One-class Damage Detector Using Fully-Convolutional Data Description for Prognostics

Takato Yasuno[*1], Masahiro Okano[*1], Riku Ogata[*1], Junichiro Fujii[*1]

[*1] Yachiyo Engineering, Co., Ltd., RIIPS.

It is important for infrastructure managers to maintain a high standard to ensure user satisfaction during a lifecycle of infrastructures. Surveillance cameras and visual inspections have enabled progress toward automating the detection of anomalous features and assessing the occurrence of the deterioration. Frequently, collecting damage data constraints time consuming and repeated inspections. One-class damage detection approach has a merit that only the normal images enables us to optimize the parameters. Simultaneously, the visual explanation using the heat map enable us to understand the localized anomalous feature. We propose a civil-purpose application to automate one-class damage detection using the fully-convolutional data description (FCDD). We also visualize the explanation of the damage feature using the up-sampling-based activation map with the Gaussian up-sampling from the receptive field of the fully convolutional network (FCN). We demonstrate it in experimental studies: concrete damage and steel corrosion and mention its usefulness and future works.

1. Introduction

1.1 Related Works for Deep Anomaly Detection

There are anomaly detection approaches based on the shallow feature map: OC-SVM [Chalapathy 2018], SVDD [Tax 2014], GMM [Theodoridis 2020], PCA [Hawkins 1974], and kernel-PCA [Hoffmann 2007]. On the other hand, there are anomaly detection approaches based on the deep feature map: Deep SVDD [Ruff 2018], FCDD [Liznerski 2021], VAEs [Kingma 2019], and AAEs [Zhou 2017]. These are divided into three categories: classification-based models, probabilistic models, and reconstruction-based models [Ruff 2020]. The reconstruction-based models are not always able to well-reconstruct the synthetic outputs, because of the weakness towards background noise. The probabilistic models requires assumptions regularized on the probability distribution. The classification-based models depends on neither synthetic reconstruction nor probabilistic assumption, so that it could be more robust anomaly detector.

We propose a civil-purpose application to automate one-class damage detection using the fully-convolutional data description (FCDD). We also visualize the explanation of the damage feature using the Gaussian up-sampling from the receptive field of the fully convolutional network (FCN). As illustrated in Figure 1, we give an overview for infrastructure damage classification using FCDD and up-sampling based heatmap explanation.

2. Damage Detection and Explanation Method

2.1 One-class Damage Classification via FCDD

Let X_i the i-th image with the size of $h \times w$, and let c the center of hyper-sphere boundary between inlier normal region and outlier anomalous region. Denote n the number of training images, and denote W the weight of FCN. Then, the Deep SVDD objective function [Ruff 2018] is formulated as next minimization problem the deep support vector data description.

Contact: Takato Yasuno, RIIPS on 3F, 5-20-8, Asakusabashi, Taito-ku, Tokyo, 111-8648, tk-yasuno@yachiyo-eng.co.jp.

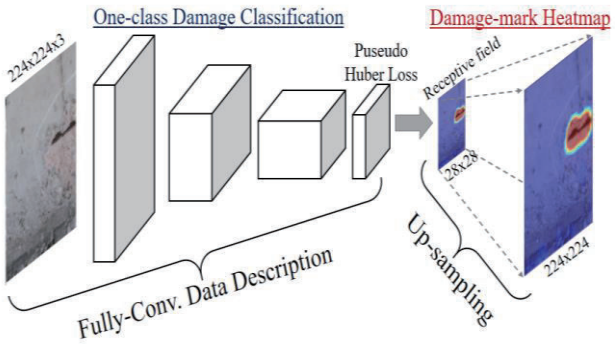

Figure 1: Overview of civil-purpose damage classification using Fully-Convolutional Data Description, damage-mark heatmap

$$\min_{w} \frac{1}{n} \sum_{i=1}^{n} \|\varphi_w(X_i) - c\|^2 \tag{1}$$

Here, denote $\varphi_w(X_i)$ mapping of FCN from input image. And the one-class classification model is formulated as follows, using the cross-entropy loss function [Ruff 2018].

$$\mathcal{L}_{DeepSVDD} = -\frac{1}{n}\sum_{i=1}^{n}(1-y_i)\log\ell\big(\varphi_w(X_i)\big) + y_i \log\big[1-\ell\big(\varphi_w(X_i)\big)\big] \tag{2}$$

Here, denote that $y_i = 1$ is i-th image's anomalous label and that $y_i = 0$ is i-th image's normal label. For more robust loss formulation, the Pseudo Huber loss function is introduced in (2).

Let $\ell(z)$ the loss function and denote the Pseudo Huber loss as:

$$\ell(z) = \exp(-H(z)), \quad H(z) = \sqrt{\|z\|^2 + 1} - 1$$

$$(2) = -\frac{1}{n}\sum_{i=1}^{n}(1-y_i)\,H\big(\varphi_w(X_i)\big) + y_i \log[1 - \exp\{-H(\varphi_w(X_i))\}] \tag{3}$$

Therefore, the FCDD loss function is formulated as:

$$\mathcal{L}_{FCDD} = \frac{1}{n}\sum_{i=1}^{n}(1-y_i)\frac{1}{uv}\sum_{x,y}H_{x,y}(\varphi_w(X_i)) - y_i \log\left[1-\exp\left\{-\frac{1}{uv}\sum_{x,y}H_{x,y}(\varphi_w(X_i))\right\}\right] \quad (4)$$

Here, denote $H_{x,y}(z)$ elements (x, y) of the receptive field with the size of $u \times v$ under the FCN [Liznerski 2021].

2.2 Heatmap Upsampling from Receptive Field

Convolutional neural network (CNN) models with millions of shared parameters achieve satisfactory performance for anomaly detection. Despite the impressive performance, the reasons for it remain unclear. This visualized heatmap technique was mainly divided into a masked sampling and an activation map approaches. The former includes the occlusion sensitivity [Zeiler 2013] and the local interpretable model-agnostic explanations (LIME) [Ribeiro 2016]. The merits of this approach do not require in-depth knowledge of the network's architecture, but the disadvantage is that it requires iterative computations per image, and running time for local partitioning, masked sampling, and output prediction. While, the latter approach includes an activation map such as the class activation map (CAM) [Zhou 2015] and gradient-based extension (Grad-CAM) [Selvaraju 2017]. The weighting feature map of the CAM is ineffective as it limits the global average pooling (GAP) and fully-connected (FC) at the concluding layer of a CNN. The merits of gradient approach is able to apply any layer structure of CNNs, so it has significant applicability. But the disadvantage is that it requires parallel computation resource and moderate running time for gradient-based heatmap.

For civil-purpose application, we select the receptive field up-sampling approach [Liznerski 2021] to visualize the explanation of the anomalous damage feature using the up-sampling-based activation map with the Gaussian up-sampling from the receptive field of the fully convolutional network. The merits of up-sampling approach require less computation resource and faster running time. This up-sampling algorithm creates an output of full-resolution anomaly heatmap from an input of low-resolution receptive field u×v. From our experiments of datasets, we set the size of receptive field 28×28, as a practical value. To visualize a damage heatmap, we have to unify a display range corresponding to anomaly score from minimum to maximum value. To strengthen the damage region as *damage-mark*, we can set a display range [min. max./4], whose quartile parameter is 0.25. This is reason that the histogram of anomaly score has long-tail shape and if we set all of anomaly score range, the color would be weaken to blue or yellow on the maximum side.

3. Applied Results

3.1 Damage Datasets in Civil Engineering

As shown in Table 1, we demonstrate a prototype method in experimental studies: pavement crack from the SDNET dataset [Dorafshan 2018], bridge of rebar exposure and steel paint peeling, volt nut corrosion, and dam embankment of janka.

Table 1: Damage datasets of inspection for road, bridge, and dam.

dataset	Patch size	Normal	Anomalous
SDNET Pavement crack	256^2	3,500	1,826
Bridge rebar exposure	224^2	306	230
Bridge steel corrosion	64^2	2,400	579
Dam exfoliation, janka	256^2	1,075	247

3.2 Training Damage Detector and Accuracy

The input size is unified to 224^2 while training. We set the mini-batch to 30 and the number of epoch to 50. We used the Adam optimizer with a learning rate of 0.0001, and set the gradient decay factor to 0.9, and the squared gradient decay factor to 0.99. Training images are partitioned so as to set the ratio 7:1:2 correspond to the number of training, calibration, and test objective. As shown in the Table 2, we build the fully-convolutional network as the initial backbone with 27 layers and learnable parameters 4.4 million. We call the CNN27 that contains either Conv-BN-ReLU or Maxpool. This initial FCN for a prototype detector has neither skip layer nor residual layer. Table 3 shows the accuracy of one-class damage detection applied to the damage dataset for road, bridge, and dam. The AUC and recall are significantly high values. This suggests that the FCDD is applicable for civil-purpose in damage inspection.

Table 2: Layer type and shape of CNN27 architecture

No	Layer type	Output Shape (S,S,C)	kernel	Params #
1	Input	224,224,3		--
2-4	Conv1-BN-Relu1	224,224,64	3	1,792
5	Maxpool1	112,112,64		--
6-8	Conv2-BN-Relu2	112,112,128	3	73,856
9	Maxpool2	56,56,128		--
10-12	Conv3-BN-Relu3	56,56,256	3	295,168
13-15	Conv4-BN-Relu4	56,56,256	3	295,168
16	Maxpool3	28,28,256		--
17-19	Conv5-BN-Relu5	28,28,512	3	1,180,160
20-22	Conv6-BN-Relu6	28,28,512	3	1,180,160
23-25	Conv7-BN-Relu7	28,28,512	3	1,180,160
26	Conv8	28,28,512	1	264,192
27	Pseudo Huber loss			
	total	Learnables		4.4M

Table 3: Accuracy of damage detection for road, bridge, and dam

dataset	**AUC**	F1	precision	**recall**
SDNET Pavement crack	**0.8955**	0.7104	0.6209	**0.8301**
Bridge rebar exposure	**0.9649**	0.9052	0.8775	**0.9347**
Bridge steel corrosion	**0.9889**	0.8803	0.7972	**0.9827**
Dam exfoliation, janka	**0.9249**	0.7831	0.7469	**0.8231**

3.3 Damage-mark Heatmap

We visualized the explanation of the damage feature using the Gaussian up-sampling from the receptive field of our CNN27. We also created the histogram of anomaly score on test images.

Figure 3 depicts that each heatmap enables to visualize the crack region of interest to succeed damage-mark explanation. Figure 4 draws that there are three overlapped bins in horizontal anomaly score, because of shadow r.w.t. crack and formwork.

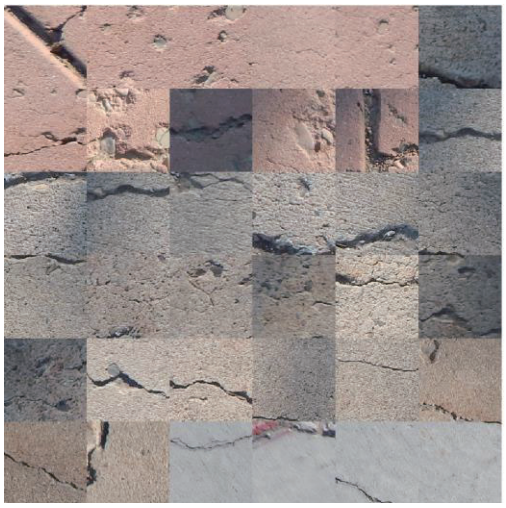

Figure 2: Input images of SDNET pavement crack

Figure 5: Input images of bridge rebar exposure

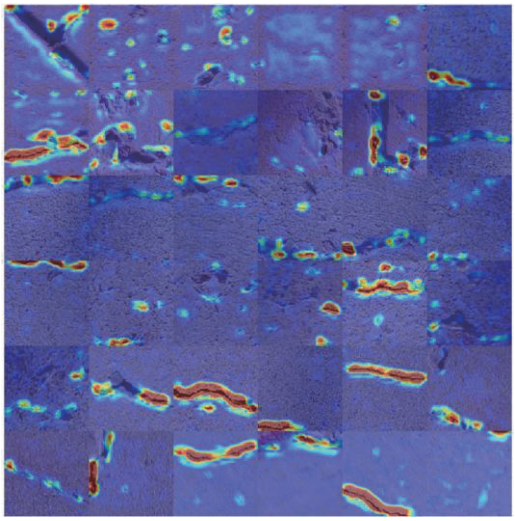

Figure 3: Heatmaps using FCDD applied to pavement crack.

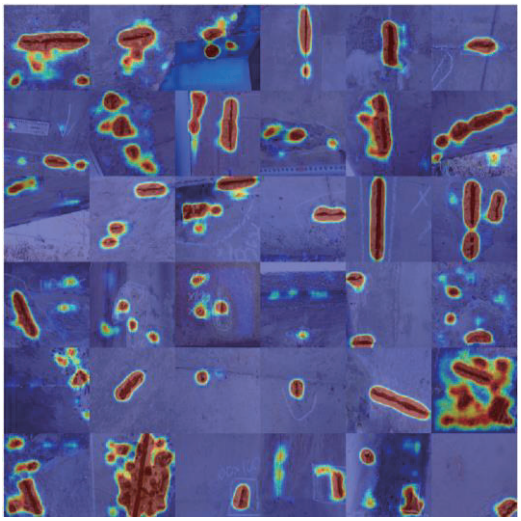

Figure 6: Heatmaps using FCDD applied to bridge rebar exposure

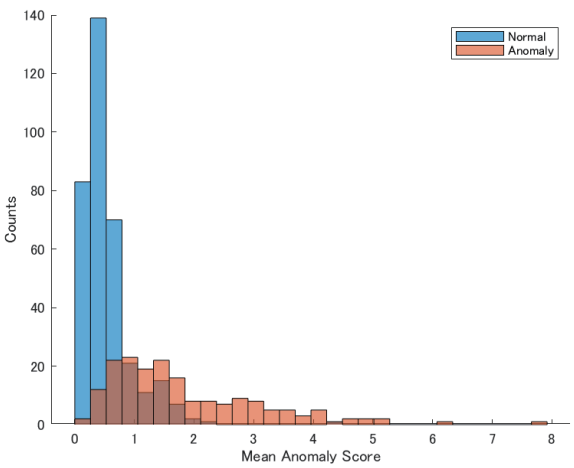

Figure 4: Anomaly score histogram via FCDD to pavement crack.

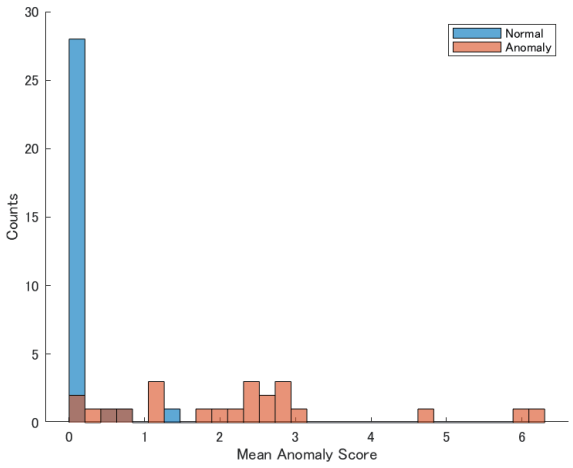

Figure 7: Anomaly score histogram via FCDD to bridge rebar exposure

Figure 6 depicts that each heatmap enables to visualize the rebar exposure with either large or small region to succeed the damage-mark explanation. Figure 7 draws that there is few overlapped bin in horizontal anomaly score, so the score range are well separated for rebar exposure detection.

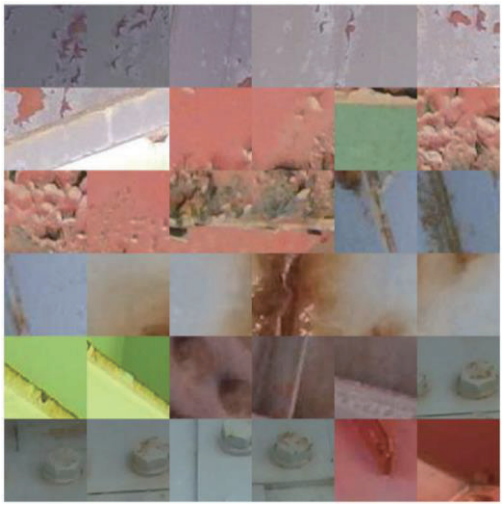

Figure 8: Input images of bridge steel corrosion, paint peeling

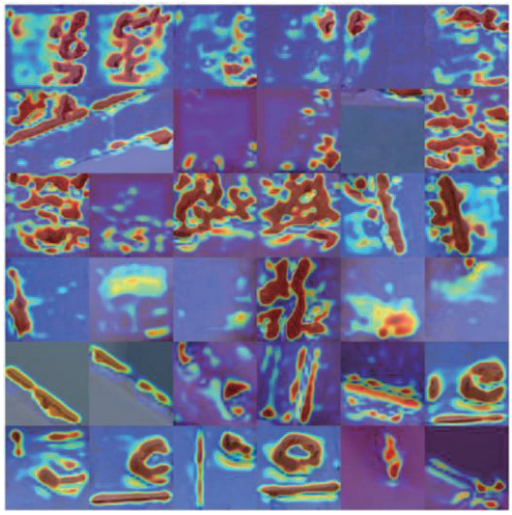

Figure 9: Heatmaps using FCDD applied to corrosion, paint peeling

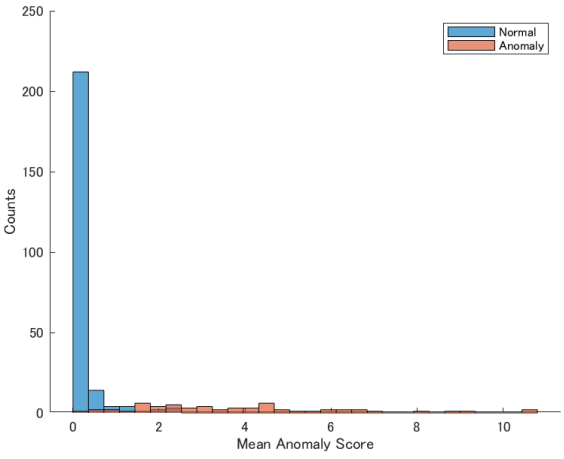

Figure 10: Anomaly score histogram FCDD to corrosion, paint peeling

Figure 9 depicts that each heatmap enables to visualize the paint peeling and volt nut corrosion to succeed the damage-mark explanation. Figure 10 draws that there is few overlapped bin in horizontal anomaly score, so the score range are well separated for steel paint peeling and corrosion detection.

Figure 11: Input images of dam surface janka

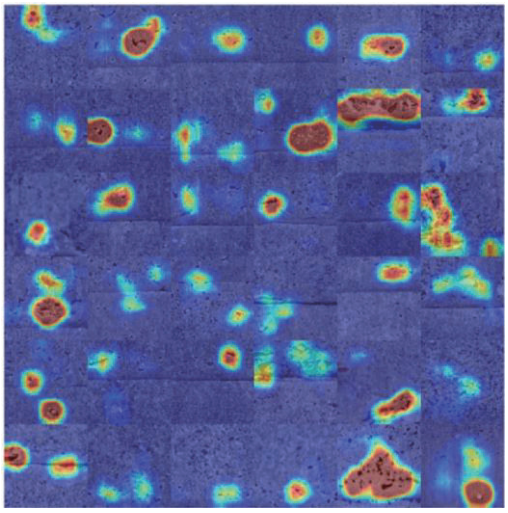

Figure 12: Heatmaps using FCDD applied to dam surface janka

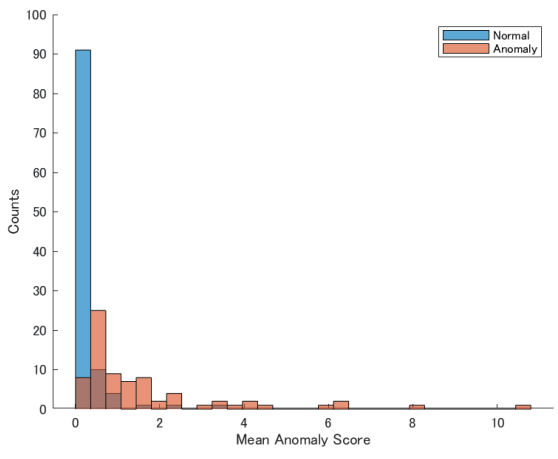

Figure 13: Anomaly score histogram via FCDD to dam surface janka

Figure 12 depicts that each heatmap enables to visualize the janka on surface of dam embankment to succeed the damage-mark explanation. Figure 13 draws that there are three overlapped bins in horizontal anomaly score, because it is visually unclear to separate janka feature and health concrete.

4. Concluding Remarks

We built a civil-purpose application to automate one-class damage detection using the FCDD with a light backbone CNN network with 27 layers that contains either Conv-BN-ReLU or maxpooling. We also visualized the damage-mark heatmap using the Gaussian up-sampling from the receptive field of FCN, directly. We demonstrated our prototype in four experimental studies: concrete pavement crack, bridge component's rebar exposure, steel corrosion, and dam embankment of exfoliation and janka. These test accuracy has high value of index AUC and recall, so that the FCDD is applicable for infrastructure damage inspection. Without any annotation toward damage region, the FCDD enable to enhance *damage-mark* for visual explanation.

Furthermore, we mention the future works for more accurate and robust application. For more accurate damage detection, it remains ablation opportunities for deeper backbone candidates: VGG16, ResNet101, Inceptionv3. For more robust training toward background noise, augmentation preprocess could be effective for one-class classification models: e.g. mixup, RICAP, cutout, and random erasing. For more flexible applicability, we have various experiment opportunities to another datasets: natural disaster e.g. typhoon, earthquake, hurricane acquired by drone, aero-photograph, and satellite imagery.

Acknowledgments The authors wish to thank the MathWorks, Takuji Fukumoto who provided us optional resources.

References

[Chalapathy 2018] R. Chalapathy, A. K. Menon, and S. Chawla, "Anomaly Detection Using One-class Neural Networks," *arXiv:1802.06360*, 2018.

[Tax 2014] D. M. J. Tax and R. P. Duin, "Support Vector Data Description," *Machine Learning*, vol. 54, no. 1, pp. 45–66, 2004.

[Theodoridis 2020] S. Theodoridis, *Machine Learning: Bayesian and Optimization Perspective, 2nd ed*. Academic Press, 2020.

[Hawkins 1974] D. M. Hawkins, "Detection of Errors in Multivariate Data Using Principal Components," *J. of American Statistical Association*, vol. 69, no. 346, pp. 340–344, 1974.

[Hoffmann 2007] H. Hoffmann, "Kernel PCA for Novelty Detection," *Pattern Recognition*, vol. 40, pp. 863–874, 2007.

[Ruff 2018] L. Ruff, R. A. Vandermeulen, N. Gornitz, L. Deecke, S. A. Siddiqui, A. Binder, E. M¨uller, and M. Kloft, "Deep One-class Classification," in *International Conference on Machine Learning*, vol. 80, pp. 4390–4399, 2018.

[Kingma 2019] D. P. Kingma and M. Welling, "An Introduction to Variational Autoencoders," *Foundations and Trends in Machine Learning*, vol. 12, no. 4, pp. 307–392, 2019.

[An 2015] J. An and S. Cho, "Variational Autoencoder Based Anomaly Detection Using Reconstruction Probability," *Special Lecture on IE*, vol. 2, pp.1–18, 2015.

[Zhou 2017] C. Zhou and R. C. Paffenroth, "Anomaly Detection with Robust Deep Autoencoders," in *International Conference on Knowledge Discovery & Data Mining*, 2017, pp. 665–674.

[Zeiler 2013] Zeiler, M.D. et al., "Visualizing and Understanding Convolutional Networks", *arXiv:1311.2901v3.*

[Ruff 2020] L.Ruff, J.R. Kauffmann, R.A. Vandermeulen, "A Unifying Review of Deep and Shallow Anomaly Detection", *arXiv:2009.11732v3*, 2020.

[Ribeiro 2016] Ribeiro, M.T., S. Sameer, "Why Should I Trust You? Explaining the Predictions of Any Classifier", *Knowledge Discovery and Data Mining*, *KDD* 2016.

[Zhou 2015] Zhou, B., Khosla, A., et al., "Learning Deep Features for Discriminative Localization", *arXiv:1512.04150v1*.

[Selvaraju 2017] Selvaraju, R.R., Cogswell M. et al., "Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization", *ICCV*, 2017.

[Liznerski 2021] P. Liznerski, L. Ruff, et al., "Explainable Deep One-Class Classification", *ICLR2021*.

[Dorafshan 2018] Dorafshan, S., Thomas, R.J. et al. "SDNET2018: An Annotated Image Dataset For Noncontact Concrete Crack", https://doi.org/10.1016/j.dib.2018.11.015, Accessed 15/12/2022.