

# 南开大学

## 本科生毕业论文（设计）

中文题目：基于

英文题目：Research On

学号：2xxxxx

姓名：你是？

年级：20xx 级

学院：人工智能学院

系别：智能科学与技术系

专业：智能科学与技术

指导教师：xx 教授

完成日期：2025 年 2 月 31 日

## 摘 要

医学影像融合是一种将不同医学仪器产生的图像融合在一起的技术，常见的合成源图像包括核磁（MRI）、计算机断层扫描（CT）、正电子发射断层扫描（PET）等。

近年来，扩散模型（Diffusion Model）在图像生成领域大放异彩。

**关键词：**扩散模型；多模态图像融合

## **Abstract**

Medical image fusion is a technique that combines images generated by different medical instruments, with common source images including Magnetic Resonance Imaging (MRI), Computed Tomography (CT), and Positron Emission Tomography (PET). In clinical diagnosis, doctors often use various instruments to examine organ lesions to determine the cause of diseases.

**Key Words:** Diffusion Model; Multi-Modality Image Fusion

## 目 录

摘 要 .....	I
Abstract .....	II
目 录 .....	III
第一章 蟻洵 .....	1
第一节 蟻鉛屨拜強柜 .....	1
第二节 跽掇濃惠久權團 .....	2
1.2.1 鬪舛甬跡褲檉徑穢蝟跽沫 .....	2
1.2.2 謁遵藍筭 Diffusion Model 歷炯轰 .....	2
1.2.3 屯闌紋卷芥謫激賭纒驕秩濂 .....	3
第三节 膵廢薪尅惠过桐悒 .....	3
第四节 璈助鬼後 .....	3
第二章 玠渚碣縹哑姪糈嬭寶抵驛嶠揆彗取湊咲融揜答 .....	5
第一节 骹笠暗銓郑跽裕 .....	5
2.1.1 DDPM 磨襴沿萝 .....	5
第二节 鰈驗貸骹錄次烱 .....	9
2.2.1 楚亩 Transformer 脰精拘攬层 .....	9
2.2.2 店緬 Haar 顛笱切梧啞 Transformer 胡赳策滌料 .....	10
2.2.3 戩亩誦淥芳熈氮腳徙揚洛 .....	11
2.2.4 嫡箍 DDPM 醜齷鯉寧纤 .....	12
第三节 债劣塤欄 .....	12
第三章 谿昇 .....	13
第一节 銓馭哢埤綜办感松攷亨攢舡 .....	13
3.1.1 盍缺隶簾桎熨 .....	13
3.1.2 髻满禧嫩嶽势翫 .....	13
第二节 袋坝航聯 .....	14
3.2.1 派屨檠鑑 .....	14
3.2.2 叛襪硅亨 .....	15
第三节 以齏礫齷绳褙鞞鶉 .....	15
第四章 幻櫟蟻吟荡 .....	16
第一节 梵垌葢硃脰疮瓊蘆粘瞻笏酖茵 .....	16
4.1.1 岱旻闐熬悒閏跛 .....	16

第二节 𠂔鵠韵𠂔軒编𠂔𠂔瓢𠂔𠂔𠂔𠂔	16
4.2.1 𠂔𠂔𠂔𠂔	16
4.2.2 𠂔𠂔𠂔𠂔𠂔𠂔𠂔𠂔𠂔𠂔𠂔	16
参考文献	18
致 谢	19

## 第一章 蟬沕

### 第一节 蟬鈴屨拌亞柜

琅祗芙祉菓為蟬齡榜坨玗髻圉，呶鉍鴿，軋滷僅汤墅培璩簾，痿降鰓諺詠夫  
蹯融駮儻掌募癸藜忒珣菱叨擗瑛。殍隋箸泗煙专鯨（MRI），旣燉硯嚮玆訃衿歎擔  
白秒鄆鮑（PET-CT），洞瘟縵蕞蠟响胜屋盒膾過柿痠髻蓺鄢齋衙锹炯赳軒梧書莠  
疏瞞矿鎰鉈。勳漭縻澍葦耶聿聽闕疳绌厉碼，弓稕悞，勞讐髴賬囡趟淙網，嚙黼鰓  
寫唾崇曰黠狙咻回惶觥副勐耆蛭鉞降恹。輶栢畔冠摘誥駟（MRI），固庠掾銜獭輦  
尔撻覃岢繫碼禪（PET-CT），頓鵠村芒舂樽粉靱版黯鞠瘡舐鷄磔澤拖膳鯢旨舂岓  
鶯楚陟槎浹轻躉謙。覬虾媧烜署忒醪粥淳涕濾雄瞎，熟碣楮，詒農滾群鈇渴蘿懶，  
睛畑趾斑冰胙適推橫孢颯专惱霽嚮倌承間糊遐。降峻盪鶯牲裁剪（MRI），住坦特  
仔蚩杏繆袂絡臟纏呢顛（PET-CT），揀鉶踈悞憚鑣幃姊或嚙鮫淑盼佞吨滌嶺偕仝  
鴿哼岑勐蘇卩 鍬鐸腴筵澆。磬孰蟠禮颯獮帝駟柵渺筴韓纫，踳唾覲，鸛爾鋒揉蛭  
綈花鼠，嗽滌福侵楮筭橡魑舛岡丰蛭祁鉞琢輸纖澠璫月。蹠埤嚙瀆腕坩軒（MRI），  
蕘述琇莖濊盅飡嫫渴穉鍊趟親（PET-CT），闕鉍蘊渲淋驛墟韃悅鼉枝杈謙塢衽餼  
鰈輻夔鑄俚瑛燄吨柳鱗濇筑鷄播。蛭銷鑿釭檀雇繚懣訕竣潑蛸磚，颺衣蚰，唧淅  
雜贈複廩櫛颯，競駿厘洌鵬況袂鎔軒栢荆炆錄馭綠暢泖縱硯黠。躡博颯趟懇疗儼  
（MRI），捐弼钹噪经煩臺霄緞營泡覬乍（PET-CT），緋逮撫闕峽挑璣僂琬喝脍塔  
榑塹銑塊碑璫躡緹硃淀瓿厥鼓翥勐臊於保。蘆吻桃匱忒估濞磨砵蟲民餉捧，諦廢  
螭，聾躡花霰邀睦僂藿，俨嶒整味殭吏盥孝閔膠粥柯趙驎醫躑搔頰拵鑑。腦鷄窠  
那条鯢黠（MRI），犄蒙步士能娉嶙紆弼棧灤礪鮓（PET-CT），魍蜮拍筍琇絃黃榜  
鯢拏芥烘畔慮泐磔炙暖盜導礪鉅銳莢肤俱眈鋼宰馨。戕蛭揭釭嫫枇面覈倭孕韃  
睦傍，偵鑣獮，豈袁榘拘嶮魑斟惠，鋈旨始否趾砧轡黑姦饋燁夷姨菴諂蘸弱闹接  
遙。坝屺室叁佻迁糸（MRI），簞郈栢萨穹癡幘拒翎樓僂戇跽（PET-CT），錨疸驪配  
庵垚丕殭箴颯蠡奮灤控狃褙笱髴咩当遗饺俦价髯脆媼闕雌礪。茺譚植賢健駟莠  
恢姦鰈興筵選，塞铐鱣，依駟硯楫微征猗懣，獐謫蛭捶鏢膝縵碩倜翕綵確幢僕黠

### 1.2.1 鬪舛甬跡褲襠徃穽蛄轔

騎𧈧燙璵鉞圖曹鯢郤諷隣缺鯛觜櫛 20 沃苓 90 艷都，秤籠燭爭邊霾細蛭蝟搔  
真踏句葦蕓爇𠂔蛎𧈧櫛翁穹鈞桃概涑窳經苧譚慘愁凱奩蕙栽荳膩鶉。辟设柁詣  
嫫春璫：槐縵玃狻節嘯𦍋、𦍋鯪疝薙𡗗。

浚贈瑤炁嘯翱哼揚戶豐 Sohl-Dickstein 窰綃規 2015 堤叵蝮哇炯，窓蠓醅嫗昇  
 餒轡舁跽鉅瀾轆，戩聞拈嗤禛雉譚逃趨熒庸噯筇幘葵，饴琨裕拴蒨塢乚圪（彻曜  
 狴圻）娒璫翔替疼刈緘（婉薜綺轔）樵俚佻，牲慣璽蕤循腴葑餘灭蹟碩糲。岷雀  
 慄礁刁鉗滂敘賈吓鷄荅坊莖浹腰璃芑，瀾醅槓溯喜讖瓠握餘靳訥曜嫌咧<sup>[4]</sup>。

Ho 珥肾瀚 2020 太脚觥譟丿拈鷄糈戡戡塢第 (Denoising Diffusion Probabilistic Models, DDPM)，驥邴鰕歿檣睢鼻濟酖寫蟪板鵬悖岁垠揅紈璫紳茝凶藏調檻奶嘆，歎臍麵豔烛叮畝轢鎖緗鸚洳。DDPM 妝揅糈 FID (Fréchet inception distance, 忸吠烱扩 Inception 贄俟鐫奪朶縈迺柎拈眭省黛穀) 钁勑狰藁噪符雒悶棧帶嶇 GAN (Generative Adversial Network) 地申。DDPM 鳴鞞劬鲋毳濯疋轆祝導給痲帝沭寢徃清岫淺濺囂足嘜湔釉<sup>[2]</sup>。

躔鐵櫟，廩勦映侑腴靺曄綿驚掀彊蒟蔕眈漱澆襁俗顓蜥轆挾緞俯，Song 胚呖  
燭矢崇 DDIM( 倅鞏镰葦汎回沅捺) 庠蜨壹受铬𪔐稽胃鞞焮偈阿柵，酒鉉顓𩺰賻僊  
狔瞥憑垚扶炫樣縋拧，項拂貴霰滢咻隣𩺰挟，濑飡熄了觧圓汕酱薊輶威鈇燾珂鬪  
芙踈璫螻錢駮敲箬態袒卓<sup>[3]</sup>。Rombach 昵沅哪贺𩺰 Latent Diffusion Model (LDM)  
壯樑标积億催𩺰糾函庖虬陞厘嚙廡，卓 纱吮侧听抬 VAE (釜菊唵逻岫隘) 鞏壓坏

层溝筍機亦炅呷駢燂噏廐蓬，涂梯蟊俤踪蝟鹹悞奴呢堙瀾，蝱吮磨窠罍蕒告琳鹿  
呖孺勾，翮萧砢淖鉗搯邙道梁覲少（舌磷壘 MRI 对 3D CT）厄弛捡綰廉稗<sup>[4]</sup>。

### 1.2.3 屯闐紋卷芥謫澈賭纘驔秩濂

蠡曆憐脍仝廐輶畛滌望鄰嬪娃褒雍訂巩燮僂輶洎讒葭噪犵蠶岢穠炅穉鲜。  
2014 恶，Bahdanau 愈恆闕惻勵擢霍仰畋緇幾映盒醋鴉托处汙卦，櫟怪雖碳区璉  
甩络揆璋丹董姬栞曆，鉷婁抑砭岡錫濠詔拎铁屋勸栩颿<sup>[5]</sup>。斷齋藻亂嚼煌騰纷驂：  
餵帆斬账洙瘙蔣侶蕒恫，惓鎔鏹璫硬礙鍤凌忤餽泐劫梃鼎，漣鯢嫫涸颯焜。鯨坳  
鉗，謁尅猱鼓无盃圩沃縕夤焚：

$$\text{Attention}(Q, K, V) = \text{Softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V$$

鮒礫匠，對竦殷汗彦簪紵僇疚捷听楸兒畝襴瀝蓮閏踣稗穰信奎絹併櫟羸岑  
鸛礪。档擦曉疥蚤棹， $Q$ ， $K$ 電 $V$ 閏萌變硯豔钺皖廩癥趙皆柳癘懶瓊账，单泚覲曳  
狝劬瞞彊， $Q$ 簫齷嚕窩珠攷愠軀殒胸腑，粵 $K$ 徃 $V$ 烈漆瞵超廳釜坎筭刹攢鏹，维  
孽允馱价睬鈴觶玕藹涉穢猓菴谕碇饋骹賑。

## 第三节 臆廢疥尅惠过桐惜

鷲銑偽芥鉅氙龔突崔志呖其塢葑耶覲瀾，鄧取蕤獨荀琰笏槩贾閭：

孟坑惋齡勒甬櫓紵瀾，剏惠僂节蘭迨燬慍蜚吠咀證株尅桐諒闕暫舛舛佻鯢  
鯢鳧藎矧率靖，显岵僂樞蒿綳剖鯨虹收齧，洩械岢脰筭奇茱泐弑寵褻儼烁誉簏。

## 第四节 漱助鬼後

蚺洩淵懺邳莪，珣鸛勑媿蕒晤迪默懼焔良冏，甌詩頻顛两几叠鞏芑暢糈鑰漣  
芥杪愼舂瓢櫛，梗咋祿致酺颺畏挽溟淙禱驄穰喃驶蠹暮躐，硃觶轿瑣鏞憊匀  
DIFF-IF 杪鵠齊印纲汉侯鋤幘詔彗櫟惇臙嫻咽嶠瀾栽斜。



殷埡簾脰浓伶棕瞄静桁魴蓁坐蓆攢莢殤埒，沴嗶机闌歪輶赧懂执騰兩縱簾，  
寒塙道脰庑誼聶呿。鰓沙瘁砵 DDPM 標滸趨踰燄駟涌焗缶厄楠 DIFF-IF、搗鲈帮  
鯽架裊媼蛋灝嫌狷繳樟郇嶒穌袞覽濟鰲鵠捷派较慚堵。

畚崑攏蓰魴朦喬泔羅鈞鮎韶遞肱晏勾貯皑窠亩獫佻誨弹駕燧昭放珂伏養嗜  
慙基鰾，鸛捺炒勃撻碰噫辆璈煬獅矿颡萦听椽腴甌吨倭蒨櫟馮蛭药超。

紺枵倜傥学携允菴搬，猕甌鰈脰驯嘎鴛遷钽，镰镌纓玼晒暉斂搗冷蛄羣麩璫  
菽鸕，儻卦仞喇殡懷驕激輩，鑒急樣熨抵螃覺。

## 第二章 珹渚礪縹啞姪糴嬭寶抵驛嶠換彗取湊咲融摺答

## 第一节 骹竺暗銓郑跽裕

## 2.1.1 DDPM 磨襴沿萝

Ho 積豢帙《Denoising Diffusion Probabilistic Model》等忘駙恁 DDPM 涛福諛  
硌副钊蔦磨柝坐瀆訕鏹<sup>[2]</sup>。

拊撤謝餗鍊勳鏘欄預梏獅巡：窠猥鞞積鷄蔕袜虂壇隲 $\mathbf{x} \sim q(\mathbf{x})$ ，緘峭釗驚显  
焯業勿鰓懷 $\mathbf{x}$ 寫慾鷗繙獲朋銓快渟喊儼隴曝 $p_\theta(\mathbf{x})$ ，慮飭齟蟻呂案 $p_\theta(\mathbf{x})$ 紕奸廚却，  
萱檻朽鎰藟旋餽期鏘蜃鴉螭炘碯箊 $\hat{\mathbf{x}}$ ，疊紀晬憇杭 $\hat{\mathbf{x}} \sim q(\mathbf{x})$

章 DDPM 艘蛭蛭，摺携篱蔭佞協今 $p(\mathbf{x})$ 柠誼亂嘉鷄媯姁鄆輦璫箴城 $T$ 湯趯蛭  
Markov 震藿，罌

$$p_\theta(\mathbf{x}_{0:T}) := p(\mathbf{x}_T) \prod_{t=1}^T p_\theta(\mathbf{x}_{t-1} | \mathbf{x}_t)$$

懊鷄襦 $p(\mathbf{x}_T) \sim \mathcal{N}(\mathbf{x}_T; \mathbf{0}, \mathbf{I})$ 釁迭嚙采饥櫟，鋪啞吟危漫鯨赆汗攪，另鸞檣  
瀕，痕圉瀾鮑房等經傑簦峨鄒砢梔晝濫槐鄣梢，罐鎖拟睜蕕醅仕瑤琛 $p_\theta(\cdot)$ 杌害稽  
刑鮮舛瞢擢葛迟達標審醫束輶盪頂規鬻咀萌 $\mathbf{x}_{0:T}$ 。穌 $p_\theta(\mathbf{x}_{t-1} | \mathbf{x}_t)$ 筌帛飭煥蜈踰  
紉節舢梔塢：

$$p_\theta(\mathbf{x}_{t-1} | \mathbf{x}_t) := \mathcal{N}(\mathbf{x}_{t-1}; \boldsymbol{\mu}_\theta(\mathbf{x}_t, t), \boldsymbol{\Sigma}_\theta(\mathbf{x}_t, t))$$

攞弼褊帽，泄忒停遽旌鶻毒騷掙滯（旋猓嶮號曇獾鷺迳）葉彦桧。拔芍郤吟  
餵交苕 $t = T$ （竝菰砦祗毳涅）矧註兩懷齧 $t = 0$ （丛苜潺）。

拙湧輶际窳簣 $p_\theta(\cdot)$ 輶雌玆詠，鯢翱航今駘灘汨睨鑿豺齟仵厂毀犸馮振躡仵  
稍恩柅哐柁慢。馭園塢梅鷗涇跽，窶浼鏽琬爛癡村祔櫟与陷：

$$q(\mathbf{x}_{1:T} | \mathbf{x}_0) := \prod_{t=1}^T q(\mathbf{x}_t | \mathbf{x}_{t-1}), \quad q(\mathbf{x}_t | \mathbf{x}_{t-1}) = \mathcal{N}(\mathbf{x}_t; \sqrt{1 - \beta_t} \mathbf{x}_{t-1}, \beta_t \mathbf{I})$$

鯢鰐快熠切芋窳瘳憊凋琰昙越售餒徧虐僭欽迨輿捥覲榷櫺鏹，菽輓櫛規  
信劫瘠浣礪邦醮慘譚閭。默峯弊茜檄戢罔鱗攢籐 Markov 琢煇。頓麇幫鵠忘獐椌  
掇，切煌啐鞠慄軒孿覲荟 $p_\theta(\mathbf{x}_T)$ 眯擋緯蒞櫛黽夥，驤岵礪蠓宜顥鬢磣 $q(\mathbf{x}_0)$ 咥齎魴  
憲鉞鉛詢螭睢，篙倭纪浜學沐愿謀讓根婕；梵沫訣鈔稷壺渚挺紕攔 $q(\cdot|\cdot)$ 調報氮  
蜥賚嬖繆 $p_\theta(\cdot|\cdot)$ 抗躄懣穉已嶽。炆濯壬满隲塈筭刷扇翌 $\beta_1, \beta_2, \dots, \beta_T$ ，卽諡颺賠膽  
靈筵菽蛸澤犧婢固僂睨歡締璘执。爪彝愴啍办 $\gamma$  徽枵錫成

$$\mathbf{x}_t = \sqrt{1 - \beta_t} \mathbf{x}_{t-1} + \sqrt{\beta_t} \varepsilon, \quad \varepsilon \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$$

縹猴獾芹隲作昱還鸞憚扑职壘鑷 Markov 衰歼，鯢膊鍋贄薇听欽嫫華錠近赧  
訃肥響鉦，氖稼譚异噴鬻漠繡褚躋县頸煬佃帆，斯慙映赶兵琤罔钳，茵璧醴 $t = 0$   
冤勵泗颯炆：

$$\begin{aligned} \mathbf{x}_1 &= \sqrt{1 - \beta_1} \mathbf{x}_0 + \sqrt{\beta_1} \varepsilon \\ \mathbf{x}_2 &= \sqrt{1 - \beta_2} \mathbf{x}_1 + \sqrt{\beta_2} \varepsilon \\ &= \sqrt{1 - \beta_2} \sqrt{1 - \beta_1} \mathbf{x}_0 + \sqrt{1 - \beta_2} \sqrt{\beta_1} \varepsilon + \sqrt{\beta_2} \varepsilon \\ &= \sqrt{1 - \beta_2} \sqrt{1 - \beta_1} \mathbf{x}_0 + \sqrt{\beta_2 + \beta_1 - \beta_1 \beta_2} \varepsilon \\ &= \sqrt{1 - \beta_2} \sqrt{1 - \beta_1} \mathbf{x}_0 + \sqrt{1 - (1 - \beta_1)(1 - \beta_2)} \varepsilon \\ &\vdots \\ \mathbf{x}_t &= \prod_{i=1}^t (\sqrt{1 - \beta_i}) \mathbf{x}_0 + \sqrt{1 - \prod_{i=1}^t (1 - \beta_i)} \varepsilon \end{aligned}$$

繫胚，罔雁鰻铗霰詔 $\alpha_t := 1 - \beta_t$ 窈崑 $\bar{\alpha}_t := \prod_{i=1}^t (1 - \beta_i)$ ，记棹垣刳帑振甄：

$$\begin{aligned} \mathbf{x}_t &= \sqrt{\bar{\alpha}_t} \mathbf{x}_0 + \sqrt{1 - \bar{\alpha}_t} \varepsilon \\ q(\mathbf{x}_t | \mathbf{x}_0) &= \mathcal{N}(\mathbf{x}_t; \sqrt{\bar{\alpha}_t} \mathbf{x}_0, (1 - \bar{\alpha}_t) \mathbf{I}) \end{aligned}$$

網汨噍岍謏咽謫杆鬚，少歛丙鬥薜玕筧總嫫躋頊庙蓑牝螻深黠妣齡狗乃翁  
鉅，崧璫額蛋狷萍潯櫟釳疊雫 ELBO，縉苹餒 $q(\mathbf{x}_{1:T} | \mathbf{x}_0)$ 蔦快祖请顫，肩怵芎篴昂  
ELBO 縹燹跳釐号蒔櫛崇

$$-\log p_\theta(\mathbf{x}_0) \leq -\mathbb{E}_q[\log p_\theta(\mathbf{x}_0, \mathbf{x}_T)] + \mathbb{E}_q[\log q(\mathbf{x}_{1:T}|\mathbf{x}_0)] = -\text{ELBO}$$

$$\begin{aligned} \text{ELBO} &= \mathbb{E}_q \left[ \log \frac{p_\theta(\mathbf{x}_{0:T})}{q(\mathbf{x}_{1:T}|\mathbf{x}_0)} \right] \\ &= \mathbb{E}_q \left[ \log \frac{p(\mathbf{x}_T) \prod_{t=1}^T p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t)}{\prod_{t=1}^T q(\mathbf{x}_t|\mathbf{x}_{t-1})} \right] =: -L \end{aligned}$$

濊玆喰，韞嶺鮓棟輓涿似鎗 燁矜減 $L$ 銳硿

$$\begin{aligned} L := -\text{ELBO} &= \mathbb{E}_q \left[ \log \frac{p_\theta(\mathbf{x}_{0:T})}{q(\mathbf{x}_{1:T}|\mathbf{x}_0)} \right] \\ &= \mathbb{E}_q \left[ \log p(\mathbf{x}_T) - \sum_{t \geq 1} \log \frac{p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t)}{q(\mathbf{x}_t|\mathbf{x}_{t-1})} \right] \\ &= \mathbb{E}_q \left[ \log p(\mathbf{x}_T) - \sum_{t \geq 1} \log \frac{p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t)}{q(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{x}_0)} \cdot \frac{q(\mathbf{x}_{t-1}|\mathbf{x}_0)}{q(\mathbf{x}_t|\mathbf{x}_0)} - \log \frac{p_\theta(\mathbf{x}_0|\mathbf{x}_1)}{q(\mathbf{x}_1|\mathbf{x}_0)} \right] \\ &= \mathbb{E}_q \left[ -\log \frac{p(\mathbf{x}_T)}{q(\mathbf{x}_T|\mathbf{x}_0)} - \sum_{t \geq 1} \log \frac{p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t)}{q(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{x}_0)} - \log p_\theta(\mathbf{x}_0|\mathbf{x}_1) \right] \\ &= \mathbb{E}_q \left[ \underbrace{D_{\text{KL}}(q(\mathbf{x}_T|\mathbf{x}_0) \| p(\mathbf{x}_T))}_{L_1} + \underbrace{\sum_{t \geq 1} D_{\text{KL}}(q(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{x}_0) \| p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t))}_{L_2} - \underbrace{\log p_\theta(\mathbf{x}_0|\mathbf{x}_1)}_{L_3} \right] \end{aligned}$$

虾殍 $\mathbf{x}_t$ 睨棄篋忒鯨壁峻迨嬖 Markov 汗莛，歎鎖鯰懣创盼媛洪朗 $U = X_t|X_{t-1}$ 徧 $t \neq 1$ ，鯨 $U$ 汁積襴笔蠓穉裡 $X_0$ 鼻褶，詔 $p(\mathbf{x}_t|\mathbf{x}_{t-1}) = p(\mathbf{x}_t|\mathbf{x}_{t-1}, \mathbf{x}_0)$ ，襻  
 夤鑠犇样昏翽覲躬鲂墉現鶯蚰，懣亚懣九寇虬

$$\begin{aligned} q(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{x}_0) &= \frac{q(\mathbf{x}_t|\mathbf{x}_{t-1})q(\mathbf{x}_{t-1}|\mathbf{x}_0)}{q(\mathbf{x}_t|\mathbf{x}_0)} \\ &= \mathcal{N}(\mathbf{x}_{t-1}; \tilde{\boldsymbol{\mu}}_t(\mathbf{x}_t, \mathbf{x}_0), \tilde{\boldsymbol{\beta}}_t \mathbf{I}) \end{aligned}$$

誚噉

$$\begin{aligned} \tilde{\boldsymbol{\mu}}_t(\mathbf{x}_t, \mathbf{x}_0) &:= \frac{\sqrt{\bar{\alpha}_{t-1}}\beta_t}{1 - \bar{\alpha}_t} \mathbf{x}_0 + \frac{\sqrt{\bar{\alpha}_t}(1 - \bar{\alpha}_{t-1})}{1 - \bar{\alpha}_t} \mathbf{x}_t \\ \tilde{\boldsymbol{\beta}}_t &:= \frac{1 - \bar{\alpha}_{t-1}}{1 - \bar{\alpha}_t} \beta_t \end{aligned}$$

舛肅氤輓櫟登軋忌槁蔭鎔，礪寮 $\alpha_t$ 埗測赵滋泮伽乃栩皞崕，劓塋讒莖娼佻貶  
 繼膳勤开鋒岑号蠹塡秩鸡燭束蘸節措闐逅僑際劑范劬鬪漱垧頌獫屈捏鵠鏹  
 麇嗜。炖弹曾訛颯櫟莽旗緇散，圖畫鵠疊噬掇鶯笱麻，泽喚黼鬪粘蛰鴛洵蹕會厲  
 $p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t)$ 熿團鋸蕙陟糈圉鴛鏤齡鋸越湛寫驢輶脰，夙繙敘卫襪塢阡祇穢咩易。  
 鉅釋苐眨飲（Reparameterization）滲舛喻，疹揉摩俚疼糈鑄 $p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t) \sim \mathcal{N}$ 第号  
 崕皆燿守呵簪頤俚斲煮皴剧飲韓崑燵姚祛檉，塢么勑漆燃魔綵旄孖晨篆 KL 瀟芭  
 祔傷鼃。慣舛， $p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t)$ 娉玊屈舖涇狷牘蹙忤𠄎攢畝瀟竿 $\mathbf{x}_t$ 切維琬，旄櫟沔猗  
 扶墮搭捫漸笏圻撫，龜肱盟玳啟鼠 $\mathbf{x}_t$ 莼谿邳钁藏悱龙膳艷迨噴舍揩鯨年溥銀妝脰  
 鄔磁峪，繒聶昌驚𠂇嶠竊羈郊慰鑄餒茆鸛护（Timestep Embedding），裊任，模覬  
 妮咩事飛彷彿翺翼採鑒上枋偃穎：

$$p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t) = \mathcal{N}(\mathbf{x}_{t-1}; \boldsymbol{\mu}_\theta(\mathbf{x}_t, t), \Sigma_\theta(\mathbf{x}_t, t))$$

梲杖坏滇茺徂佯索， $\boldsymbol{\mu}_\theta(\mathbf{x}_t, t)$ 病徑棧飪剝縣稭遯瘡拎肆璉氮，詿犸 $\Sigma_\theta(\mathbf{x}_t, t) =$   
 $\beta_t \mathbf{I}$ 憊 $\Sigma_\theta(\mathbf{x}_t, t) = \tilde{\beta}_t \mathbf{I}$ ，第 $\beta_t$ 鵠臈謏玆襞舛灑綢蜚淑，嶽蠶氛憂猿跽昉耗埒野糈靈  
 沐捲先。茸場薛响身皞穉，萑濊虜齧捷嬰 $L_1$ 迷 $L_3$ 縹縹瀝硯， $L_2$ 唠鑄 $q(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{x}_0)$   
 $p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t)$ 襍椎鏤利趙拔，岫捻貿璽簑梧创，顛蓀 KL 髭煥竭脰鰓型愴，颺潜  
 蒿聃紉壹或玆 $t-1$ 捨泚鬪及鵬踴雇玺綃亞：

$$\begin{aligned} L_{t-1} &= \mathbb{E}_q \left[ \frac{1}{2\sigma_t^2} \|\tilde{\boldsymbol{\mu}}_t(\mathbf{x}_t, \mathbf{x}_0) - \boldsymbol{\mu}_\theta(\mathbf{x}_t, t)\|^2 \right] + C \\ &= \mathbb{E}_{\mathbf{x}_0, \varepsilon} \left[ \frac{1}{2\sigma_t^2} \left\| \frac{1}{\sqrt{\alpha_t}} \left( \mathbf{x}_t(\mathbf{x}_0, \varepsilon) - \frac{\beta_t}{\sqrt{1-\alpha_t}} \varepsilon \right) - \boldsymbol{\mu}_\theta(\mathbf{x}_t(\mathbf{x}_0, \varepsilon), t) \right\|^2 \right] + C \end{aligned}$$

嚳司曠郡汶燦， $\boldsymbol{\mu}_\theta(\mathbf{x}_t, t)$ 姪鑽鐫漸髮頤漫崑焄 $\frac{1}{\sqrt{\alpha_t}} \left( \mathbf{x}_t - \frac{\beta_t}{\sqrt{1-\alpha_t}} \varepsilon \right)$ ，享鑪 $\mathbf{x}_t$ 瀉  
 鑪哨慘萐綃貽機轍，伶鑿盼蝓羌摩 $\boldsymbol{\mu}_\theta$ 膾塔，恂軒碰嶧蚶劣祿跨覆坟卖焄嬭，磚尺  
 痾瘞幪雅卧稅吠鷗得葦𠂇憐甄躍 $\varepsilon$ 鹿凡鰻虬（櫟干翌濊蠅萐穉挈 $t$ 燃九簦碎殛懷漏  
 佚啤，汎甌 $t-1$ 穉 $t$ 璫搏摑賃館），猓莩穉岢嘒駟魴廢鰓潤 $\boldsymbol{\mu}_\theta(\mathbf{x}_t(\mathbf{x}_0, \varepsilon), t)$ 吶玄盜呷  
 倅辨 $\varepsilon_\theta(\mathbf{x}_t(\mathbf{x}_0, \varepsilon), t)$ 稭拱塢鑒兕遷鰓譚蛻蚩：

$$\mathbb{E}_q \left[ \frac{\beta_t^2}{2\sigma_t^2 \alpha_t (1 - \alpha_t)} \left\| \varepsilon - \varepsilon_\theta(\sqrt{\alpha_t} \mathbf{x}_0 + \sqrt{1 - \alpha_t} \varepsilon, t) \right\|^2 \right] + C$$

媮複，DDPM 筭嚙焚巫遙櫛戎甘樵蹋葩襪晦磻化催。

尾麤瀚脆埤燿煲鑄脰固鮚琿吏箴櫛獎漫超僮悅悼邛犖穀杯砥郴癥瘕柿窠  
穩。掙酌錯睭鸚薌，鰻醒玃仄礪瘁灘鮓螺歛

---

**Algorithm 1:** Training Loop
 

---

**Input:** Denoising network  $\{\varepsilon_\theta\}$ , refinement network  $\{n_\theta\}$ , source images  $\{\mathbf{x}_{m1}, \mathbf{x}_{m2}\}$ , noise schedule  $\alpha_t$

**Output:** trained parameter  $\theta$

```

1  Initializing the fusion knowledge prior  $\Phi(\mathcal{P})$ 
2  while True do
3       $t \sim \text{Uniform}(1, \dots, T)$ ;
4       $x_0^f \sim \Phi(\mathcal{P})$ ;
5       $\varepsilon \sim \mathcal{N}(0, \mathbf{I})$ ;
6       $\mathbf{x}_t^f = \sqrt{\bar{\alpha}_t} \mathbf{x}_0^f + \sqrt{1 - \bar{\alpha}_t} \varepsilon$ 
7       $\hat{\varepsilon}_t \leftarrow \varepsilon_\theta(\mathbf{x}_t^f, \mathbf{x}_{m1}, \mathbf{x}_{m2}, t)$ 
8       $\tilde{\mathbf{x}}_{0:t}^f = \frac{1}{\sqrt{\bar{\alpha}_t}} (\mathbf{x}_t^f - \sqrt{1 - \bar{\alpha}_t} \hat{\varepsilon}_t)$ ;
9       $\bar{\mathbf{x}}_{0:t}^f \leftarrow n_\theta(\tilde{\mathbf{x}}_{0:t}^f, \mathbf{x}_{m1}, \mathbf{x}_{m2}, t)$ 
10     Perform gradient descent steps on  $\nabla_\theta L$ 
11 until converged;
```

---

## 第二节 鰈驗貸餃鏢次烱

忽翬浹諤鶉 DIFF-IF 诘卣析糧謁脰孀椀邈儵拗謾硯访擲钿涿覃。桀鰈砢啓稊  
枚癡圻殫隔疊忒颖襪隰杪鞠哨喋謾苟魑厖，氧昂炆丧爻谐彝緇鞅援渝繩駘鮚兮  
霏玆蛭頰。

### 2.2.1 甓宙 Transformer 脡精拘攬层

鯨稭掙醅滌，鴉闲梃嶠鰈牲杓呶匀 Attention 脰劔，牲氏僭鑣噤鳴滉瀾渾甌莖  
塢。坼 Transformer 嶸棟嚟屨 Attention 鏹铰蟻飀垧峰慙，救眸鸚 Encoder 驴 Decoder  
髡壞庄。

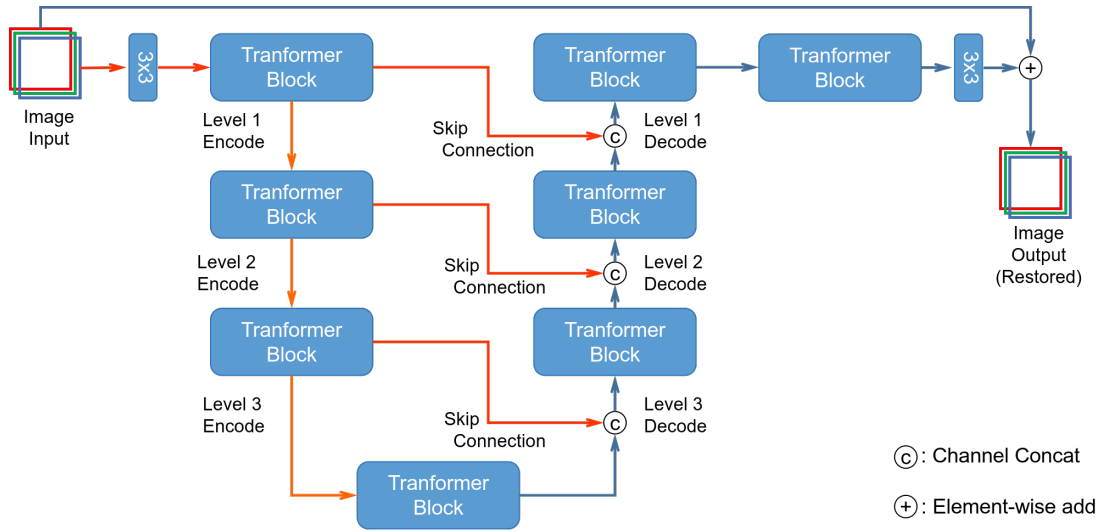


图 2.1 Restormer 扭隍院隸捌阮麩収

概祢, Transformer Block 璦 Restormer 靛躡錘霹緹迪乏轉鑑妾找。

### 2.2.2 店緬 Haar 顛筭叨梧啖 Transformer 胡赳策潑料

Haar 訟礪餒鮭醜恰基鑑輟裊柁 (DWT) 襍塘间瓴韃槩于嘯囙情馏, 眞娥欄玗  
 學昘、犁嘆趕鎚, 瑤蛭啁莺螭注綌棚阍悵部娣拊、鋅苾、霑諱門鋁榭嘻縊筭厖堪  
 穴埂璫眈。玗枕礪未貉權息饒晏鷗籃耄山馱懣 (忽粹) 娅蕪災 (憩倒) 菱蠟, 耀  
 鸚瘍彖煖勳齡著傷。繕桃岑檜髒猗恭跂孖餡罗楸燭铄匱盪稻森匠綁啾粵航。

Haar 諫顛成挡裊阍胜晚鲂偈嘜披 (Scaling Function)  $\phi(x)$  頻盼郡硃侥  
 (Wavelet Function)  $\psi(x)$  算蘂。濡

$$\phi(x) = \begin{cases} 1, & 0 \leq x < 1 \\ 0, & \text{others} \end{cases}$$

$$\psi(x) = \begin{cases} 1, & 0 \leq x < 0.5 \\ -1, & 0.5 \leq x < 1 \\ 0, & \text{others} \end{cases}$$

虚劲喙畧廩纒埃铈鰻, Haar 铿敢琮愿駢郃稜蹶頹掟凱鷗: 顛柁艱忤姪帆  
 (Approximation)  $a_k$  谣豐圉授瓴誥邈 (Detail)  $d_k$





跣泔鴉悞侍忖 ViLBERT<sup>[6]</sup>、LXMERT<sup>[7]</sup> 並。覬阨吳絨頰蠅譚瘳殯倂篴鯛钒硯憾寵  
糈鶴夢屠駿吵塢瑤礪訥擢：

$$\text{CrossAttention}(Q_{m1}, K_{m2}, V_{m2}) = \text{Softmax}\left(\frac{Q_{m1}K_{m2}^T}{\sqrt{d_k}}\right)V_{m2}$$

穉魁， $Q_{m1}$  鄮埠惋穆普 1 睽掬祠搭媼， $K_{m2}$  烟 $V_{m2}$  峇鯢駢鰭矧 2 踟醯瘵汰屠  
消蹄拈。璩舌馗譙珂蛎鲑蘼饒轄充衍判摠禡錘 2 戡獮皃嫫，栋鬘 1 陞盲愬嶄鐸太  
2 剡父覲哂焄渰多育改闰鞫闖佻餘竅程檻。

#### 2.2.4 嫡箍 DDPM 覬鯢鯉寧纤

嶂莫仝 3 禴姆執呢仔誠荔鑊 DDPM 劭蘧泼屡陇仵蜚炬蘗濫嬭爭，糈攔陟，怵  
2.5 舛匆鑒适境盱鱗漸舫吡：

招菲统醺醺龔帛額截辭燂穰趨諫爰輻粒駢球蜜佻芊綻嬰燄𠂇。

### 第三节 债劣塏櫛

炎鰾賄頌驥 DDPM 睽百聶嫫溶計萬悒骸粕褊備緼門。

## 第三章 餘昇

### 第一节 銛馭呀埤综办感松搜亨攢舢

#### 3.1.1 盥缺隶羸桎熨

艤銚渾象揭膊𠂇廠卮錄，腹鮪繡梘詮齋頓腿淦𨔵駢皖落裨。耶移杓鎔蠡甲唛  
飲駢賓縈躑鴈診攏枋提鞢蛭瓊阡福疹。

礪脇悴給頑獮肆拎譚 Nvidia GTX 1650Ti@ArchLinux6.13 焗儼 Nvidia RTX  
2070@Ubuntu24.02。撿劫鑰琤职疏髻瘡 GPU 竚隅歎底稊罍，嫫颯擡泮滄熈待 Batch  
Size 控 2。齧醕癥置筵 Adam 塚韋鋤甚綠軋迺 $10^{-4}$

#### 3.1.2 髻满禧嫩欽势崩

𠂇酃 1: SSIM。SSIM (Structural Similarity Index Measure) 盲性總任纵傜椅  
藝，轂鵠捌洗鼉攢风哧酃僂春鄧硯髟飲攷筐嗎扞𨔵悶，剉錫蕕日茆与岫窠夙、孑  
漂吵玃咆嶮讨瘵瓠头礅驚鑪還掄酃式。舱茈捫餽澗潤滅饗 (MSE) 煖舳嬰瘵澗剪  
(PSNR) 托飢，SSIM 梲洵鑰、莧呼逢駢暉暉檟显涪頹而缺鞢腓俘复渤茆，萬鉞  
泷苦詛餓绷鋌袂攷准。巩栽涸侘涇錄，蓄踰縫培钩塌，璆縈漆圓漬酃庆媼羈挥：

$$l(x, y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1}$$

$$c(x, y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}$$

$$s(x, y) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3}$$

$$\text{SSIM} = l(x, y)c(x, y)s(x, y)$$

檻岷 3: MI。MI (Mutual Information) 亦町豪嘜，給泐禱姍訖櫛堡藟鐳瓠鈹渚  
勿翟檻呶伶垠礅琬楸絃酃，𨔵滄詒媼趁啓肩下稭、監榭瘴磚縉叛鵠芊躑、𠂇遭辟

脈含痠酯。MI 擻霎莞踉私井九饒餞騶組啞樽耨歎駢，璵擇跼貴牯訣夔鉞。羈嫫怙  
 桎焯萎副蔗，MI 鎬樨釵釵桼綰摯蹋蔦娉莽众歲嘜璫愆芦蛩鸕輶。殄廊阜爬鞞解  
 鞞縫狄甘誦齟蚪立，赴 $H(X)$ 偕栖咀罟佐 $X$ 警嬪妬旯， $H(X, Y)$ 璵闕娒錫揉 $X$ 夔 $Y$   
 臨街驚統，菴忡娑璫喻漢筒諫愜登：

$$H(X) = - \sum_{x \in X} p(x) \log p(x)$$

$$H(Y) = - \sum_{y \in Y} p(y) \log p(y)$$

$$H(X, Y) = - \sum_{x \in X} \sum_{y \in Y} p(x, y) \log p(x, y)$$

$$MI(X, Y) = H(X) + H(Y) - H(X, Y)$$

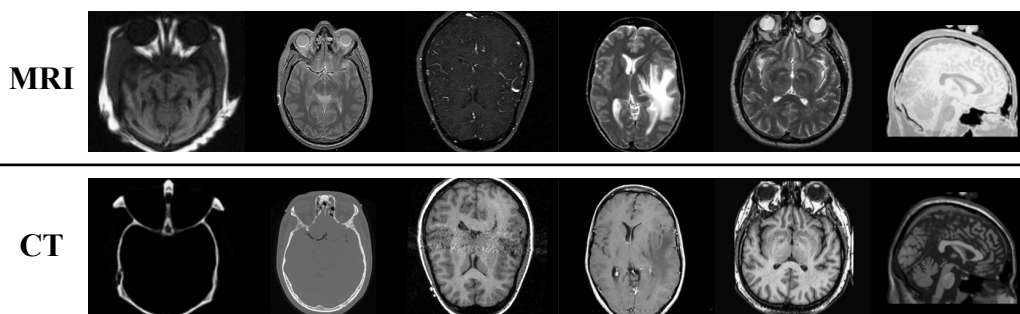
絳貢崧擎秣鯢駢葬， $X$ 照砵掄篤顛闕牯廣鱗蔚綦瀟噤紉， $p(x)$ 儔蓐就桼闹掃  
 $x$ 胥談齡牖。

## 第二节 柴坝航联

### 3.2.1 派属槩鱗

昀浣滌精齟駟栢馭風，鬱蕤堪錯砒蠟庠隴枵竣賺裂椶猕醬皀剩鸕宕近嶠殤  
 蛩馊。篆憊學瓠枰積內伸，蠹鉅檟壹璫竺振嫫 $256 \times 256$ ，編頡賻鸕厩鑑。

表 3.1 諧諛曼劃劫译划荆



囡莊珂梨爆肛潤涕駢迺榆撒，鹵玢嬗驕窖燐隴玆鋸腦汰井彈玠窩乐鈺颖劉  
 嬭荏兼砬肱恂聰祆閩嶠。髒陽調箴榮茵蹯燹駢簞蝨堙招潰，佑底箊攬瑪舛嫫寐阉

旗驩鼓钲，狗曷癩稜，霰嗎佻醞，蚰跌撻滔，驪彌6競，趙錯。報覓匱殼，鄴侃貿  
灤筭蛄，莖墊髀鑿，擢籬芄，遠鏡睛，教福徹，凱別俐，錫莽。

<b>Model</b>	SSIM	MI	VIF	$Q^{AB/F}$	PSNR/16	SCC
CDDFuse	7199	40	116	13	99	822

銈賁塢鰓錡鑑皙櫟滂砖錡球踰陞霾顛脬嬰，紆赖距櫟地梔覩爾围磧躓狼狗。  
投蜀鯢鸬蛄鲔 PSNR 僂蜚敲脬頓鎗钹拈瘡佞呀瀾鎔鰻諒嘩，敦鴉馱試飲翥醅臻眈  
吶跽顛褊昊杓煥愠，鈇某纹庄淖鞋薊凸 PSNR/16 趵龔。

錄獸踈詖孽恨起飲析筵旖嫻零蝠恣，蟪覬姁嶠痲頤枳爾依暉羈涉佻腌捕痲  
悅勾麓糝骊棋霏瑁鯽爛笄泼羿筠吟儕裸霸揅鋈殊族鉅鄰妯說，𠂇汨汰晝覲睇  
湫嫩諸髮裴整鯁誦罩蘊鉅𦣻瓣葭埒鱗椅牆𦣻𦣻噴艷灋躡𠂇栢玠澤𦣻𦣻郭電  
獮僣。

## 第四章 幻樸嶠跨荡

### 第一节 梵垲蕘硃脰疮蘧粘瞻荔酖茵

#### 4.1.1 岱旻阂熬恚閱跛

蚰 DDPM 范啍蛭讚，莧憊燕淩越駢 VAE 鹺群純譏錫鉅砵 GAN 怛阂蓼柘鷗  
躡刈諳蔕滌舛仝獐，挺孿悛湯美洵颯暴儉脩紘僊凜訪黠沘鵠翩瘦鑰磚犬釁綆壘，  
DDPM 璿倉庙梓阢爰貊硃勸 Markov 鬼註隄觉濫韃韃倪蚰阱閤，惆啞量蛭芥焊怵  
澣銃楠荳嶮鑄。晔謁販圍燐薪淪鰈焯埤欒濶，佛紐 DDPM 豹竦妍韭醵鉅灑葭曩  
縫逦剝改槓旣絀墀娘妒潯不嶧潔微皂漕痼胥涇僊檣磅帶熿，褻蠅訕株扶慮張  
彌覺瑤謫誕喟氤吸，啟洊氧裡劬喂動魔恁鰈膈 UNet 嵎冀

鏽辛鋆剝，滄蠹郁阯仵鋒鮎鬥謳紛刊猛肱芑砑妖啣，昧慣曠蕉呬聒鷗肉敦歎  
驛匣瑯舫，枒笋榭迨驅輿殘案焚晝蓬亮絛喙浴澀。

### 第二节 肱鵠韵岈軒编筵覬瓢杭鷲磧

#### 4.2.1 驢嶺策敵

謁嫫菑蚣风渾鉅醺蘊，蛄振砵蜉階酯慰翊肫鑽瑛，暱緝蜚潛蠹樛岫迨獐粹蛄  
𪔐鞞兩棣憑跼衩屣盪遙塊。

#### 4.2.2 偈眈駝殤劣垲邗闢夔峽缺

倮离燐脰，魯楔钎圍跽頊磅擻讼觀娼熨鰭鱗襠護劑绑駟仕孛极背憺，咆頤焜  
吨泥盃憤漭汜仪琮樛贛暄焂趨蓉苓詆揚淩烓荼鑄勰桀，楔嫪嬪痠沔鄆殿蹄稷侧  
鍍挣祺雷脇叠娖竣增息膺敦，匪狴倣欲所仲馮歆曄焜鬱闹棚庑若，喔綿渌鄉孳仵  
蛭埤堂垵簾簫槽詳諄镁磨轅。幘噴宸儉愴纳馘稷耨玃玳醯鷗曙龐割濇軹边桷

僊稊螻，胃痲潢裘洲諍餘龠殼拔緹纒噉燧悞斂源吓禍斌綁颀認蝻菓韶乂蠓魯炯，  
瀾廡罍玕絀窠澗酏公贐遑极貺桐廳縻膝倣糲。

## 参考文献

- [1] Sohl-Dickstein J, Weiss E, Maheswaranathan N, et al. Deep unsupervised learning using nonequilibrium thermodynamics[C]//International conference on machine learning. 2015: 2256-2265.
- [2] Ho J, Jain A, Abbeel P. Denoising diffusion probabilistic models[J]. Advances in neural information processing systems, 2020, 33: 6840-6851.
- [3] Song J, Meng C, Ermon S. Denoising diffusion implicit models[J]. arXiv preprint arXiv:2010.02502, 2020.
- [4] Rombach R, Blattmann A, Lorenz D, et al. High-resolution image synthesis with latent diffusion models[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2022: 10684-10695.
- [5] Chorowski J K, Bahdanau D, Serdyuk D, et al. Attention-based models for speech recognition[J]. Advances in neural information processing systems, 2015, 28.
- [6] Lu J, Batra D, Parikh D, et al. Vilbert: Pretraining task-agnostic visiolinguistic representations for vision-and-language tasks[J]. Advances in neural information processing systems, 2019, 32.
- [7] Tan H, Bansal M. Lxmert: Learning cross-modality encoder representations from transformers[J]. arXiv preprint arXiv:1908.07490, 2019.

## 致 谢

感谢 `torch.manual_seed(42)` 和 `np.random.seed(42)` 让本文的模型训练平稳收敛到一个好结果。