedding layer, denoising layer)

freeze training



input → task 반복 / Denoise) → 단순한 fusion을 통해 → 한번의 task로 결합.

feature extraction & Denoising

~~input x를 Diffusion Network에 X.~~

$\frac{x_0}{4}$

⇒ feature (from Backbone)를 input으로 사용하는 것. ▷ LDM 논문 내용.

feature의 Denoising ⇒ ~~noise~~ robust feature를 위한 것

~~목적~~ ⇒ ~~noise~~ adapted 학습시 ▷ noise Diffusion ▷ 학습 (recursive!)
Backbone ▷ Denoising layer 학습목적.

해결방법 ⇒ Backbone에 Denoise layer를 free.

Step: 여기서 inference step.

⇒ feature 추출 ▷ 학습된 Diffusion model 활용목적.

▷ Denoising layer into next embedded layer.

↳ $y = Wx + b$.

Experiments

only Backbone VS Denoise Backbone ⇒ 성능향상

즉, Denoise Backbone ⇒ feature (Robust)하게 학습됨