Deep Vision in Smart Manufacturing: MODERN Framework for Intelligent Quality Monitoring and Diagnosis

Yicheng Kang

Farmer School of Business
Miami University

June 18, 2024

This is joint work with Y. Jiao (WHU), X. Geng (U of Miami) and M. Nagarajan (UBC).

Background & Motivation

Today's manufacturing often comes with high degree of automation using robotics, computer vision and sensor technology - an example of smart manufacturing in Industry 4.0.

Image-based SPC harnesses the power of smart manufacturing. It replaces human operators with the following benefits:

- accurately captures product specifics - product dimension, surface finish and spatial patterns
- eliminates human subjectivity and ensures consistency
- able to keep up with high production rates
- avoid labor cost

Image-Based Quality Control

Images are increasingly used in industrial quality control.

- Anomaly detection of rolling processes in the steel industry (Feng and Qiu 2018).
- Surface grading in the ceramic tile industry (López et al. 2008; Koosha et al. 2017).
- Dimension control in the web production industry (Lyu and Chen, 2009).
- Inspection of surface uniformity in the liquid crystal display industry (Jiang et al. 2005).

Existing Methods

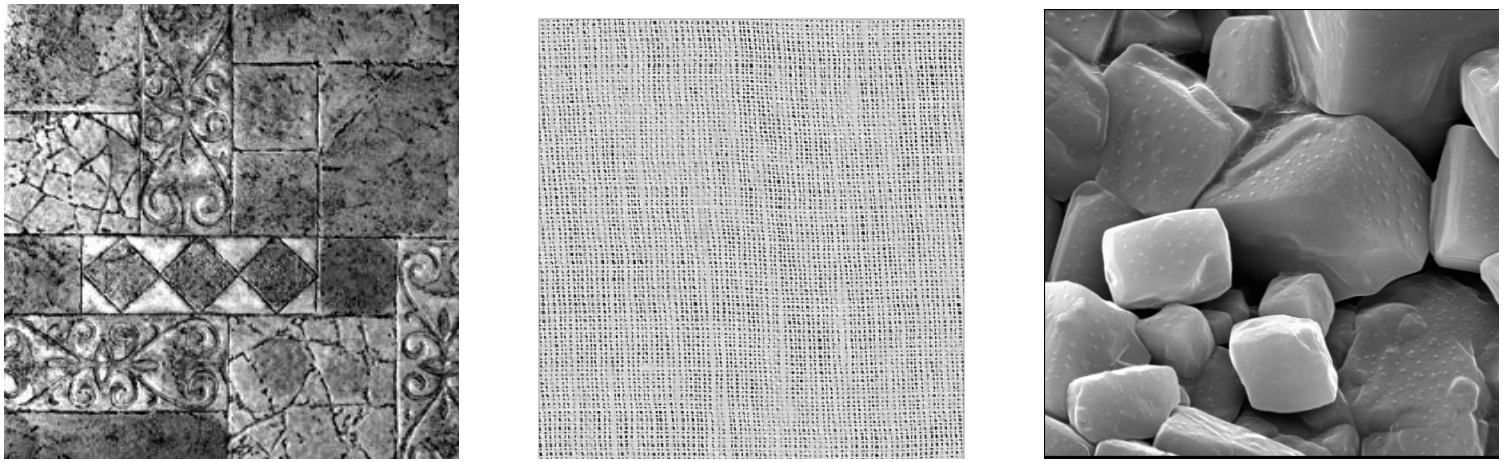


Figure 1: An illustration for the existing image-based SPC. Left: A tile surface. Middle: A textile fabric sample. Right: Microscopic structures of salt.

- There exists a gold standard e.g., the tile image, (Lin, 2007; Megahed et al., 2012; Koosha et al., 2017; Feng and Qiu, 2018).
- Two images cannot be matched pixel-to-pixel (e.g., the textile image) but satisfy the Markov Field (MF) assumption: stationarity and locality (Bui and Apley, 2018).
- The salt image doesn't meet the MF assumption.

Our Contribution

- We propose MODERN – Monitoring with Deep Residual Network – a framework for CNN-based SPC without making restrictive assumptions.
  - MODERN-Chart
  - MODERN-Diagnosis
- Transfer monitoring – extending pretrained MODERN-Chart to different production settings.
- Minimax asymptotic optimality and its implied managerial interpretation.

Deep CNN Architecture

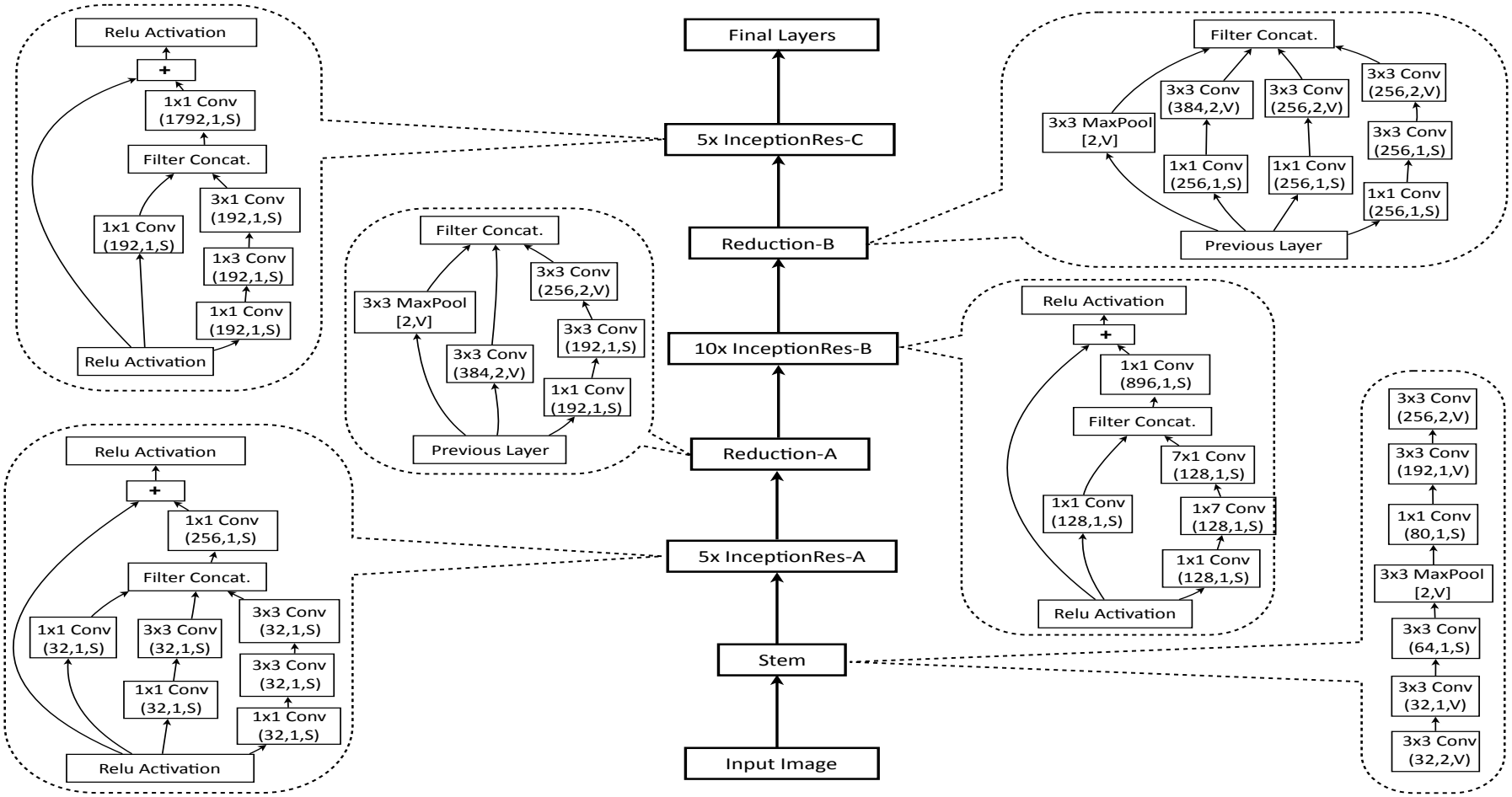


Figure 2: Architecture schematics of inception net with residual connections (Szegedy et al. 2017, 23.4m parameters).

Charting Statistics

Let $\{\widehat{p}_t = \mathscr{F}(\boldsymbol{X}_t; \widehat{\mathcal{V}}) : t \geq 1\}$. The charting statistics is defined by

$$E_t = \max\left\{0, \lambda(\widehat{p}_t - \mu_{\texttt{IC}}) + (1-\lambda)E_{t-1}\right\},$$

where $E_0 = 0$, $\lambda \in (0,1)$ is a weighting parameter, and $\mu_{\texttt{IC}}$ is the mean of $\widehat{p}_t$ when the process is IC. A signal of upward shift should be sent if

$$E_t > \rho\sqrt{\frac{\lambda}{2-\lambda}\sigma^2_{\texttt{IC}}},$$

where $\rho > 0$ is a parameter chosen to achieve a given ARL_0. To estimate $\mu_{\texttt{IC}}$ and $\sigma^2_{\texttt{IC}}$, let $\mathcal{C}_0 = \{\boldsymbol{X}_k^{\texttt{IC}} : k = 1, \cdots, n_{\texttt{IC}}\}$.

$$\widehat{\mu}_{\texttt{IC}} = \frac{1}{n_{\texttt{IC}}}\sum_{k=1}^{n_{\texttt{IC}}} \mathscr{F}\left(\boldsymbol{X}_k^{\texttt{IC}}; \widehat{\mathcal{V}}\right), \qquad \widehat{\sigma}^2_{\texttt{IC}} = \frac{1}{n_{\texttt{IC}}-1}\sum_{k=1}^{n_{\texttt{IC}}}\left(\mathscr{F}\left(\boldsymbol{X}_k^{\texttt{IC}}; \widehat{\mathcal{V}}\right) - \widehat{\mu}_{\texttt{IC}}\right)^2.$$

Control Limit

Augmented bootstrap (AB) algorithm:

- Set $b = 1$ and repeat:
  - randomly draw with replacement an IC image $\boldsymbol{X}_b$;
  - randomly apply an image transformation (rotate by 0, 90, 180, 270 degrees or flip horizontally or vertically), resulting in $\tilde{\boldsymbol{X}}_b$;
  - compute $E_b = \max\left\{0, \lambda\left(\mathscr{F}\left(\tilde{\boldsymbol{X}}_b; \widehat{\mathcal{V}}\right) - \widehat{\mu}_{\mathtt{IC}}\right) + (1-\lambda)E_{b-1}\right\}$;
  - $b \leftarrow b + 1$.

  until $E_b \geq \rho\sqrt{(\lambda\widehat{\sigma}^2_{\mathtt{IC}})/(2-\lambda)}$.
- Let $\mathrm{RL}(r) = b - 1$.
- Compute $\mathrm{ARL}(\rho) = 1/R\sum_{r=1}^{R}\mathrm{RL}(r)$.

Diagnosis: Faulty Region Estimation

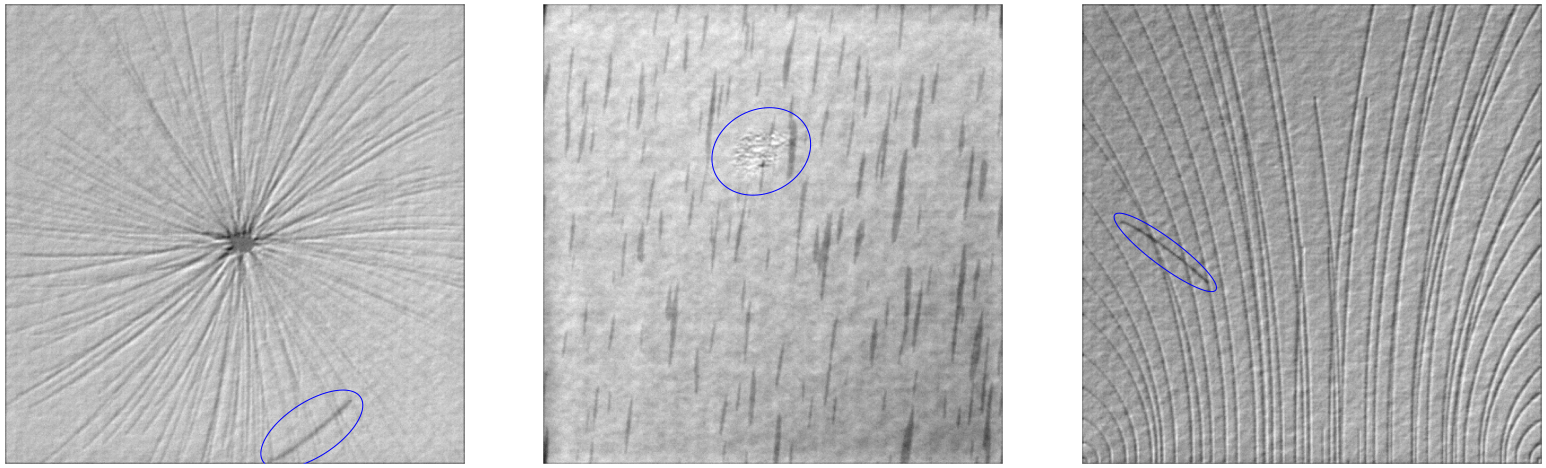


Figure 3: Examples of OC images.

Altering the neural net's final layer so that it minimizes the L_1 loss:

$$\frac{1}{n_{\mathsf{OC}}}\sum_{i=1}^{n_{\mathsf{OC}}} L_1\left(\boldsymbol{z}_i, \mathscr{G}\left(\boldsymbol{X}_i^{\mathsf{OC}};\mathcal{U}\right)\right) = \frac{1}{n_{\mathsf{OC}}}\sum_{i=1}^{n_{\mathsf{OC}}} \boldsymbol{w}'\left|\mathscr{G}\left(\boldsymbol{X}_i^{\mathsf{OC}};\mathcal{U}\right) - \boldsymbol{z}_i\right|,$$

where $\boldsymbol{z}_i$ is the five parameters of the ellipse in $\boldsymbol{X}_i^{\mathsf{OC}}$, and $\boldsymbol{w} = (w_1, \cdots, w_5)'$ are pre-specified weights.

Transfer Monitoring: Applicability Test

Let $\mathcal{S}_0 = \{\boldsymbol{S}_1^{\mathtt{IC}}, \cdots, \boldsymbol{S}_{m_0}^{\mathtt{IC}}\}$ and $\mathcal{S}_1 = \{\boldsymbol{S}_1^{\mathtt{OC}}, \cdots, \boldsymbol{S}_{m_1}^{\mathtt{OC}}\}$ be the set of IC and OC images in a new setting.

$$\widehat{\mathcal{P}}_0 = \left\{\mathscr{F}\left(\boldsymbol{S}_k^{\mathtt{IC}}; \widehat{\mathcal{V}}\right) : 1 \le k \le m_0\right\}, \quad \widehat{\mathcal{P}}_1 = \left\{\mathscr{F}\left(\boldsymbol{S}_k^{\mathtt{OC}}; \widehat{\mathcal{V}}\right) : 1 \le k \le m_1\right\}.$$

$$\overline{\mathcal{P}}_0 = \frac{1}{m_0}\sum_{k=1}^{m_0} \mathscr{F}(\boldsymbol{S}_k^{\mathtt{IC}}; \widehat{\mathcal{V}}), \ \overline{\mathcal{P}}_1 = \frac{1}{m_1}\sum_{k=1}^{m_1} \mathscr{F}(\boldsymbol{S}_k^{\mathtt{OC}}; \widehat{\mathcal{V}}),$$

$$\widehat{\sigma}_{\mathtt{app}}^2 = \frac{\sum_{k=1}^{m_0}\left(\mathscr{F}(\boldsymbol{S}_k^{\mathtt{IC}}; \widehat{\mathcal{V}}) - \overline{\mathcal{P}}_0\right)^2 + \sum_{k=1}^{m_1}\left(\mathscr{F}(\boldsymbol{S}_k^{\mathtt{OC}}; \widehat{\mathcal{V}}) - \overline{\mathcal{P}}_1\right)^2}{m_0 + m_1 - 2},$$

$$t_{\mathtt{app}} = \frac{\overline{\mathcal{P}}_1 - \overline{\mathcal{P}}_0}{\sqrt{\widehat{\sigma}_{\mathtt{app}}^2\left(\frac{1}{m_0} + \frac{1}{m_1}\right)}}.$$

Transfer Monitoring: Control Limit Redetermination

The applicability test only measures the separability of $\widehat{\mathcal{P}}_1$ from $\widehat{\mathcal{P}}_0$. The distribution of $\{\mathscr{F}(\boldsymbol{X}_k^{\mathtt{IC}}; \widehat{\mathcal{V}}) : \boldsymbol{X}_k \in \mathcal{C}_0\}$ in the old setting and the distribution of $\{\mathscr{F}(\boldsymbol{S}_k^{\mathtt{IC}}; \widehat{\mathcal{V}}) : \boldsymbol{S}_k \in \mathcal{S}_0\}$ in the new setting may be quite different.

- Recalculate the control limit using augmented bootstrap with $\mathcal{C}_0$ replaced by $\mathcal{S}_0$.
- Only need a few hundred images to achieve a commonly used ARL_0 values (tens of thousands would be needed to retrain the neural net sufficiently).

Asymptotic Properties - Likelihood Estimation

Let $f^*(\boldsymbol{X})$ be the log-odds of the OC probability given image $\boldsymbol{X}$, and $\mathcal{D}$ and $\mathcal{W}$ be the depth and width of the neural net.

Theorem 1. Assume the following conditions hold: (i) $f^*(x)$ is Lipschitz continuous; (ii) $\mathcal{DW} = \mathcal{O}(N^{d/(2d+4)})$; and (iii) $(\mathcal{DW})^2 \log(\mathcal{DW}) < Ne^{-1}$, where $N = n_{\texttt{IC}} + n_{\texttt{OC}}$ is the training sample size. Then we have

$$\mathbb{E}\left[\left\|\log\left(\frac{\mathscr{F}(\cdot;\widehat{\mathcal{V}})}{1-\mathscr{F}(\cdot;\widehat{\mathcal{V}})}\right) - f^*\right\|^2_{L^2(\nu)}\right] = \mathcal{O}\left(N^{-2/(2+d)}\log^{3/2} N\right).$$

Remark. The convergence rate is minimax optimal.

Asymptotic Properties - Fault Diagnosis

Suppose the relationship between an OC image $\boldsymbol{X}$ and its faulty region is given by $g^* : \mathbb{R}^d \to \mathbb{R}^5$.

Theorem 2. Suppose that $g^*(x)$ is Lipschitz continuous. Assume the conditions in Theorem 1 with N replaced by $n_{\mathtt{OC}}$ If we further assume that there exists $a_1, a_2 > 0$ such that for any $|c| \leq a_1$,

$$\left|F_{\boldsymbol{z}|\boldsymbol{X}^{\mathtt{OC}}}\left(g^*(x)+c\right) - F_{\boldsymbol{z}|\boldsymbol{X}^{\mathtt{OC}}}\left(g^*(x)\right)\right| \geq a_2|c|, \text{ a.s. },$$

where $F_{\boldsymbol{z}|\boldsymbol{X}^{\mathtt{OC}}}(\cdot)$ is the cumulative distribution function of $\boldsymbol{z}$ given $\boldsymbol{X}^{\mathtt{OC}}$, then we have

$$\mathbb{E}\left[\left\|\mathscr{G}(\cdot;\widehat{\mathcal{U}}) - g^*\right\|^2_{L^2(\nu)}\right] = \mathcal{O}\left(n_{\mathtt{OC}}^{-1/(2+d)} \log^{3/2} n_{\mathtt{OC}}\right).$$

Managerial Interpretation

- The more training images (both IC and OC) the better.
- Upgrading the monitoring equipment (i.e., higher resolution) may not help.
- Should manufacturers never upgrade?
  - No need to upgrade if labeling fidelity is good enough.
  - Upgrade if low resolution causes mislabeling.

Benchmark Dataset - DAGM

- Ten classes of textured surfaces (512×512): 1000 IC and 150 OC images each class (1 – 6), 2000 IC and 300 OC images each class(7 – 10).
- Artificially created for an industrial image processing competition.

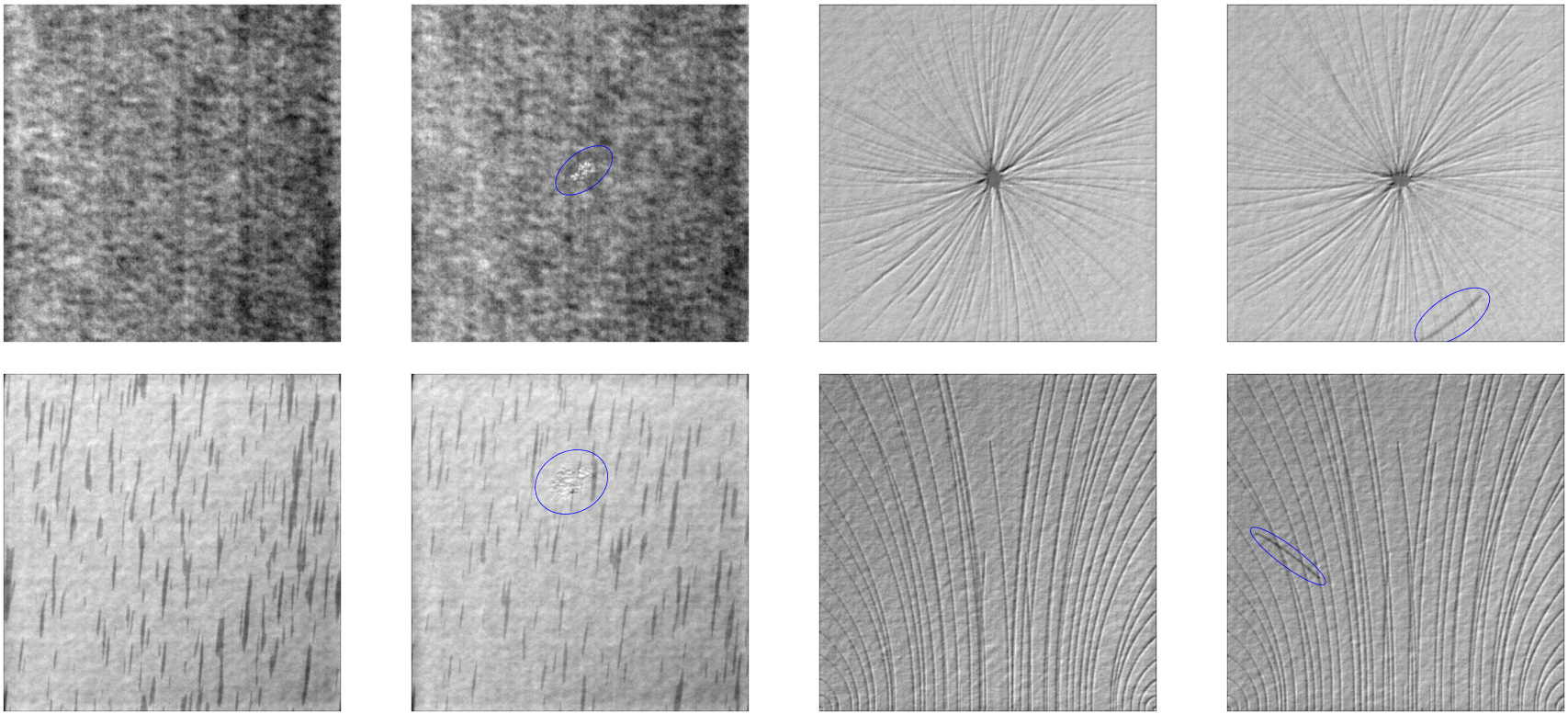


Figure 4: Four groups of IC/OC images from class 3, 4, 7 and 10 in DAGM are shown Left-to-right and top-to-bottom.

Training Details

- Data partition: 60% (training), 25% (validation), 15% (testing).
- Data Augmentation: Apply all 5 transformation to each OC image. Apply one of the 5 (randomly) to each IC image. About 2:1 IC/OC ratio.
- Batch size: 16. Learning rate follows a triangular cyclical schedule:

$$\alpha_j = \begin{cases} \alpha_L + (\alpha_U - \alpha_L)\frac{j-l\cdot C}{C}, & l\cdot C < j \le (l+1)C \\ \alpha_U - (\alpha_U - \alpha_L)\frac{j-(l+1)\cdot C}{C}, & (l+1)C < j \le (l+2)C \end{cases}, \; l = 0, 2, 4, \cdots,$$

  where $\alpha_L = 0.001$, $\alpha_U = 0.01$, and $C = 2000$.
- Transfer learning the diagnosis neural net $\mathscr{G}(\cdot;\mathcal{U})$ using $\widehat{\mathcal{V}}$ in the initialization.

Stopping Criterion

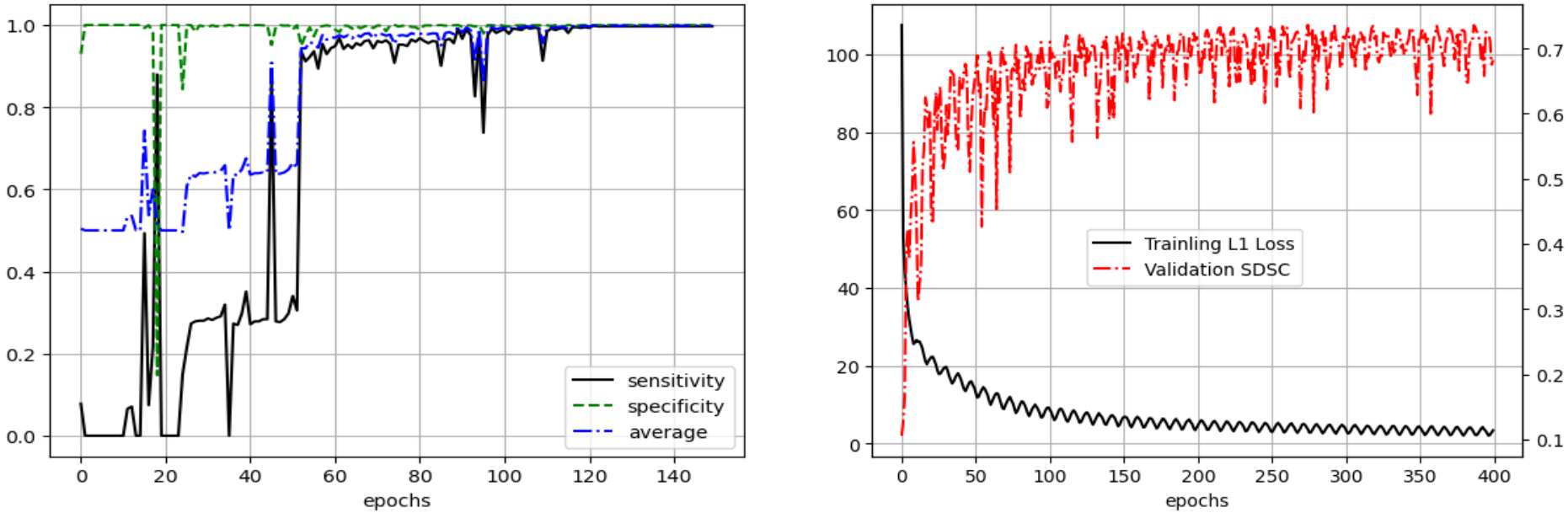


Figure 5: The training progression of $\mathscr{F}(\cdot;\mathcal{V})$ (left) and $\mathscr{G}(\cdot;\mathcal{U})$ (right).

Sorensen-Dice similarity coefficient (SDSC) used as the validation metric when training $\mathscr{G}(\cdot;\mathcal{U})$:

$$\text{SDSC} = \frac{2 \times \text{Area}(\text{EstimatedRegion} \cap \text{TrueRegion})}{\text{Area}(\text{EstimatedRegion}) + \text{Area}(\text{TrueRegion})}.$$

DAGM Testing Performance I

The change point is at $t = 21$.

Class	ARL_1	$sdARL_1$	Prop.Early	Prop.On	SDSC	sdSDSC
1	20.64	0.2533	0.02	0.98	0.8293	0.0280
2	20.96	0.0400	0.01	0.99	0.7157	0.0270
3	21.00	0.0000	0.00	1.00	0.7974	0.0131
4	20.82	0.1218	0.03	0.97	0.7696	0.0348
5	21.00	0.0000	0.00	1.00	0.8719	0.0135
6	21.00	0.0000	0.00	1.00	0.6548	0.0425
7	20.45	0.2618	0.06	0.94	0.8820	0.0133
8	21.01	0.0100	0.00	0.99	0.6036	0.0336
9	20.88	0.1104	0.02	0.98	0.8352	0.0101
10	20.88	0.0844	0.02	0.98	0.7072	0.0159

Table 1: Column sdARL and sdSDSC are standard errors based on 100 repeated simulations.

DAGM Testing Performance II

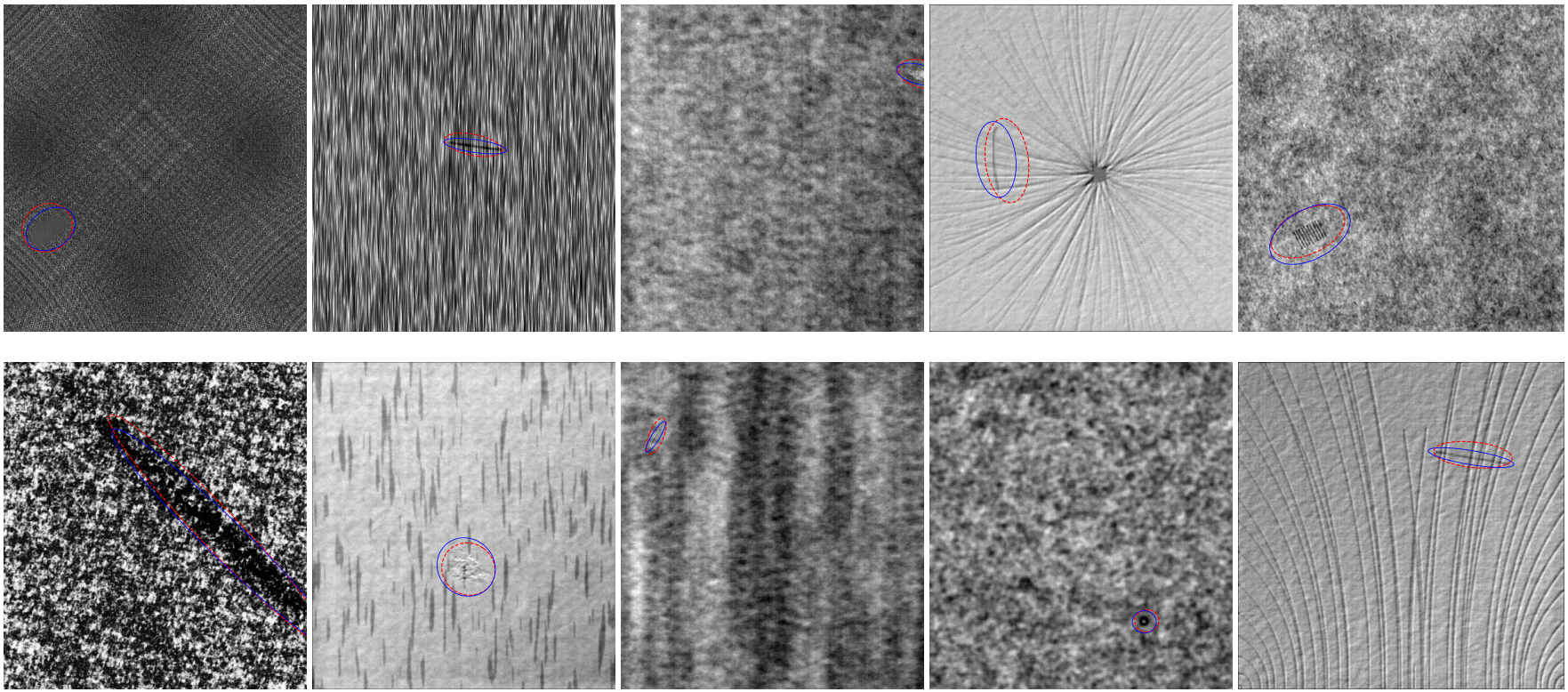


Figure 6: The blue and red ellipses are the true and estimated faulty regions, respectively. Each image corresponds to the median SDSC value in its class.

Compare with MF Method

Simulate images by the following auto-regressive (AR) model:

$$\boldsymbol{X}(i,j) = \phi_1 \boldsymbol{X}(i-1,j) + \phi_2 \boldsymbol{X}(i,j-1) + \varepsilon(i,j), \quad 1 \le i,j \le n,$$

- When the production process is IC, we set $\phi_1 = 0.65$, $\phi_2 = 0.35$ and $\sigma = 0.1$.
- Type-1 defect: $\phi_1 = 0.65$, $\phi_2 = 0.35$ and $\sigma = 10^{-6}$.
- Type-2 defect: $\phi_1 = \phi_2 = 0$ and $\sigma = 0.1$.
- The change point occurs at $t = 21$.
- The ARL_0 is set equal to 200.

Applicability to AR Images

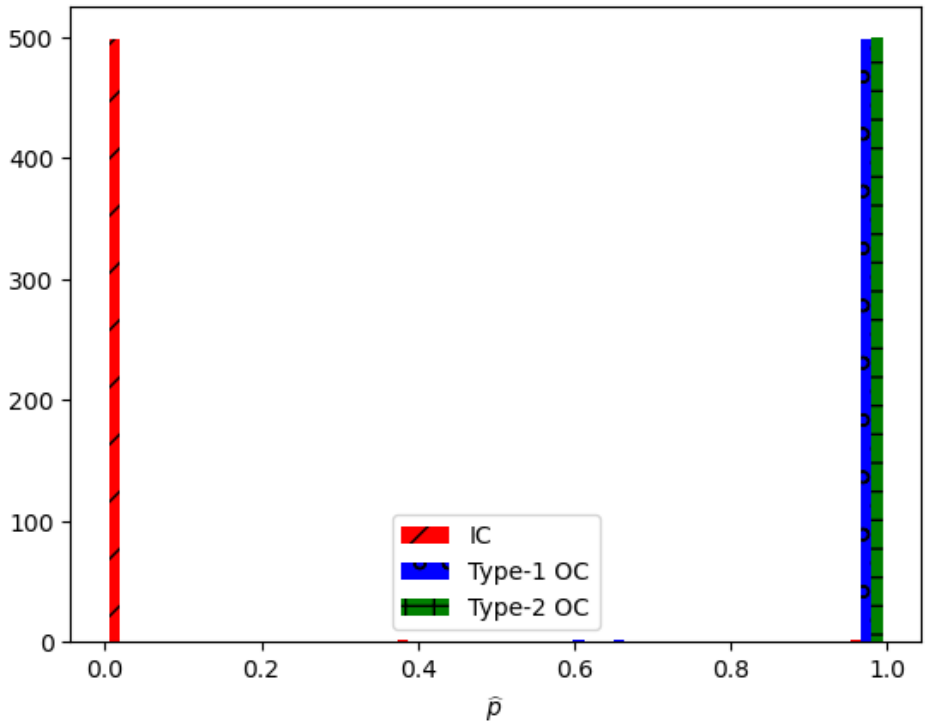


Figure 7: The histogram of the applicability test in which $\mathscr{F}(\cdot;\widehat{\mathcal{V}})$ is applied to 500 IC images and the two types of OC images (500 each type) generated by the AR model.

Comparison Result

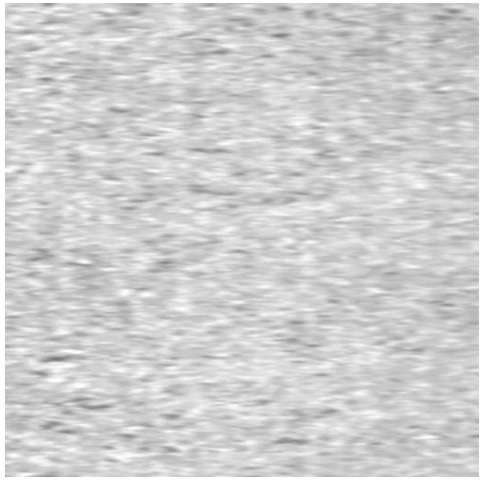
Table 2: Comparison with the MF method using AR images.

Fault Type	Method	ARL_1	$sdARL_1$	Prop.On	SDSC	sdSDSC
1	MODERN	21.00	0.0000	1.00	0.6286	0.0084
	MF	153.75	6.7987	0.01	0.0000	0.0000
2	MODERN	21.00	0.0000	1.00	0.6208	0.0057
	MF	150.58	6.3956	0.00	0.5762	0.0097

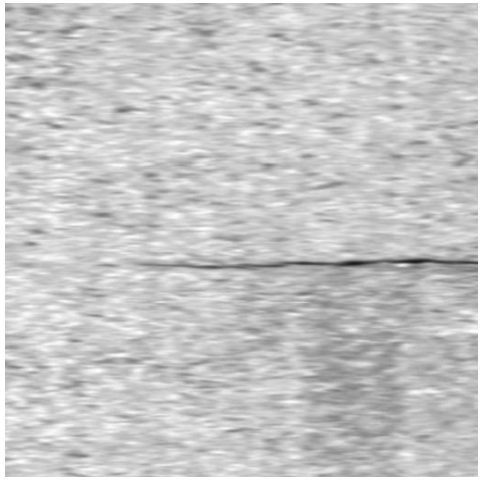
Application to Manufacturing of Electric Commutators

Electric commutators are key components of electric motors. Our sample includes 14 IC images and 1 OC image of commutators.

- 10 IC images are used for estimating the IC parameters and control limit.
- The monitoring phase shows the three IC images first for $t \leq 3$ and then the OC image for $t > 3$.
- ARL_0 is set equal to 25.



One of the 14 IC images



The OC image

Figure 8: Representative images in the electric commutator manufacturing.

Commutator Manufacturing: Monitoring and Diagnosis

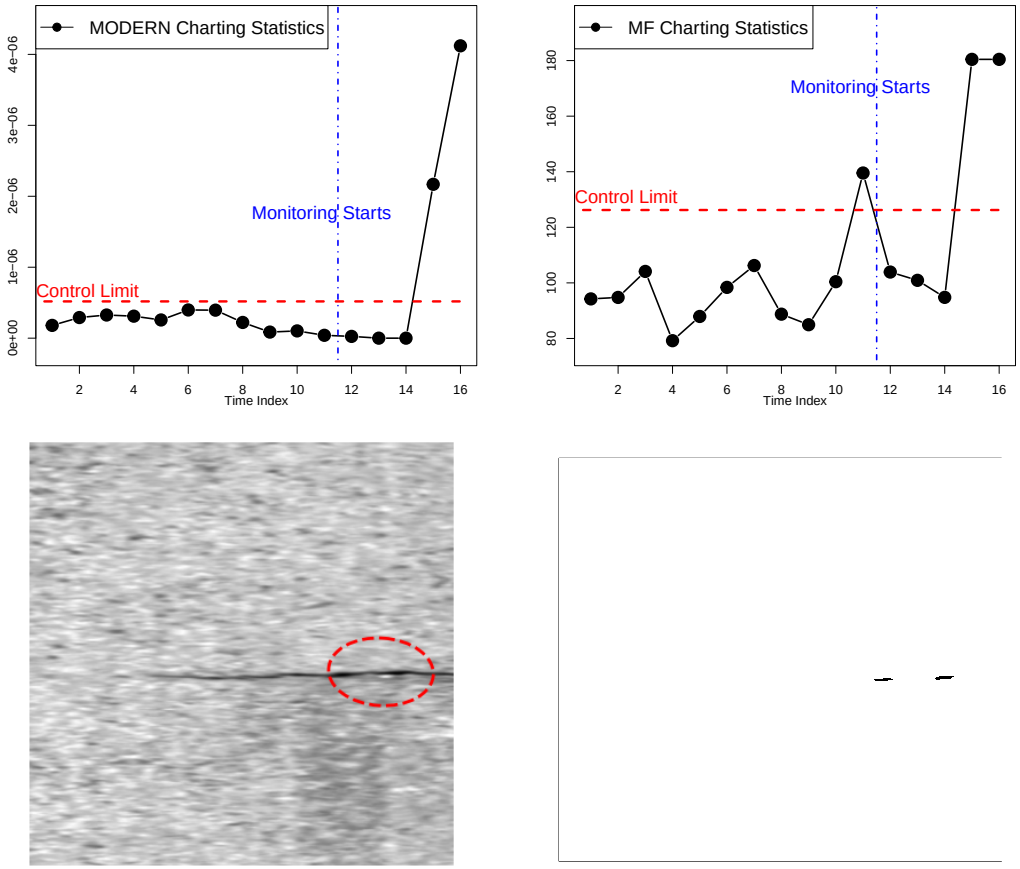


Figure 9: Comparison with the MF method in the electric commutator manufacturing data.

Summary

- Proposed a CNN-based framework, MODERN, for image monitoring and diagnosis (no need of a gold standard or the MF assumptions).
- Suggested transfer monitoring for assessing applicability and reducing sample size requirement.
- Established minimax optimality and its managerial interpretation.


Thank you!

References

Bui, A. T. and Apley, D. W. (2018). A monitoring and diagnostic approach for stochastic textured surfaces. *Technometrics*, 60(1):1–13.

Feng, L. and Qiu, P. (2018). Difference detection between two images for image monitoring. *Technometrics*, 60(3):345–359.

Jiang, B., Wang, C.-C., and Liu, H.-C. (2005). Liquid crystal display surface uniformity defect inspection using analysis of variance and exponentially weighted moving average techniques. *International Journal of Production Research*, 43(1):67–80.

Koosha, M., Noorossana, R., and Megahed, F. (2017). Statistical process monitoring via image data using wavelets. *Quality and Reliability Engineering International*, 33(8):2059–2073.

Lin, H. (2007). Automated visual inspection of ripple defects using wavelet characteristic based multivariate statistical approach. *Image and Vision Computing*, 25(11):1785–1801.

López, F., Valiente, J. M., Prats, J. M., and Ferrer, A. (2008). Performance evaluation of soft color texture descriptors for surface grading using experimental design and logistic regression. *Pattern Recognition*, 41(5):1744–1755.

Lyu, J. and Chen, M. (2009). Automated visual inspection expert system for multivariate statistical process control chart. *Expert Systems with Applications*, 36(3):5113–5118.

Megahed, F. M., Wells, L. J., Camelio, J. A., and Woodall, W. H. (2012). A spatiotemporal method for the monitoring of image data. *Quality and Reliability Engineering International*, 28(8):967–980.

Szegedy, C., Ioffe, S., Vanhoucke, V., and Alemi, A. (2017). Inception-v4, inception-resnet and the impact of residual connections on learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 31, pages 4278–4284.