



(12) 发明专利

(10) 授权公告号 CN 116664562 B

(45) 授权公告日 2023. 11. 03

(21) 申请号 202310935992.6

G01N 21/88 (2006.01)

(22) 申请日 2023.07.28

(56) 对比文件

(65) 同一申请的已公布的文献号

US 2022156578 A1, 2022.05.19

申请公布号 CN 116664562 A

WO 2020015167 A1, 2020.01.23

(43) 申请公布日 2023.08.29

CN 113344903 A, 2021.09.03

(73) 专利权人 香港理工大学深圳研究院

CN 112416643 A, 2021.02.26

CN 115456929 A, 2022.12.09

地址 518057 广东省深圳市南山区粤海街

杨雪. 基于变分自编码器和分布式应变的桥梁表面裂缝异常检测.《中国优秀硕士学位论文全文数据库 工程科技 II 辑》.2022, (第3期), C034-903.

道高新技术产业园南区粤兴一道18号

香港理工大学产学研大楼205室

(72) 发明人 魏元昊 倪一清 王友武

TINGTING CHEN et al. Unsupervised Anomaly Detection of Industrial Robots Using Sliding-Window Convolutional Variational Autoencoder .《IEEEAccess》.2020, 第47072-47081页.

(74) 专利代理机构 深圳市君胜知识产权代理事

务所(普通合伙) 44268

专利代理师 谢松

(51) Int. Cl.

G06T 7/00 (2017.01)

G06N 3/088 (2023.01)

G06V 10/766 (2022.01)

审查员 于芝枝

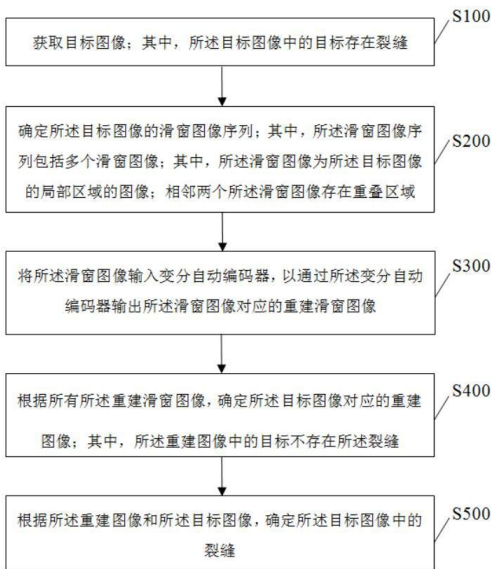
权利要求书2页 说明书10页 附图3页

(54) 发明名称

一种基于滑窗变分自编码器的非监督裂缝识别方法

(57) 摘要

本发明公开了一种基于滑窗变分自编码器的非监督裂缝识别方法,包括步骤:获取目标图像;目标图像中的目标存在裂缝;确定目标图像的滑窗图像序列;滑窗图像序列包括多个滑窗图像;将滑窗图像输入变分自动编码器,以通过变分自动编码器输出滑窗图像对应的重建滑窗图像;根据所有重建滑窗图像,确定目标图像对应的重建图像;重建图像中的目标不存在裂缝;根据重建图像和目标图像,确定目标图像中的裂缝。由于本申请是利用变分自动编码器消除目标图像中的异常值,得到不存在裂缝的重建图像,从而确定目标图像中的裂缝。本申请不需要正样本训练或学习特定于裂缝的图像特征,从而避免了缺乏训练数据或不平衡数据集带来的裂缝识别准确性较低的问题。



1. 一种基于滑窗变分自编码器的非监督裂缝识别方法,其特征在于,包括步骤:

获取目标图像;其中,所述目标图像中的目标存在裂缝;所述目标图像为铁路的旋转双耳组件的图像;

确定所述目标图像的滑窗图像序列;其中,所述滑窗图像序列包括多个滑窗图像;其中,所述滑窗图像为所述目标图像的局部区域的图像;相邻两个所述滑窗图像存在重叠区域;

将所述滑窗图像输入变分自动编码器,以通过所述变分自动编码器输出所述滑窗图像对应的重建滑窗图像;所述变分自动编码器用于消除滑窗图像中的裂缝信息;

根据所有所述重建滑窗图像,确定所述目标图像对应的重建图像;其中,所述重建图像中的目标不存在所述裂缝;

根据所述重建图像和所述目标图像,确定所述目标图像中的裂缝;

所述变分自动编码器包括:编码器、采样层和解码器;

所述将所述滑窗图像输入变分自动编码器,以通过所述变分自动编码器输出所述滑窗图像对应的重建滑窗图像,包括:

将所述滑窗图像输入所述编码器,通过所述编码器输出所述滑窗图像对应的似然分布;

将所述似然分布输入采样层,以得到所述滑窗图像对应的隐变量;

将所述隐变量输入所述解码器,通过所述解码器输出所述滑窗图像对应的重建滑窗图像;

所述将所述似然分布输入采样层,以得到所述滑窗图像对应的隐变量,包括:

将所述似然分布输入采样层,通过所述采样层输出采样的隐变量,对所述采样的隐变量进行高斯过程回归,得到高斯过程回归后的均值向量,并作为滑窗图像对应的隐变量;

所述变分自动编码器的损失函数为:

$$L = \sum_1^n \left\{ -E_{q_\phi(\mathbf{z}_w^i | \mathbf{x}_w^i)} [\log p_\theta(\hat{\mathbf{x}}_w^i | \mathbf{z}_w^i)] + KL(q_\phi(\mathbf{z}_w^i | \mathbf{x}_w^i) \parallel p(\mathbf{z}_w^i)) \right\};$$

其中, L 表示损失函数, \sum 表示求和操作, n 表示滑窗图像的数量, $E_{q_\phi(\mathbf{z}_w^i | \mathbf{x}_w^i)}(\cdot)$ 表示期望, $p_\theta(\cdot | \cdot)$ 表示解码器, θ 表示解码器的参数, $\hat{\mathbf{x}}_w^i$ 表示第*i*个重建滑窗图像, \mathbf{z}_w^i 表示第*i*个采样的隐变量, w 表示滑窗图像的宽度, \mathbf{x}_w^i 表示第*i*个滑窗图像, \log 表示取对数操作, $KL(\cdot \parallel \cdot)$ 表示KL散度, $q_\phi(\cdot | \cdot)$ 表示编码器, ϕ 表示编码器的参数, $p(\cdot)$ 表示先验分布;

所述高斯过程回归后的均值向量为:

$$\mu_{z_{gp}}^i = k(t_i, \mathbf{t})(K + \sigma_n^2 I)^{-1} \mathbf{z}_w^i;$$

$$k(t_i, \mathbf{t}) = \begin{bmatrix} \phi(t_i, t_1) \\ \phi(t_i, t_2) \\ \vdots \\ \phi(t_i, t_n) \end{bmatrix}^T;$$

$$K = \begin{bmatrix} \phi(t_1, t_1) & \dots & \phi(t_1, t_n) \\ \dots & \dots & \dots \\ \phi(t_n, t_1) & \dots & \phi(t_n, t_n) \end{bmatrix};$$

$$\phi(t, t') = \alpha \exp\left(-\frac{(t - t')^2}{2l^2}\right);$$

$$\mathbf{t} = [t_1, t_2, \dots, t_i, \dots, t_n];$$

其中, $\mu_{z_{gp}}^i$ 表示第 i 个高斯过程回归后的均值, $k(\cdot, \cdot)$ 表示核函数, \mathbf{t} 表示采样的隐变量的序号, K 表示核函数矩阵, t_i 表示第 i 个采样的隐变量的序号, $\phi(\cdot, \cdot)$ 表示径向基核函数, \mathbf{t}, \mathbf{t}' 分别表示采样的隐变量的序号 \mathbf{t} 中任意两个, α 和 l 是径向基核函数中的超参数, 分别表示径向基核函数的带宽和特征长度尺度, σ_n^2 表示采样的隐变量 \mathbf{z}_w^i 的噪音, \mathbf{I} 表示单位矩阵;

所述根据所有所述重建滑窗图像, 确定所述目标图像对应的重建图像, 包括:

根据第一个滑窗图像对应的重建滑窗图像的上半部分图像、最后一个滑窗图像对应的重建滑窗图像的下半部分图像以及其余滑窗图像对应的重建滑窗图像的中间部分图像, 组合形成重建图像; 其中, 所述中间部分图像的宽度根据预设窗口滑动的距离确定; 所述预设窗口滑动的距离为 1 个像素时, 所述中间部分图像为 1 行像素的均值 $\mu_{\hat{\mathbf{x}}_w^i}$, 即重建滑窗图像的中间。

2. 根据权利要求 1 所述的基于滑窗变分自编码器的非监督裂缝识别方法, 其特征在于, 所述先验分布被配置为标准正态分布;

所述重建滑窗图像中每个像素由一个单独的正态分布表示。

3. 根据权利要求 2 所述的基于滑窗变分自编码器的非监督裂缝识别方法, 其特征在于, 所述确定所述目标图像的滑窗图像序列, 包括:

将预设窗口配置在所述目标图像的一角, 并将所述目标图像在预设窗口内的图像作为滑窗图像序列的第一个滑窗图像;

滑动所述预设窗口, 并将所述目标图像在所述预设窗口内的图像作为所述滑窗图像序列的滑窗图像, 直至预设窗口遍及所述目标图像, 得到所述目标图像的滑窗图像序列。

4. 根据权利要求 1 所述的基于滑窗变分自编码器的非监督裂缝识别方法, 其特征在于, 所述预设窗口沿所述目标图像的长度方向滑动, 所述预设窗口的长度与所述目标图像的宽度相等。

5. 根据权利要求 1-4 任一项所述的基于滑窗变分自编码器的非监督裂缝识别方法, 其特征在于, 所述根据所述重建图像和所述目标图像, 确定所述目标图像中的裂缝, 包括:

将所述目标图像中每个像素的像素值与所述重建图像对应像素的像素值作差, 得到差值图像;

当所述差值图像中的连续多个像素的像素值大于预设像素值时, 将所述目标图像中对应的像素确定为裂缝。

一种基于滑窗变分自编码器的非监督裂缝识别方法

技术领域

[0001] 本发明涉及裂缝识别技术领域,尤其涉及的是一种基于滑窗变分自编码器的非监督裂缝识别方法。

背景技术

[0002] 结构系统内存在的裂缝可能导致其刚度降低,最终导致重大危险。一般来说,基于深度学习的裂缝检测方法可以分为两大类:监督学习和非监督学习。当有足够大的数据集可用于模型训练时,基于监督学习的裂缝识别方法已证明具有出色的性能。现有技术中,在训练数据不足或正负样本严重不平衡的情况下,基于监督学习的裂缝识别方法的识别裂缝的准确性较低。

[0003] 因此,现有技术还有待于改进和发展。

发明内容

[0004] 本发明要解决的技术问题在于,针对现有技术的上述缺陷,提供一种基于滑窗变分自编码器的非监督裂缝识别方法,旨在解决现有技术中在训练数据不足或正负样本严重不平衡的情况下基于监督学习的裂缝识别方法的识别裂缝的准确性较低的问题。

[0005] 本发明解决技术问题所采用的技术方案如下:

[0006] 一种基于滑窗变分自编码器的非监督裂缝识别方法,其中,包括步骤:

[0007] 获取目标图像;其中,所述目标图像中的目标存在裂缝;

[0008] 确定所述目标图像的滑窗图像序列;其中,所述滑窗图像序列包括多个滑窗图像;其中,所述滑窗图像为所述目标图像的局部区域的图像;相邻两个所述滑窗图像存在重叠区域;

[0009] 将所述滑窗图像输入变分自动编码器,以通过所述变分自动编码器输出所述滑窗图像对应的重建滑窗图像;

[0010] 根据所有所述重建滑窗图像,确定所述目标图像对应的重建图像;其中,所述重建图像中的目标不存在所述裂缝;

[0011] 根据所述重建图像和所述目标图像,确定所述目标图像中的裂缝。

[0012] 所述的基于滑窗变分自编码器的非监督裂缝识别方法,其中,所述变分自动编码器包括:编码器、采样层和解码器;

[0013] 所述将所述滑窗图像输入变分自动编码器,以通过所述变分自动编码器输出所述滑窗图像对应的重建滑窗图像,包括:

[0014] 将所述滑窗图像输入所述编码器,通过所述编码器输出所述滑窗图像对应的似然分布;

[0015] 将所述似然分布输入采样层,通过所述采样层输出所述滑窗图像对应的隐变量;

[0016] 将所述隐变量输入所述解码器,以得到所述滑窗图像对应的重建滑窗图像。

[0017] 所述的基于滑窗变分自编码器的非监督裂缝识别方法,其中,所述将所述似然分

布输入采样层,以得到所述滑窗图像对应的隐变量,包括:

[0018] 将所述似然分布输入采样层,将所述采样层输出采样的隐变量,作为滑窗图像对应的隐变量;或者

[0019] 将所述似然分布输入采样层,通过所述采样层输出采样的隐变量,对所述采样的隐变量进行高斯过程回归,得到所述滑窗图像对应的隐变量。

[0020] 所述的基于滑窗变分自编码器的非监督裂缝识别方法,其中,所述变分自动编码器的损失函数为:

$$[0021] \quad L = \sum_1^n \left\{ -E_{q_{\phi}(z_w^i | x_w^i)} [\log p_{\theta}(\hat{x}_w^i | z_w^i)] + KL(q_{\phi}(z_w^i | x_w^i) \parallel p(z_w^i)) \right\};$$

[0022] 其中, L 表示损失函数, \sum 表示求和操作, n 表示滑窗图像的数量, $E_{q_{\phi}(z_w^i | x_w^i)}(\cdot)$ 表示期望, $p_{\theta}(\cdot | \cdot)$ 表示解码器, θ 表示解码器的参数, \hat{x}_w^i 表示第 i 个重建滑窗图像, z_w^i 表示第 i 个采样的隐变量, w 表示滑窗图像的宽度, x_w^i 表示第 i 个滑窗图像, \log 表示取对数操作, $KL(\cdot \parallel \cdot)$ 表示KL散度, $q_{\phi}(\cdot | \cdot)$ 表示编码器, ϕ 表示编码器的参数, $p(\cdot)$ 表示先验分布;

[0023] 所述高斯过程回归后的均值向量为:

$$[0024] \quad \mu_{z_{gp}}^i = k(t_i, \mathbf{t})(K + \sigma_n^2 \mathbf{I})^{-1} \mathbf{z}_w^i;$$

$$[0025] \quad k(t_i, \mathbf{t}) = \begin{bmatrix} \phi(t_i, t_1) \\ \phi(t_i, t_2) \\ \vdots \\ \phi(t_i, t_n) \end{bmatrix}^T;$$

$$[0026] \quad K = \begin{bmatrix} \phi(t_1, t_1) & \dots & \phi(t_1, t_n) \\ \dots & \dots & \dots \\ \phi(t_n, t_1) & \dots & \phi(t_n, t_n) \end{bmatrix};$$

$$[0027] \quad \phi(t, t') = \alpha \exp \left(-\frac{(t - t')^2}{2l^2} \right);$$

$$[0028] \quad \mathbf{t} = [t_1, t_2, \dots, t_i, \dots, t_n];$$

[0029] 其中, $\mu_{z_{gp}}^i$ 表示第 i 个高斯过程回归后的均值, $k(\cdot, \cdot)$ 表示核函数, \mathbf{t} 表示采样的隐变量的序号, K 表示核函数矩阵, t_i 表示第 i 个采样的隐变量的序号, $\phi(\cdot, \cdot)$ 表示径向基核函数, \mathbf{t}, \mathbf{t}' 分别表示采样的隐变量的序号 \mathbf{t} 中任意两个, α 和 l 是径向基核函数中的超参数,分别表示径向基核函数的带宽和特征长度尺度, σ_n^2 表示采样的隐变量 \mathbf{z}_w^i 的噪音, \mathbf{I} 表示单位矩阵。

[0030] 所述的基于滑窗变分自编码器的非监督裂缝识别方法,其中,所述先验分布被配置为标准正态分布;

[0031] 所述重建滑窗图像中每个像素由一个单独的正态分布表示。

[0032] 所述的基于滑窗变分自编码器的非监督裂缝识别方法,其中,所述确定所述目标图像的滑窗图像序列,包括:

[0033] 将预设窗口配置在所述目标图像的一角,并将所述目标图像在预设窗口内的图像作为滑窗图像序列的第一个滑窗图像;

[0034] 滑动所述预设窗口,并将所述目标图像在所述预设窗口内的图像作为所述滑窗图像序列的滑窗图像,直至预设窗口遍及所述目标图像,得到所述目标图像的滑窗图像序列。

[0035] 所述的基于滑窗变分自编码器的非监督裂缝识别方法,其中,所述根据所有所述重建滑窗图像,确定所述目标图像对应的重建图像,包括:

[0036] 根据第一个滑窗图像对应的重建滑窗图像的上半部分图像、最后一个滑窗图像对应的重建滑窗图像的下半部分图像以及其余滑窗图像对应的重建滑窗图像的中间部分图像,组合形成重建图像;其中,所述中间部分图像的宽度根据所述预设窗口滑动的距离确定。

[0037] 所述的基于滑窗变分自编码器的非监督裂缝识别方法,其中,所述预设窗口滑动的距离为1个像素。

[0038] 所述的基于滑窗变分自编码器的非监督裂缝识别方法,其中,所述预设窗口沿所述目标图像的长度方向滑动,所述预设窗口的长度与所述目标图像的宽度相等。

[0039] 所述的基于滑窗变分自编码器的非监督裂缝识别方法,其中,所述根据所述重建图像和所述目标图像,确定所述目标图像中的裂缝,包括:

[0040] 将所述目标图像中每个像素的像素值与所述重建图像对应像素的像素值作差,得到差值图像;

[0041] 当所述差值图像中的连续多个像素的像素值大于预设像素值时,将所述目标图像中对应的像素确定为裂缝。

[0042] 所述的基于滑窗变分自编码器的非监督裂缝识别方法,其中,所述目标图像为铁路的旋转双耳组件的图像。

[0043] 有益效果:由于本申请是利用变分自动编码器消除目标图像中的异常值,得到不存在裂缝的重建图像,从而确定目标图像中的裂缝。本申请不需要正样本训练或学习特定于裂缝的图像特征,从而避免了缺乏训练数据或不平衡数据集带来的裂缝识别准确性较低的问题。

附图说明

[0044] 图1是本发明实施例中基于滑窗变分自编码器的非监督裂缝识别方法的原理图。

[0045] 图2是本发明实施例中基于滑窗变分自编码器的非监督裂缝识别方法的流程图。

[0046] 图3是本发明实施例中基于滑窗变分自编码器的非监督裂缝识别方法的另一原理图。

具体实施方式

[0047] 为使本发明的目的、技术方案及优点更加清楚、明确,以下参照附图并举实施例对本发明进一步详细说明。应当理解,此处所描述的具体实施例仅仅用以解释本发明,并不用于限定本发明。

[0048] 请同时参阅图1-图3,本发明提供了一种基于滑窗变分自编码器的非监督裂缝识别方法的一些实施例。

[0049] 研究表明,无监督学习可以有效地用于识别时间序列数据中的异常值。异常值的存在会破坏时间序列中数据点之间的时间结构。然而,具有稳健特性的概率模型可以减轻偶然异常值的影响。此外,由于环境因素和数据采集设备的影响,噪声往往伴随着采集的时间序列。因此,将收集到的时间序列视为随机变量,用概率模型对其进行建模更为合理。本申请旨在将图像转换为模拟时间序列数据的序列信息。因此,图像中存在的裂缝可以被视为异常值,并通过无监督学习使用稳健的概率模型来消除。

[0050] 作为深度学习算法的一种形式,变分自动编码器(Variational Autoencoders, VAE)展示了无监督学习能力和概率建模特性。此外,损失函数中Kullback-Leibler(KL)散度的存在赋予VAE模型出色的正则化特性,使其对异常值具有鲁棒性。因此,本申请采用VAE对输入图像信息进行概率建模。为了序列化图像信息,本申请提出了一种将图像数据分割成片段的滑动窗口策略,这些片段随后以顺序方式输入到VAE模型中。这样,图像中的裂缝信息可以被视为时间序列中的异常值,并被鲁棒滑动窗口VAE(Sliding Window VAE, SW-VAE)丢弃。最后,通过将输入图像与SW-VAE输出的重建图像进行比较,可以揭示裂缝信息。值得注意的是,本申请不需要正样本训练或学习特定于裂缝的图像特征,从而避免了缺乏训练数据或不平衡数据集带来的裂缝识别准确性较低的问题。

[0051] 如图2所示,本发明实施例的基于滑窗变分自编码器的非监督裂缝识别方法,包括以下步骤:

[0052] 步骤S100、获取目标图像;其中,所述目标图像中的目标存在裂缝。

[0053] 步骤S200、确定所述目标图像的滑窗图像序列;其中,所述滑窗图像序列包括多个滑窗图像;其中,所述滑窗图像为所述目标图像的局部区域的图像;相邻两个所述滑窗图像存在重叠区域。

[0054] 步骤S300、将所述滑窗图像输入变分自动编码器,以通过所述变分自动编码器输出所述滑窗图像对应的重建滑窗图像。

[0055] 步骤S400、根据所有所述重建滑窗图像,确定所述目标图像对应的重建图像;其中,所述重建图像中的目标不存在所述裂缝。

[0056] 步骤S500、根据所述重建图像和所述目标图像,确定所述目标图像中的裂缝。

[0057] 具体地,目标图像是指目标的图像,目标可以是基础设施结构等,例如,目标图像为铁路的旋转双耳组件的图像,对旋转双耳组件进行拍摄,得到目标图像,并识别目标图像中是否存在裂缝,实现结构健康监测(Structural Health Monitoring, SHM)。由于裂缝的存在,目标图像中裂缝处的像素值会出现突变,表现为异常值。

[0058] 基于目标图像得到滑窗图像序列,以模拟时间序列的数据,以便后续通过变分自动编码器进行处理。滑窗图像序列中有多个滑窗图像,滑窗图像按照一定顺序排列,具体是按照滑窗依次滑过目标图像的顺序排列。每个滑窗图像并未覆盖整个目标图像,滑窗图像是目标图像部分区域的图像,滑窗图像滑动的距离并不会超过滑窗图像的尺寸,因此,相邻两个滑窗图像存在重叠区域,滑窗的滑动距离越小,重叠区域越大,滑窗的滑动距离越大,重叠区域越小。

[0059] 变分自动编码器是从对象的学习分布中生成新的采样数据,本申请采用变分自

动编码器消除目标图像中的异常值,得到不存在裂缝的重建图像。将滑窗图像输入变分自动编码器,得到对应的重建滑窗图像,由于滑窗图像是目标图像的一部分,重建滑窗图像也是一部分内容,需要再根据所有重建滑窗图像,得到重建图像。重建图像的尺寸与目标图像的尺寸相同,重建图像和目标图像中均存在目标,不同的是,重建图像中的目标不存在裂缝,而目标图像中的目标存在裂缝。最后根据重建图像和目标图像,就可以确定目标图像中的裂缝,可以获得裂缝的相关信息,例如裂缝的位置,裂缝的尺寸,裂缝的形状等信息。

[0060] 由于本申请是利用变分自动编码器消除目标图像中的异常值,得到不存在裂缝的重建图像,从而确定目标图像中的裂缝。变分自编码器可以用于输入数据的重构,而由于其概率模型的本质和鲁棒性的特点,其具备自动忽略异常值的能力。因此对比通过变分编码器的输入和输出就可以实现裂缝的识别。无需任何数据对变分自动编码器进行训练,无需考虑训练样本数量,正负样本平衡等问题。此发明可以将目标图像直接输入模型,而如有裂缝,则裂缝会自动输出。由于其无需大量数据训练,因此成本更低,更简单。此外,由于变分自动编码器不是由某一特定目标的数据训练而来,其泛化能力也因此更好。

[0061] 步骤S200具体包括:

[0062] 步骤S210、将预设窗口配置在所述目标图像的一角,并将所述目标图像在预设窗口内的图像作为滑窗图像序列的第一个滑窗图像。

[0063] 步骤S220、滑动所述预设窗口,并将所述目标图像在所述预设窗口内的图像作为所述滑窗图像序列的滑窗图像,直至预设窗口遍及所述目标图像,得到所述目标图像的滑窗图像序列。

[0064] 具体地,预设窗口的尺寸可以根据需要设定,预设窗口的尺寸给定之后,滑窗图像的尺寸也就确定了,因此,可以根据所需要的滑窗图像的尺寸来确定预设窗口的尺寸。如图1所示,目标图像为矩形,预设窗口为矩形(虚线框表示预设窗口),预设窗口的一边长与目标图像的一边长相等,预设窗口的另一边长小于目标图像的另一边长,例如,预设窗口的长度与目标图像的宽度一致,预设窗口的宽度小于目标图像的长度,则滑动方向为目标图像的长度方向。预设窗口的任意一边长小于目标图像的任意一边长,则预设窗口可以沿两个方向滑动。

[0065] 预设窗口自目标图像的一角开始滑动,每滑动一下得到一个滑动图像,预设窗口遍及目标图像所有区域后,得到滑窗图像序列,即 $x_w^1, x_w^2, \dots, x_w^i, \dots, x_w^n$, 其中, x_w^i 表示第 i 个滑窗图像, w 表示滑窗图像的宽度。滑动的预设窗口的使用有效地限制了VAE输入层的神经节点数量,增强了方法的鲁棒性。

[0066] 如表1所示,变分自动编码器包括:编码器、采样层和解码器。编码器包括:输入层、隐藏层1、隐藏层2、隐变量均值输出层以及隐变量方差输出层。输入层与隐藏层1连接,隐藏层1采用全连接层,隐藏层1的输入节点有 m 个, m 表示滑窗图像的像素点的总数,隐藏层1的输出节点有10个,隐藏层1与隐藏层2连接,隐藏层2采用全连接层,隐藏层2的输入节点有10个,隐藏层2的输出节点有2个,分别连接隐变量均值输出层及隐变量方差输出层。

[0067] 解码器包括:隐藏层3、隐藏层4、重构均值输出层以及重构方差输出层。隐藏层3采用全连接层,隐藏层3的输入节点有1个,隐藏层3的输出节点有10个,隐藏层3与隐藏层4连接,隐藏层4采用全连接层,隐藏层4的输入节点有10个,隐藏层2的输出节点有 $2m$ 个,分别连接隐变量均值输出层及隐变量方差输出层。

[0068] 表1变分自动编码器的神经网络层及参数

[0069]

		神经网络层	输入节点	输出节点	激活函数
编码器		输入层→隐藏层 1	m	10	Relu
		隐藏层 1→隐藏层 2	10	10	Relu
		隐藏层 2→隐变量均值输出层	10	1	Linear
		隐藏层 2→隐变量方差输出层	10	1	Softplus
解码器		滑窗图像对应的隐变量→隐藏层 3	1	10	Relu
		隐藏层 3→隐藏层 4	10	10	Relu
		隐藏层 4→重构均值输出层	10	m	Linear
		隐藏层 4→重构方差输出层	10	m	Softplus

[0070] 其中, m 代表滑窗内像素点的总数, 即滑窗片段的长乘宽。

[0071] 步骤S300具体包括:

[0072] 步骤S310、将所述滑窗图像输入所述编码器, 通过所述编码器输出所述滑窗图像对应的似然分布。

[0073] 步骤S320、将所述似然分布输入采样层, 通过所述采样层输出所述滑窗图像对应的隐变量。

[0074] 步骤S330、将所述隐变量输入所述解码器, 通过所述解码器输出所述滑窗图像对应的重建滑窗图像。

[0075] 具体地, 将滑窗图像 \mathbf{x}_w^i 输入编码器, 编码器输出滑窗图像对应的似然分布 (似然分布具有均值 μ_z 和方差 σ_z), 然后基于该似然分布进行采样以得到滑窗图像对应的隐变量, 再将滑窗图像对应的隐变量输入解码器, 得到重建滑窗图像 $\hat{\mathbf{x}}_w^i$ 。可以将采样层输出的采样的隐变量 \mathbf{z}_w^i , 直接作为滑窗图像对应的隐变量, 也可以对采样层输出的采样的隐变量 \mathbf{z}_w^i 进行处理, 得到滑窗图像对应的隐变量。变分自动编码器还可以包括: 高斯过程回归模块, 对采样层输出的采样的隐变量 \mathbf{z}_w^i 进行高斯过程回归处理, 得到滑窗图像对应的隐变量。

[0076] 将多个滑窗图像自输入层依次输入隐藏层1, 隐藏层1的输出特征输入隐藏层2, 隐藏层2输出隐变量均值和隐变量方差, 每个滑窗图像得到若干个似然分布, 每个似然分布具有1个隐变量均值和1个隐变量方差, 基于似然分布进行采样, 得到采样的隐变量 \mathbf{z}_w^i (每个似然分布可以得到一个采样的隐变量), 以确定滑窗图像对应的隐变量。然后将滑窗图像对应的隐变量输入隐藏层3, 隐藏层3的输出特征输入隐藏层4, 隐藏层4输出 m 个重构均值 $\mu_{\hat{\mathbf{x}}_w^i}$ 和 m 个重构方差 $\sigma_{\hat{\mathbf{x}}_w^i}$, 基于 m 个重构均值 $\mu_{\hat{\mathbf{x}}_w^i}$ 得到滑窗图像对应的重建滑窗图像 $\hat{\mathbf{x}}_w^i$ 。

[0077] 在VAE算法中, 将生成图像 \mathbf{x} 视为通过隐变量 \mathbf{z} 生成的随机变量。根据贝叶斯定律, 可以通过以下方式获得隐变量 \mathbf{z} 的后验分布:

$$[0078] \quad p(\mathbf{z}|\mathbf{x}) = \frac{p(\mathbf{x}|\mathbf{z})p(\mathbf{z})}{p(\mathbf{x})}; \quad (1)$$

[0079] 其中, $p(\mathbf{z}|\mathbf{x})$ 表示后验分布, $p(\mathbf{x}|\mathbf{z})$ 表示似然分布, $p(\mathbf{z})$ 表示先验分布, 通常被配置为标准正态分布, $p(\mathbf{x})$ 表示边际似然分布。由于边际似然分布 $p(\mathbf{x})$ 和似然分布 $p(\mathbf{x}|\mathbf{z})$ 都

是未知的,则后验分布 $p(\mathbf{z}|\mathbf{x})$ 无法由计算直接得出。因此变分方法通过引入 $q(\mathbf{z}|\mathbf{x})$ 来逼近真正的后验 $p(\mathbf{z}|\mathbf{x})$ 。将 $q(\mathbf{z}|\mathbf{x})$ 引入 $p(\mathbf{x})$ 的对数边际似然,可得:

$$\begin{aligned} \log p(\mathbf{x}) &= \int q(\mathbf{z}|\mathbf{x}) \log p(\mathbf{x}) d\mathbf{z} \\ [0080] \quad &= \int q(\mathbf{z}|\mathbf{x}) \log \frac{p(\mathbf{x}, \mathbf{z}) q(\mathbf{z}|\mathbf{x})}{q(\mathbf{z}|\mathbf{x}) p(\mathbf{z}|\mathbf{x})} d\mathbf{z} ; \quad (2) \\ &= \int q(\mathbf{z}|\mathbf{x}) \log \frac{p(\mathbf{x}, \mathbf{z})}{q(\mathbf{z}|\mathbf{x})} d\mathbf{z} + KL(q(\mathbf{z}|\mathbf{x}) \parallel p(\mathbf{z}|\mathbf{x})) \end{aligned}$$

[0081] 其中, $q(\mathbf{z}|\mathbf{x})$ 表示引入的后验分布,具体可以是神经网络输出的用来逼近真实后验分布 $p(\mathbf{z}|\mathbf{x})$ 的后验分布, $p(\mathbf{x}, \mathbf{z})$ 表示联合分布, $KL(\cdot \parallel \cdot)$ 表示KL散度, \log 表示取对数操作。

[0082] 将等式右侧非负的KL散度省略后,则有:

$$\begin{aligned} \log p(\mathbf{x}) &\geq \int q(\mathbf{z}|\mathbf{x}) \log \frac{p(\mathbf{x}, \mathbf{z})}{q(\mathbf{z}|\mathbf{x})} d\mathbf{z} \\ [0083] \quad &\geq \int q(\mathbf{z}|\mathbf{x}) \log p(\mathbf{x}|\mathbf{z}) d\mathbf{z} + \int q(\mathbf{z}|\mathbf{x}) \log \frac{p(\mathbf{z})}{q(\mathbf{z}|\mathbf{x})} d\mathbf{z} ; \quad (3) \\ &\geq E_{q(\mathbf{z}|\mathbf{x})}[\log p(\mathbf{x}|\mathbf{z})] - KL(q(\mathbf{z}|\mathbf{x}) \parallel p(\mathbf{z})) \end{aligned}$$

[0084] 其中, $E_{q(\mathbf{z}|\mathbf{x})}(\cdot)$ 表示期望。

[0085] 余下不等式右侧的项被称为变分下届(Evidence Lower Bound, ELBO)。由于KL散度的非负性,最大化ELBO可以使得 $q(\mathbf{z}|\mathbf{x})$ 靠近 $p(\mathbf{z}|\mathbf{x})$ (最大化ELBO就是迫使KL趋近于0,当KL等于0意味着 $q(\mathbf{z}|\mathbf{x})$ 和 $p(\mathbf{z}|\mathbf{x})$ 一样,那么用 $q(\mathbf{z}|\mathbf{x})$ 接近 $p(\mathbf{z}|\mathbf{x})$ 的目的就达成了,也就是说式(2)中除了 $KL(q(\mathbf{z}|\mathbf{x}) \parallel p(\mathbf{z}|\mathbf{x}))$ 以外的其它项最大化,会自动逼迫 q 趋近于 p ,所以只要保证式(3)最大即可,这样被省略的 $KL(q(\mathbf{z}|\mathbf{x}) \parallel p(\mathbf{z}|\mathbf{x}))$ 会自动趋近于0从而,在这里假设 $\log p(\mathbf{x}|\mathbf{z})$ 不变,那么最大化ELBO就是挤占非负 $KL(q(\mathbf{z}|\mathbf{x}) \parallel p(\mathbf{z}|\mathbf{x}))$ 的值使之变为0。将上述公式结合到一个编码器-解码器结构中,也就是将从 \mathbf{x} 到 \mathbf{z} 的过程视为 $p(\mathbf{z}|\mathbf{x})$ (编码器),而将从 \mathbf{z} 重构 \mathbf{x} 的过程视为 $p(\mathbf{x}|\mathbf{z})$ (解码器),那么可以得到变分自编码器的损失函数为:

$$[0086] \quad \log p(\mathbf{x}) \geq E_{q_{\phi}(\mathbf{z}|\mathbf{x})}[\log p_{\theta}(\mathbf{x}|\mathbf{z})] - KL(q_{\phi}(\mathbf{z}|\mathbf{x}) \parallel p(\mathbf{z})); \quad (4)$$

[0087] 其中, ϕ 表示编码器的参数, θ 表示解码器的参数。

[0088] 编码器和解码器均采用神经网络,由于神经网络只能梯度下降,所以最大化上式(4)等效于最小化下面的相反式子。

$$[0089] \quad L = -E_{q_{\phi}(\mathbf{z}|\mathbf{x})}[\log p_{\theta}(\mathbf{x}|\mathbf{z})] + KL(q_{\phi}(\mathbf{z}|\mathbf{x}) \parallel p(\mathbf{z})); \quad (5)$$

[0090] 由于将目标图像形成了滑窗图像序列,则变分自动编码器的损失函数为:

$$[0091] \quad L = \sum_1^n \left\{ -E_{q_\phi(z_w^i|x_w^i)} [\log p_\theta(\hat{x}_w^i|z_w^i)] + KL(q_\phi(z_w^i|x_w^i) \parallel p(z_w^i)) \right\}; \quad (6)$$

[0092] 其中, L 表示损失函数, \sum 表示求和操作, n 表示滑窗图像的数量, $E_{q_\phi(z_w^i|x_w^i)}(\cdot)$ 表示期望, $p_\theta(\cdot|\cdot)$ 表示解码器, θ 表示解码器的参数, \hat{x}_w^i 表示第 i 个重建滑窗图像, z_w^i 表示第 i 个采样的隐变量, w 表示滑窗图像的宽度, x_w^i 表示第 i 个滑窗图像, \log 表示取对数操作, $KL(\cdot\parallel\cdot)$ 表示KL散度, $q_\phi(\cdot|\cdot)$ 表示编码器, ϕ 表示编码器的参数, $p(\cdot)$ 表示先验分布。

[0093] 具体地, 所述先验分布 $p(z)$ 被配置为标准正态分布, 得到的重建滑窗图像中每个像素由一个单独的正态分布表示, 存在均值 $\mu_{\hat{x}_w^i}$ 和方差 $\sigma_{\hat{x}_w^i}$ 。为方便起见, 使用均值 $\mu_{\hat{x}_w^i}$ 来重建图像, 得到重建滑窗图像 \hat{x}_w^i 。

[0094] 步骤S320具体包括:

[0095] 步骤S321、将所述似然分布输入采样层, 将所述采样层输出的采样的隐变量, 作为滑窗图像对应的隐变量。

[0096] 具体地, 如图1所示, 将滑窗图像 x_w^i 输入编码器, 编码器输出滑窗图像对应的似然分布, 然后基于该似然分布进行采样以得到采样的隐变量 z_w^i , 并将采样的隐变量 z_w^i 作为滑窗图像对应的隐变量。

[0097] 步骤S320具体包括:

[0098] 步骤S322、将所述似然分布输入采样层, 通过所述采样层输出采样的隐变量, 对所述采样的隐变量进行高斯过程回归, 得到所述滑窗图像对应的隐变量。

[0099] 具体地, 如图3所示, 将滑窗图像 x_w^i 输入编码器, 编码器输出滑窗图像对应的似然分布, 然后基于该似然分布进行采样以得到采样的隐变量 z_w^i , 并对采样的隐变量 z_w^i 进行高斯过程回归, 得到高斯过程回归后的均值向量 $\mu_{z_{gp}}^i$, 并将高斯过程回归后的均值向量 $\mu_{z_{gp}}^i$ 作为滑窗图像对应的隐变量。

[0100] 高斯过程回归后的均值向量为:

$$[0101] \quad \mu_{z_{gp}}^i = k(t_i, \mathbf{t})(K + \sigma_n^2 \mathbf{I})^{-1} \mathbf{z}_w^i;$$

$$[0102] \quad K = \begin{bmatrix} \phi(t_1, t_1) & \dots & \phi(t_1, t_n) \\ \dots & \dots & \dots \\ \phi(t_n, t_1) & \dots & \phi(t_n, t_n) \end{bmatrix};$$

$$[0103] \quad \phi(t, t') = \alpha \exp\left(-\frac{(t - t')^2}{2l^2}\right);$$

$$[0104] \quad \mathbf{t} = [t_1, t_2, \dots, t_i, \dots, t_n];$$

[0105] 其中, $\mu_{z_{gp}}^i$ 表示第 i 个高斯过程回归后的均值, $k(\cdot, \cdot)$ 表示核函数, \mathbf{t} 表示采样的隐变量的序号, K 表示核函数矩阵, t_i 表示第 i 个采样的隐变量的序号, $\phi(\cdot, \cdot)$ 表示径向基核函数, t, t' 分别表示采样的隐变量的序号 \mathbf{t} 中任意两个, α 和 l 是径向基核函数中的超参数, 分别表示径向基核函数的带宽和特征长度尺度, σ_n^2 表示采样的隐变量 z_w^i 的噪音, \mathbf{I} 表示单位

矩阵。

[0106] 当各个滑窗所提取的信息被压缩到低维空间,对采样的隐变量 \mathbf{z}_w^i 进行高斯过程回归,可以强制增强图像信息之间的连续性,使得滑窗之间过渡的更为平滑,更有利于强制忽略其中的异常信息。

[0107] 现有技术中高斯回归通常是对时间序列回归,建立每个时间点所对应值的相关性,尤其是相近时间点之间的平滑关系。本申请中在滑窗的作用下形成滑窗图像序列,并按照滑窗图像序列中滑窗图像对应的序号,依次将滑窗图像输入变分自动编码器,得到对应的采样的隐变量,即形成采样的隐变量的序列,每个采样的隐变量具有序号,通过高斯过程回归,使得各个滑窗图像对应的隐变量是更加平滑变化的过渡的,这样突变的裂缝就可以被自动鲁棒的充分舍弃。

[0108] 步骤S400具体包括:

[0109] 步骤S410、根据第一个滑窗图像对应的重建滑窗图像的上半部分图像、最后一个滑窗图像对应的重建滑窗图像的下半部分图像以及其余滑窗图像对应的重建滑窗图像的中间部分图像,组合形成重建图像;其中,所述中间部分图像的宽度根据所述预设窗口滑动的距离确定。

[0110] 具体地,由于重建滑窗图像有n个,即 $\hat{\mathbf{x}}_w^1, \hat{\mathbf{x}}_w^2, \dots, \hat{\mathbf{x}}_w^i, \dots, \hat{\mathbf{x}}_w^n$,其中, $\hat{\mathbf{x}}_w^i$ 表示第i个重建滑窗图像,各重建滑窗图像的均值 $\mu_{\hat{\mathbf{x}}_w^i}$ 之间有很多重叠,因此在通过所有重建滑窗图像组合形成重建图像时,一些重叠部分可以不予考虑。第一个重建滑窗图像(即第一个滑窗图像对应的重建滑窗图像)和最后一个重建滑窗图像(即最后一个滑窗图像对应的重建滑窗图像)均取一半图像,具体地,第一个重建滑窗图像取上半部分图像,最后一个重建滑窗图像取下半部分图像,其余重建滑窗图像取中间部分图像,最后根据上半部分图像、下半部分图像以及所有中间部分图像组成重建图像 $\hat{\mathbf{x}}$ 。中间部分图像的尺寸根据预设窗口滑动的距离确定,例如,预设窗口滑动的距离为1个像素时,中间部分图像为1行像素的均值 $\mu_{\hat{\mathbf{x}}_w^i}$,即重建滑窗图像的中间线。

[0111] 步骤S500具体包括:

[0112] 步骤S510、将所述目标图像中每个像素的像素值与所述重建图像对应像素的像素值作差,得到差值图像。

[0113] 步骤S520、当所述差值图像中的连续多个像素的像素值大于预设像素值时,将所述目标图像中对应的像素确定为裂缝。

[0114] 具体地,重建图像与目标图像极为相似,由于重建图像中不存在裂缝,而目标图像中存在裂缝,两者的差异之处在于裂缝。重建图像与目标图像形成差值图像后,差值图像中裂缝区域之外的像素值为0或者趋近于0(小于或等于预设像素值),而差值图像中裂缝区域的像素值大于预设像素值,因此,可以根据差值图像中像素值确定裂缝。

[0115] 此外,用于检测高速铁路接触网系统旋转双耳中的裂缝。旋转双耳组件的灰度图像如图1所示,重建图像中裂缝被自动删除,从而可以识别出旋转双耳组件的裂缝。本申请使用滑动窗口将图像信息转换为时间序列数据,从而将目标图像中的裂缝视为序列中的异常值。由于SW-VAE的鲁棒性,这些序列化图像中的裂缝信息被视为异常值并自动删除。因此,可以通过将SW-VAE的重建图像与目标图像进行比较来揭示裂缝。

[0116] 本申请SW-VAE的优点是不需要使用包含裂缝的数据集进行模型训练,也不需要模型通过没有裂缝信息的正样本来学习图像的正常模式。因此,当训练数据集稀缺时,SW-VAE为裂纹检测提供了一个可行的选择。极端条件下只用一个单独的图像数据,其中,裂缝依然可以被识别。

[0117] 应当理解的是,本发明的应用不限于上述的举例,对本领域普通技术人员来说,可以根据上述说明加以改进或变换,所有这些改进和变换都应属于本发明所附权利要求的保护范围。

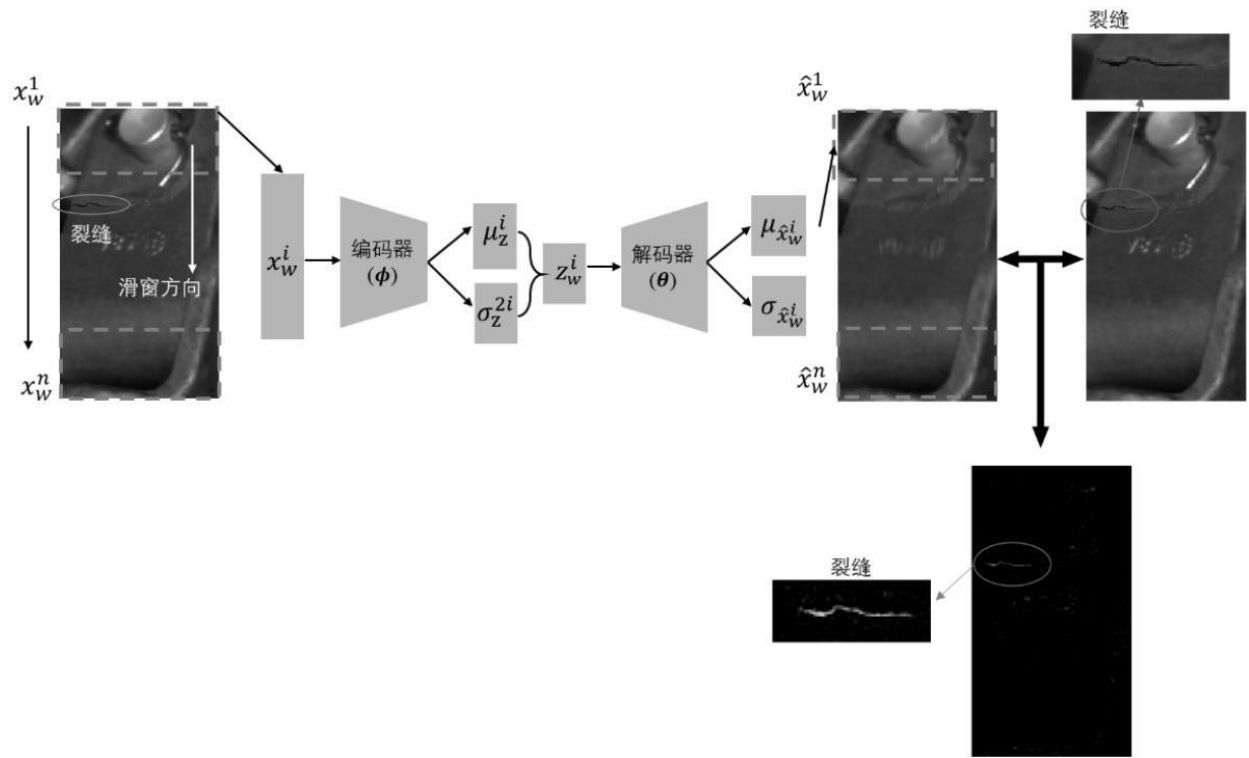


图 1

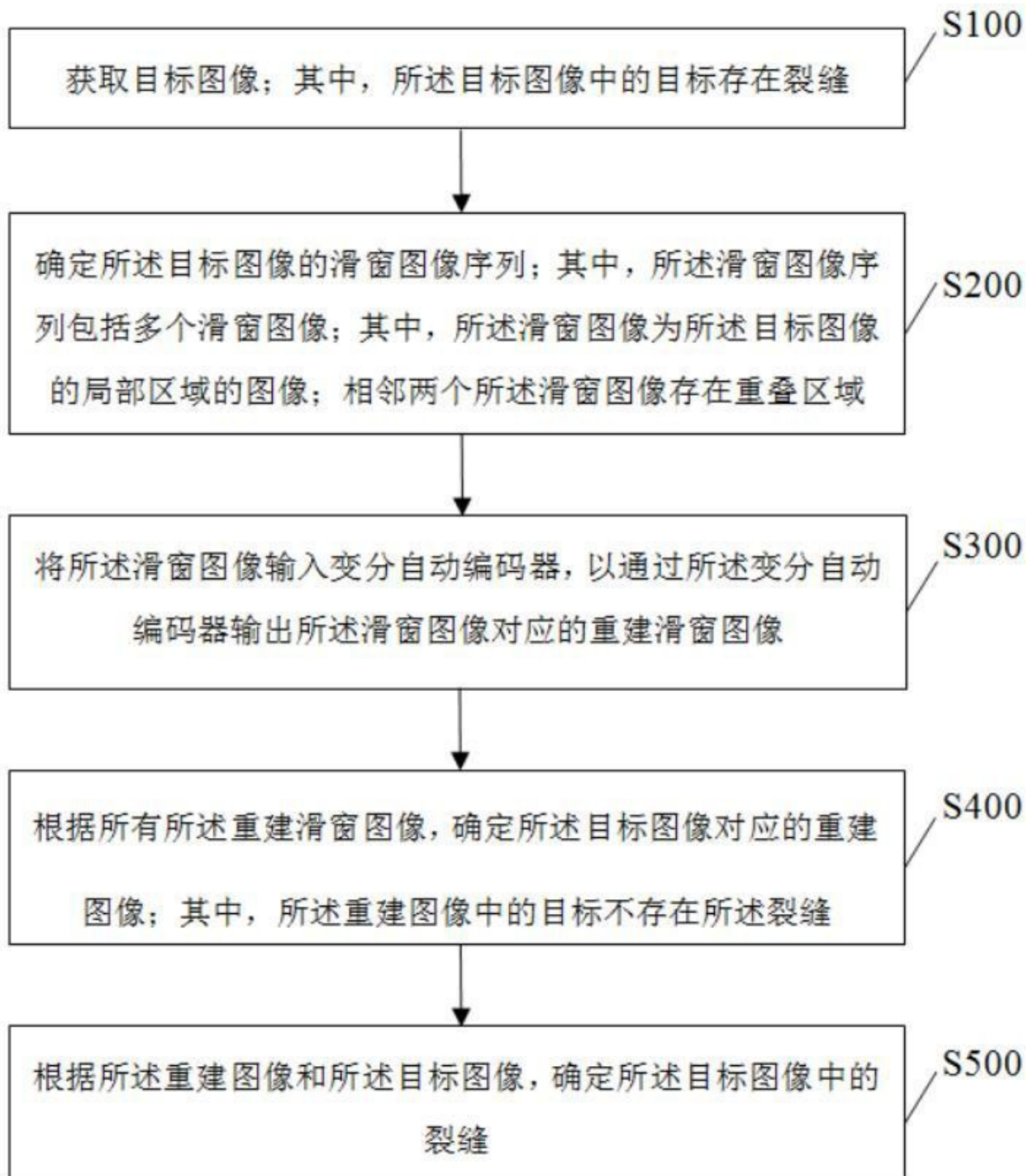


图 2

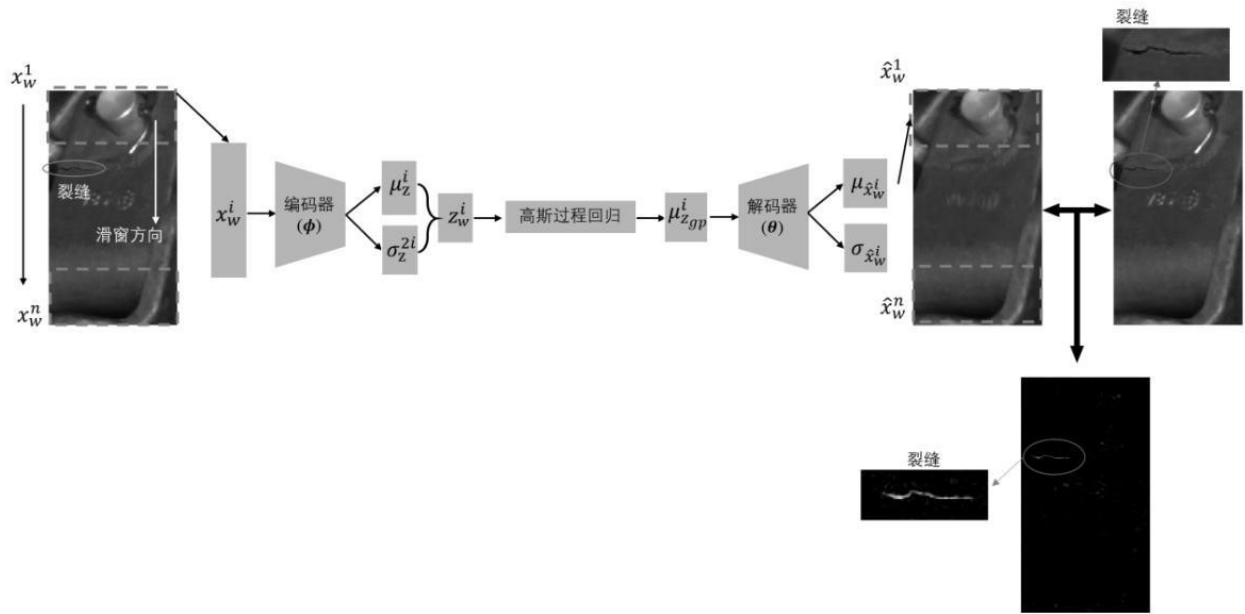


图 3