

Boost-Up Efficiency of Defective Solar Panel Detection With Pre-Trained Attention Recycling

YeongHyeon Park ^{1,2}, Myung Jin Kim ², Uju Gim ², Juneho Yi ¹

Department of Electrical and Computer Engineering, Sungkyunkwan University ¹

IoT Solution Business Group, SK Planet Co., Ltd. ²



IEEE TRANSACTIONS ON
**INDUSTRY
APPLICATIONS**



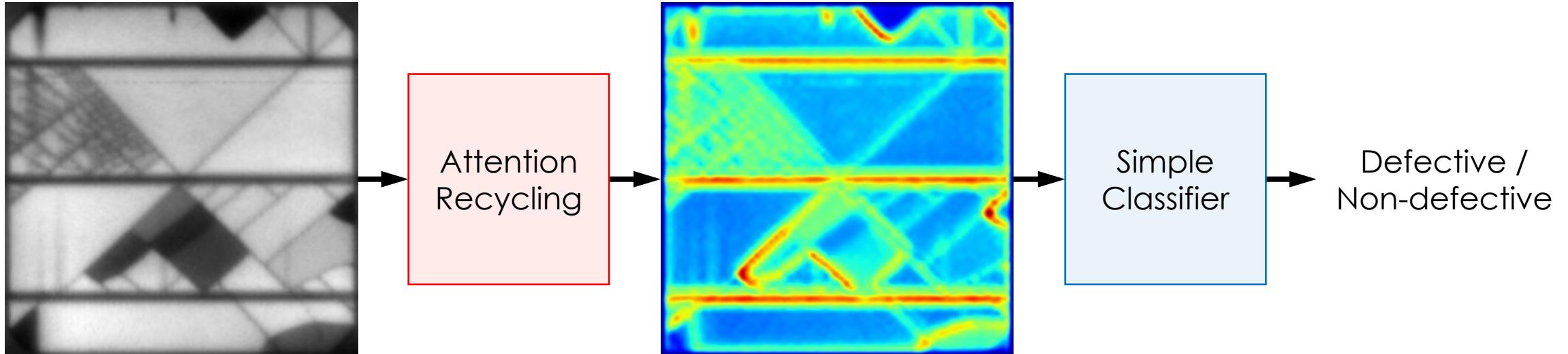
Motivation

- ❑ Solar panel defects significantly degrade energy conversion efficiency [1,2]
- ❑ It is necessary to develop a practically deployable method
 - To solve real-world problems
 - To avoid blindly employing end-to-end deep learning methods



[1] A. Taşçıoğlu et al., "A power case study for monocrystalline and polycrystalline solar panels in Bursa city, Turkey," *International Journal of Photoenergy*, 2016
[2] S. Deitsch et al., "Automatic classification of defective photovoltaic module cells in electroluminescence images," *Solar Energy*, 2019

Key idea

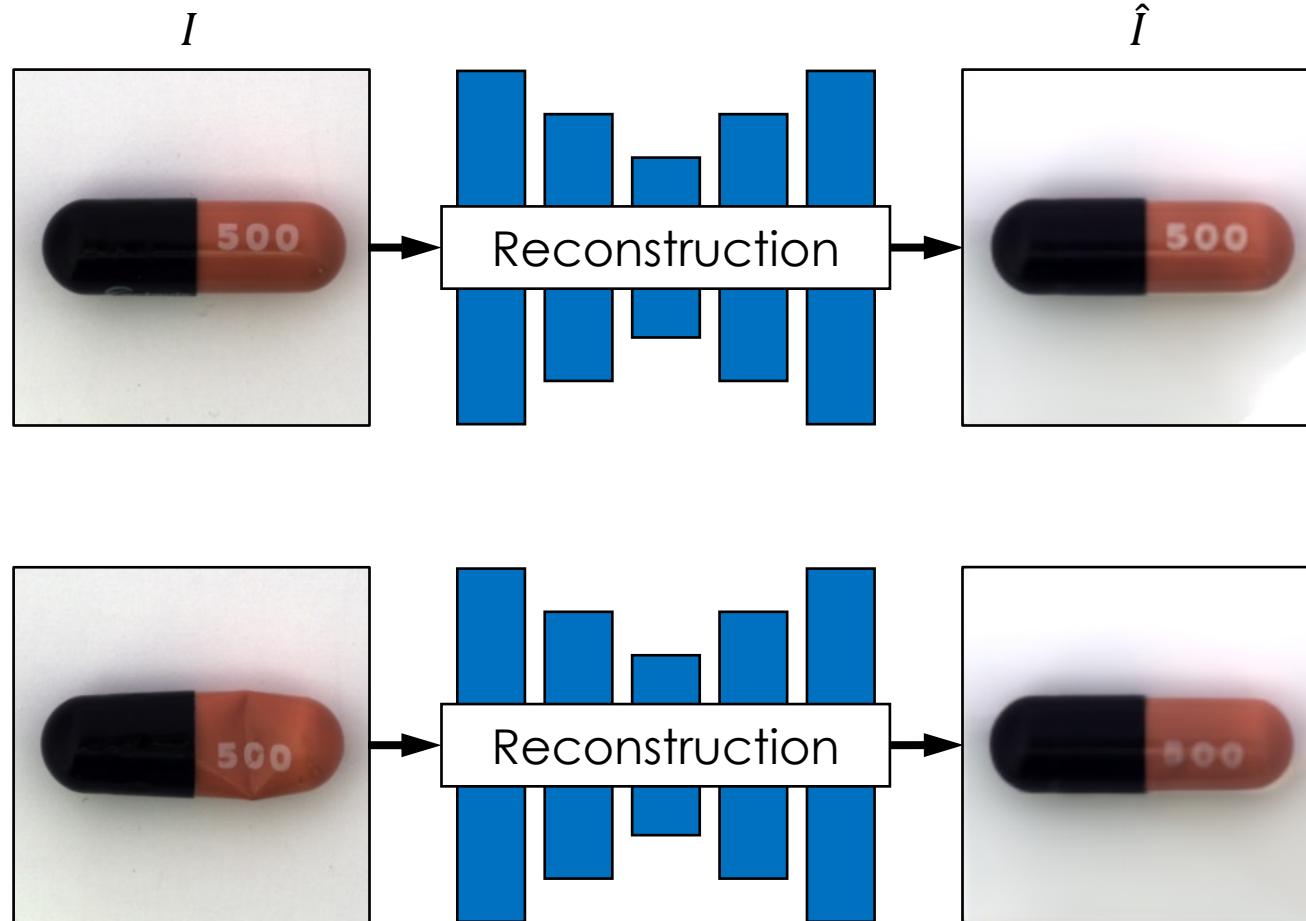


Recycling of a pre-trained attention mechanism

Boost-Up Efficiency of Defective Solar Panel Detection With Pre-Trained Attention Recycling

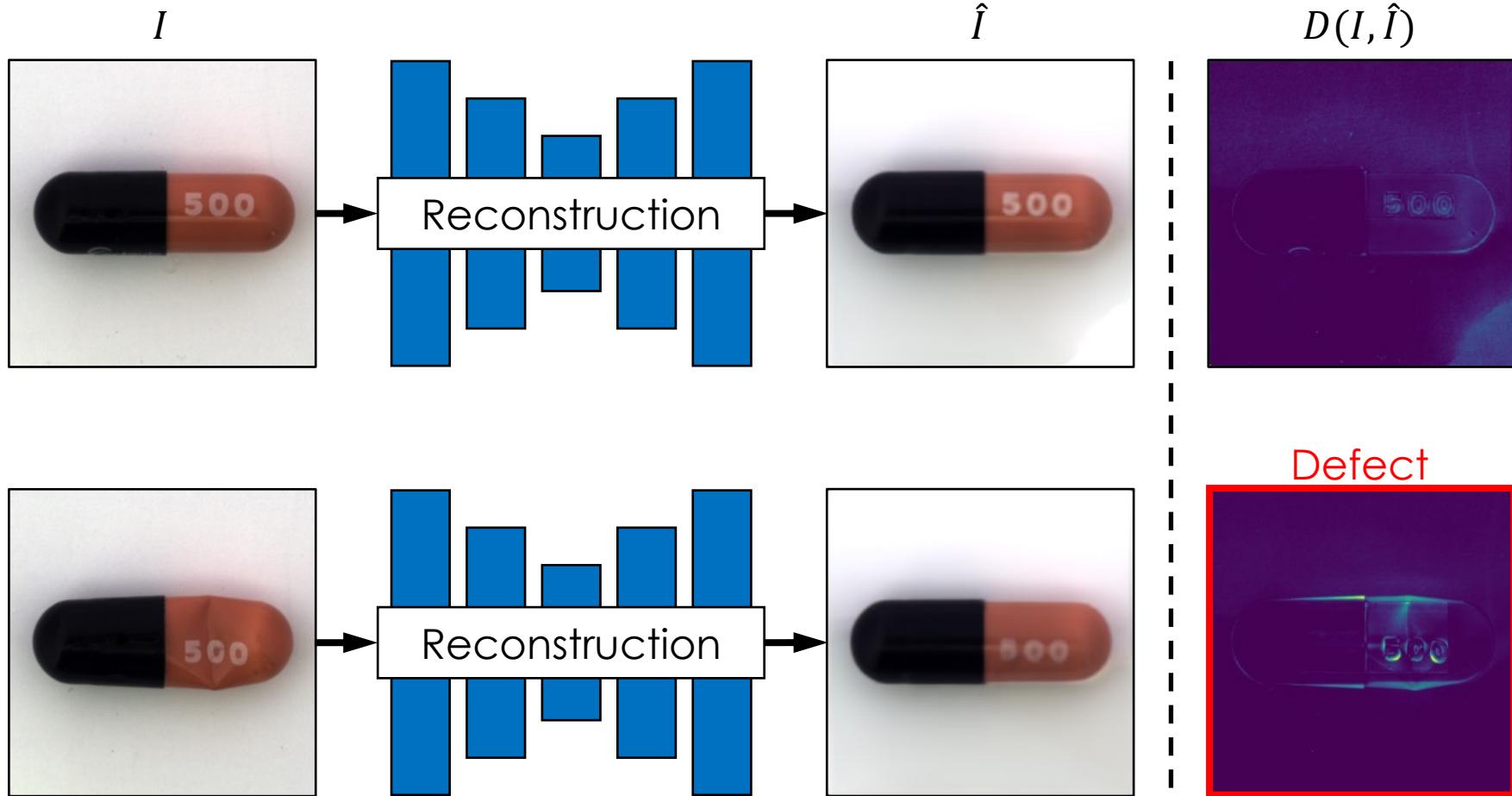
[Details](#)

Anomaly Detection



[3] Y. Park et al., "Neural Network Training Strategy to Enhance Anomaly Detection Performance: A Perspective on Reconstruction Loss Amplification," arXiv, 2023

Anomaly Detection



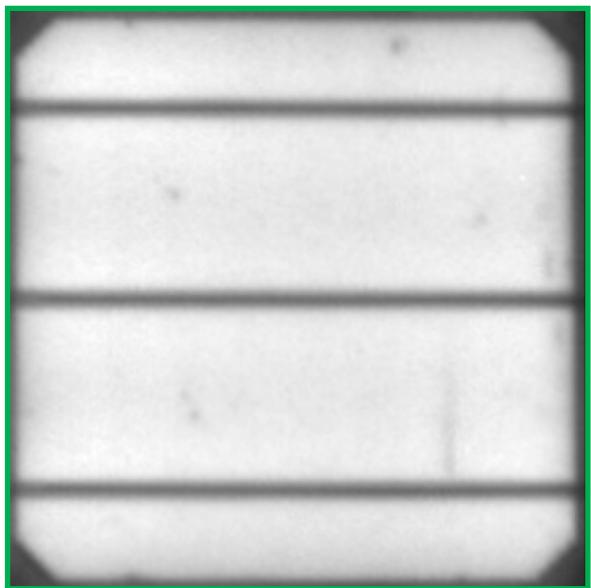
[3] Y. Park et al., "Neural Network Training Strategy to Enhance Anomaly Detection Performance: A Perspective on Reconstruction Loss Amplification," arXiv, 2023

Anomaly Detection

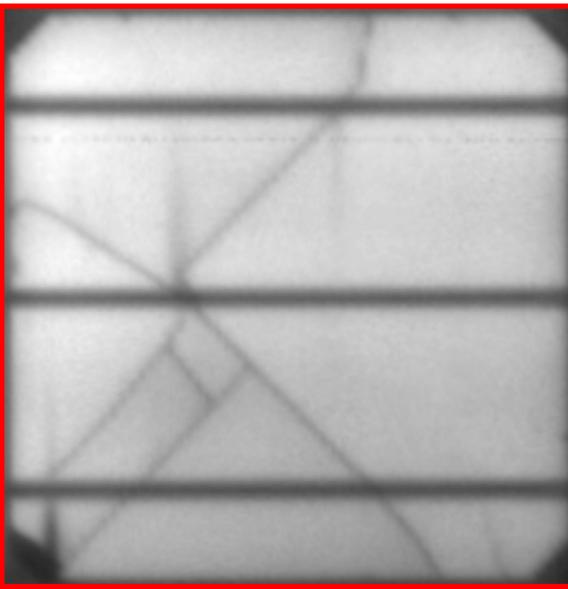
- Unsupervised anomaly detection based on reconstruction error
 - + Eases class imbalance problem
 - + Training with non-defective samples only
 - Deep neural networks consume lots of power
 - Requires a large-scale dataset

- Defective solar panel detection
 - + Class balanced dataset
 - + Easy to understand without deep knowledge
 - Small quantity of samples

Solar panels [4]



Non-defective



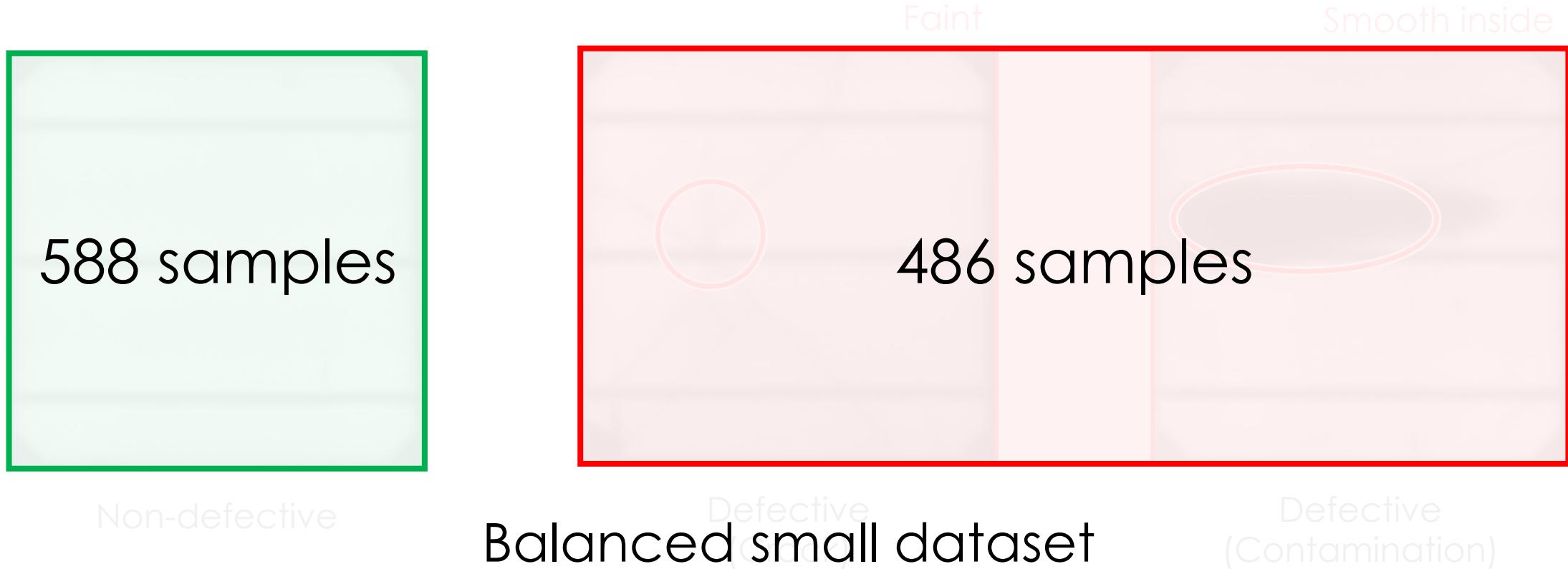
Defective
(Crack)



Defective
(Contamination)

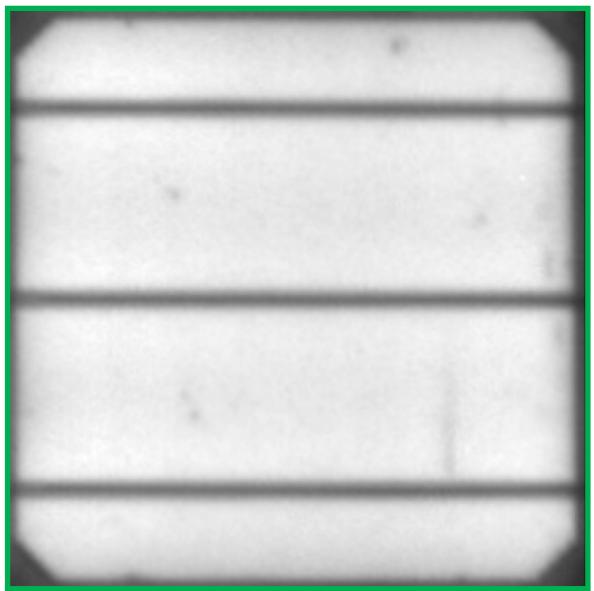
[4] C. Buerhop-Lutz et al., "A benchmark for visual identification of defective solar cells in electroluminescence imagery," EU PVSEC, 2018

Solar panels [4]

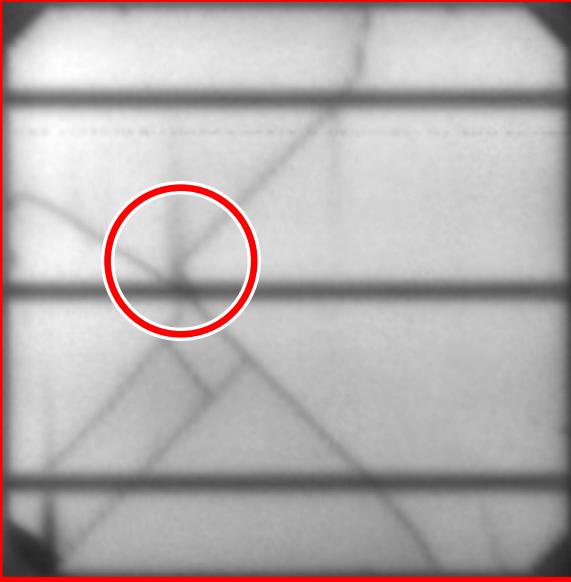


[4] C. Buerhop-Lutz et al., "A benchmark for visual identification of defective solar cells in electroluminescence imagery," EU PVSEC, 2018

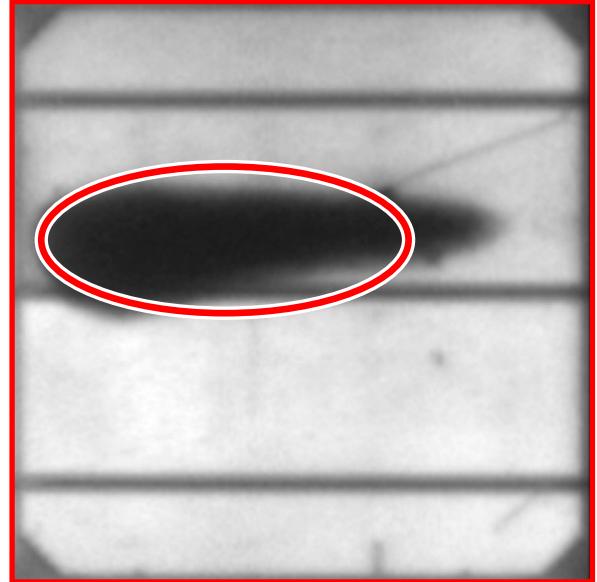
Solar panels [4]



Non-defective



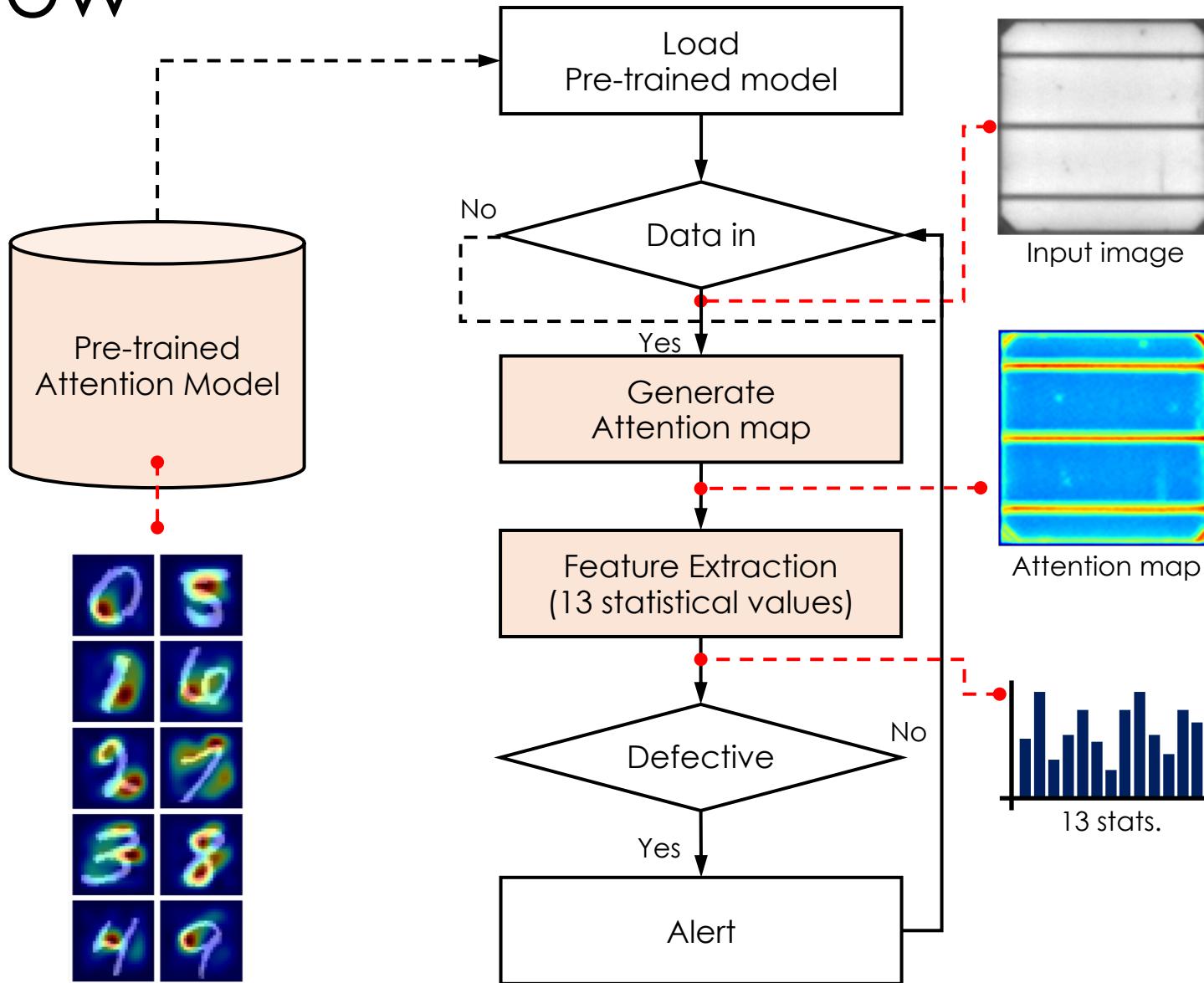
Defective
(Crack)



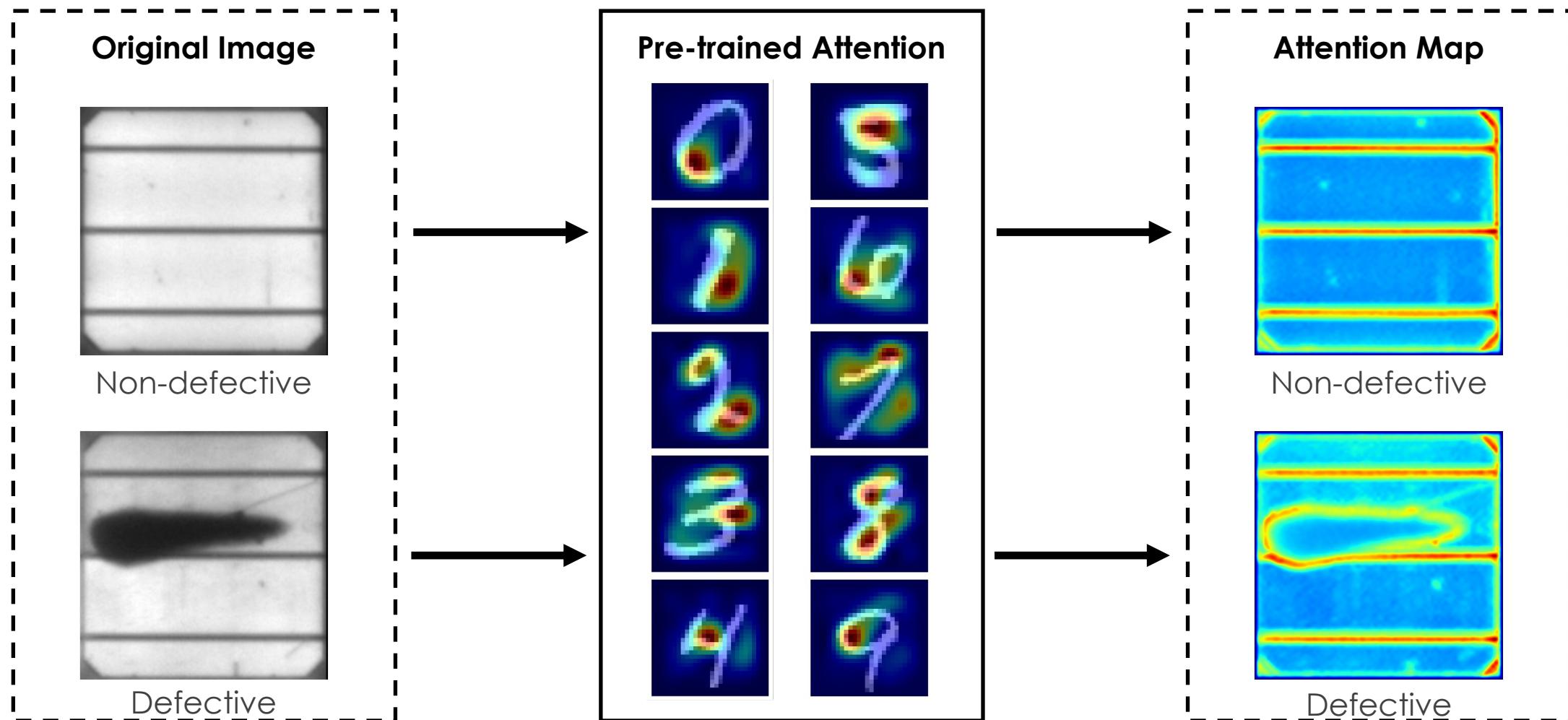
Defective
(Contamination)

[4] C. Buerhop-Lutz et al., "A benchmark for visual identification of defective solar cells in electroluminescence imagery," EU PVSEC, 2018

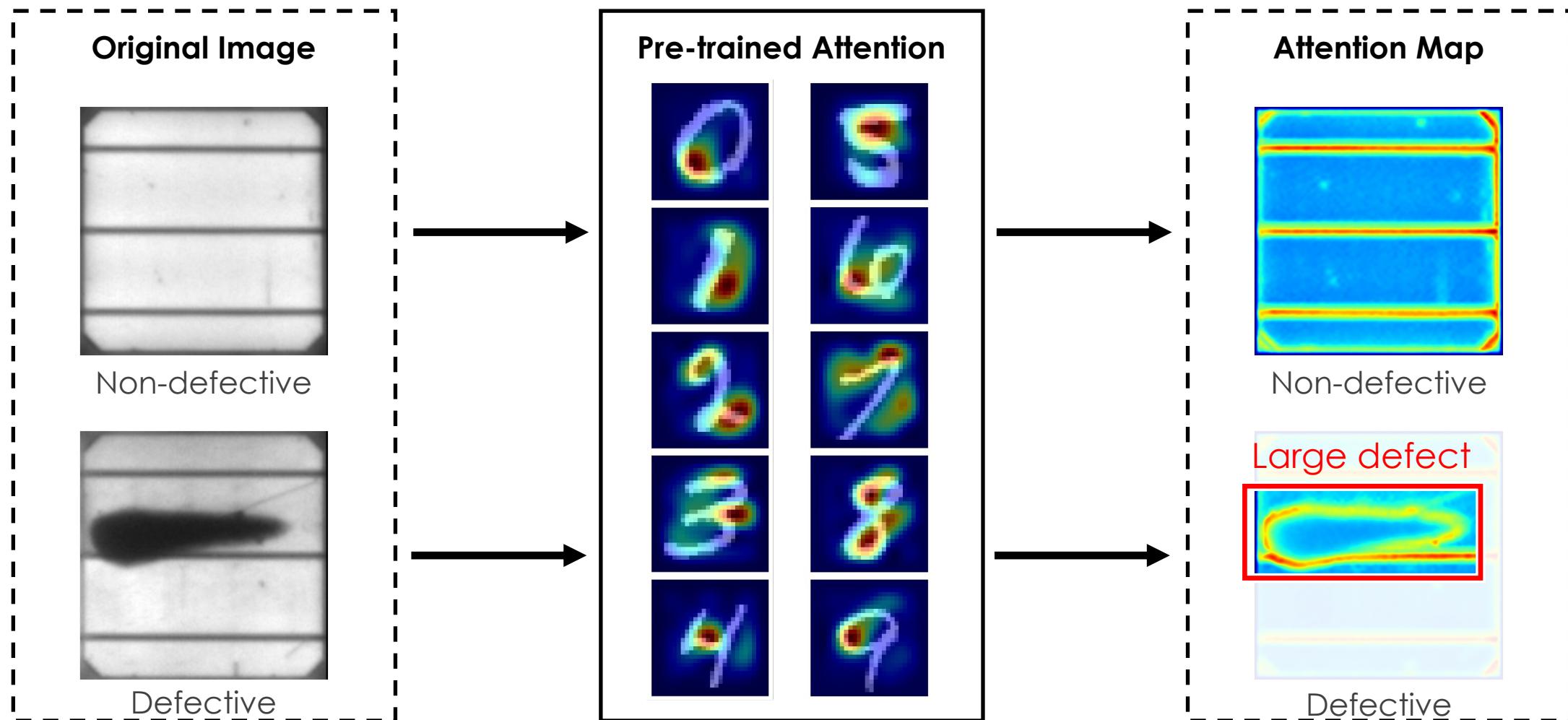
Workflow



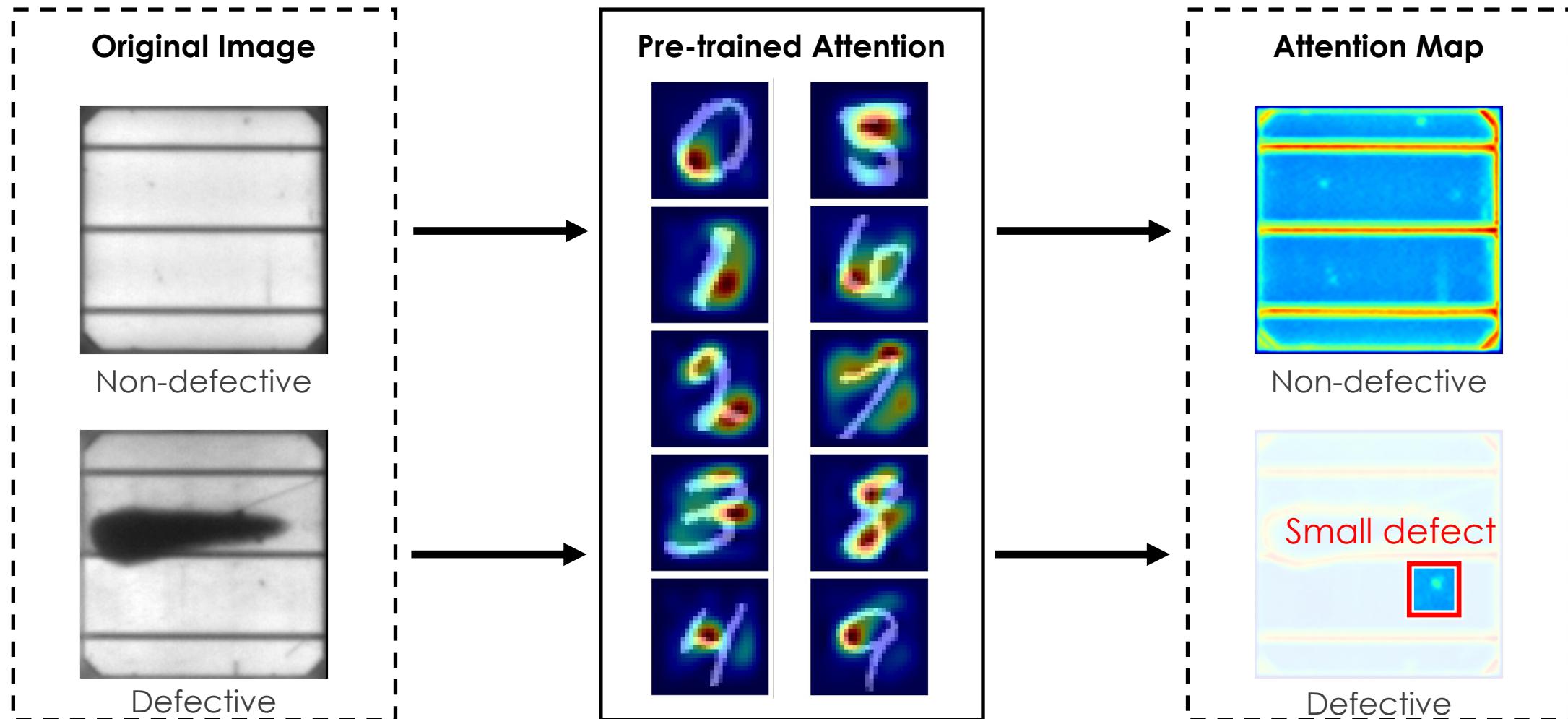
Emphasizing the defect



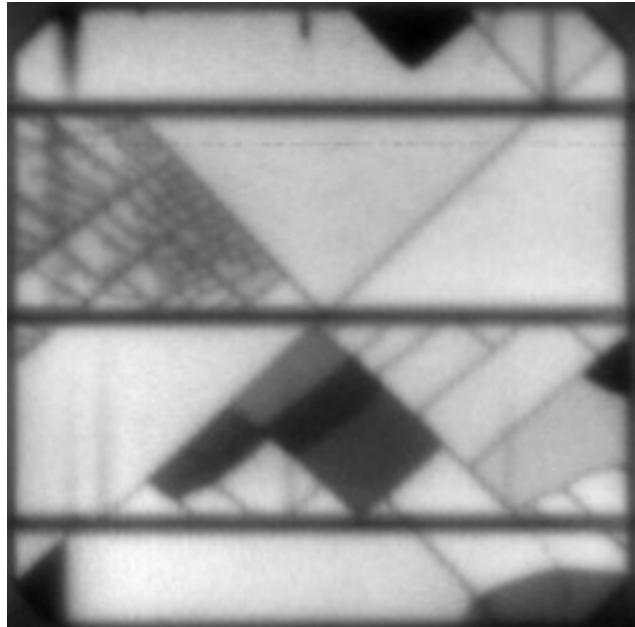
Emphasizing the defect



Emphasizing the defect

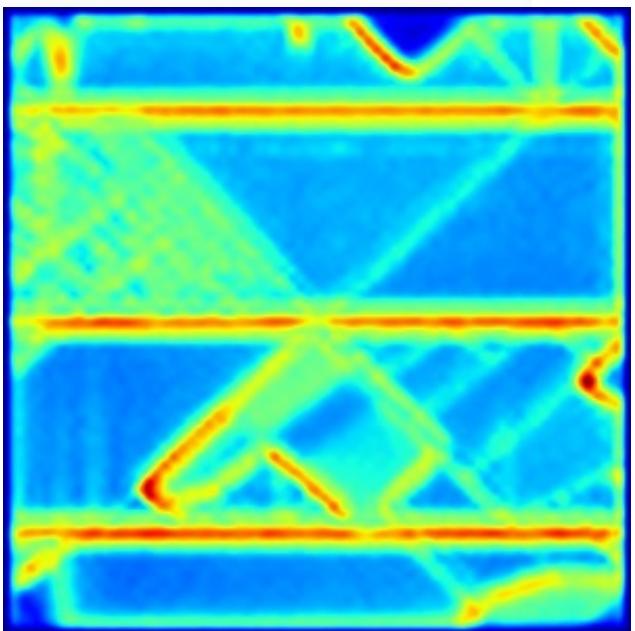


Statistical feature extraction



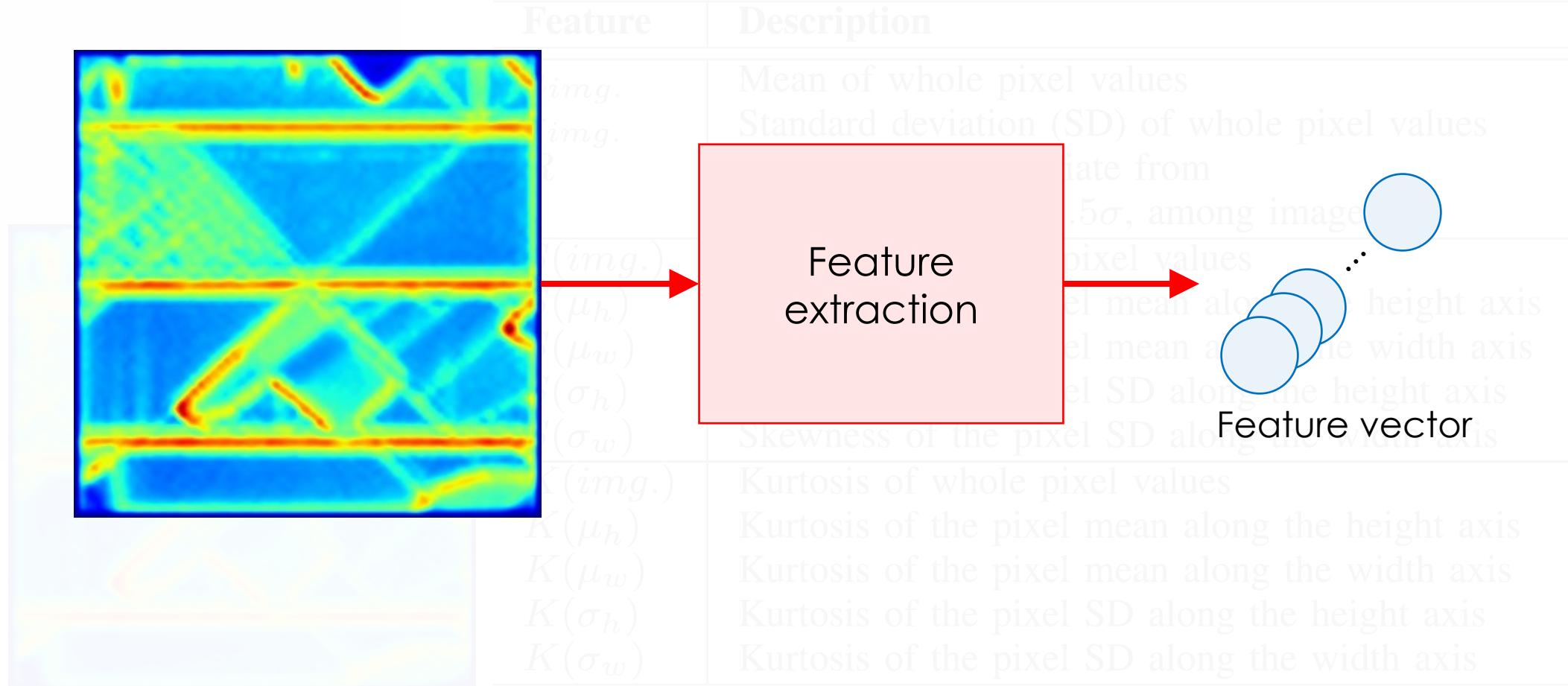
Feature	Description
$\mu_{img.}$	Mean of whole pixel values
$\sigma_{img.}$	Standard deviation (SD) of whole pixel values
R	Outlier rate that deviate from the threshold, $\mu \pm 1.5\sigma$, among image
$S(img.)$	Skewness of whole pixel values
$S(\mu_h)$	Skewness of the pixel mean along the height axis
$S(\mu_w)$	Skewness of the pixel mean along the width axis
$S(\sigma_h)$	Skewness of the pixel SD along the height axis
$S(\sigma_w)$	Skewness of the pixel SD along the width axis
$K(img.)$	Kurtosis of whole pixel values
$K(\mu_h)$	Kurtosis of the pixel mean along the height axis
$K(\mu_w)$	Kurtosis of the pixel mean along the width axis
$K(\sigma_h)$	Kurtosis of the pixel SD along the height axis
$K(\sigma_w)$	Kurtosis of the pixel SD along the width axis

Statistical feature extraction

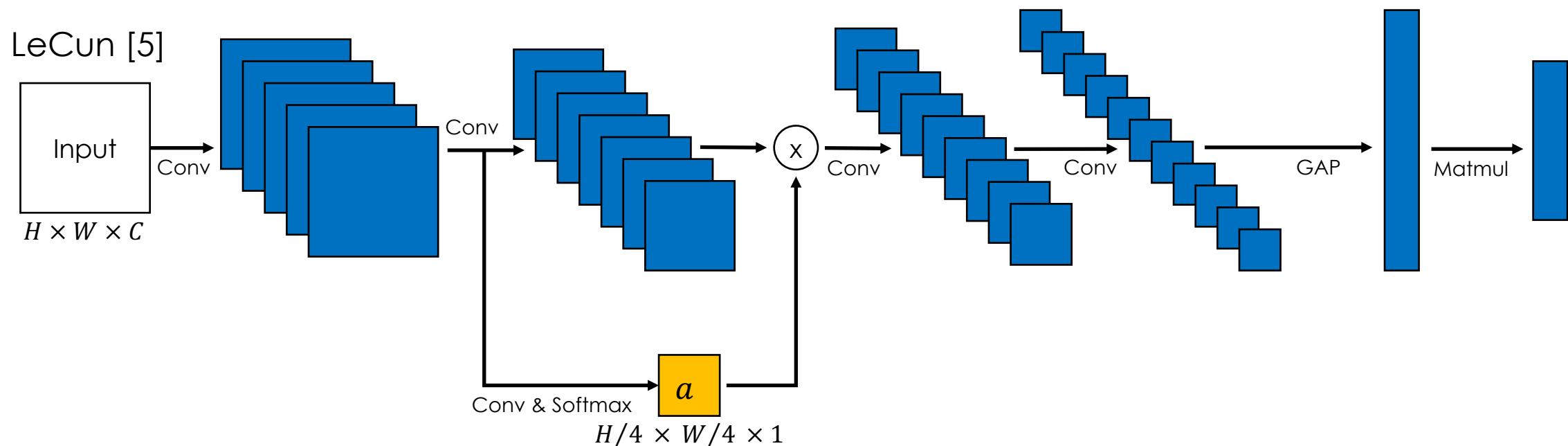


Feature	Description
$\mu_{img.}$	Mean of whole pixel values
$\sigma_{img.}$	Standard deviation (SD) of whole pixel values
R	Outlier rate that deviate from the threshold, $\mu \pm 1.5\sigma$, among image
$S(img.)$	Skewness of whole pixel values
$S(\mu_h)$	Skewness of the pixel mean along the height axis
$S(\mu_w)$	Skewness of the pixel mean along the width axis
$S(\sigma_h)$	Skewness of the pixel SD along the height axis
$S(\sigma_w)$	Skewness of the pixel SD along the width axis
$K(img.)$	Kurtosis of whole pixel values
$K(\mu_h)$	Kurtosis of the pixel mean along the height axis
$K(\mu_w)$	Kurtosis of the pixel mean along the width axis
$K(\sigma_h)$	Kurtosis of the pixel SD along the height axis
$K(\sigma_w)$	Kurtosis of the pixel SD along the width axis

Statistical feature extraction

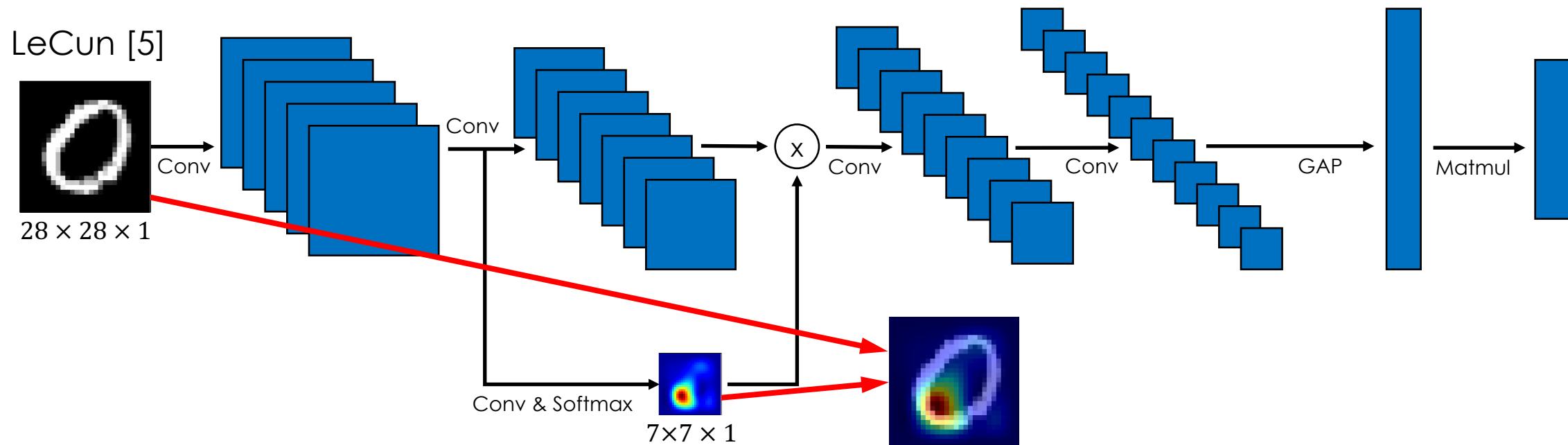


Pre-trained attention



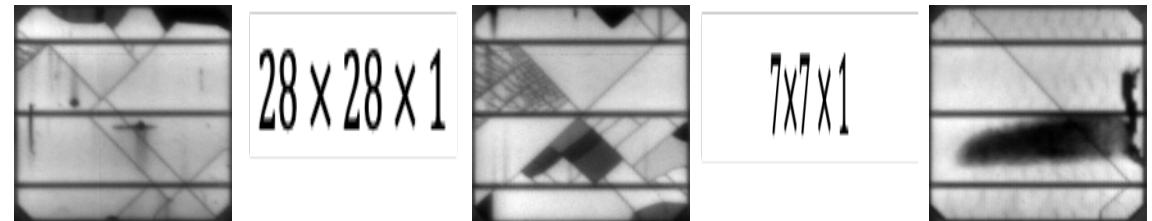
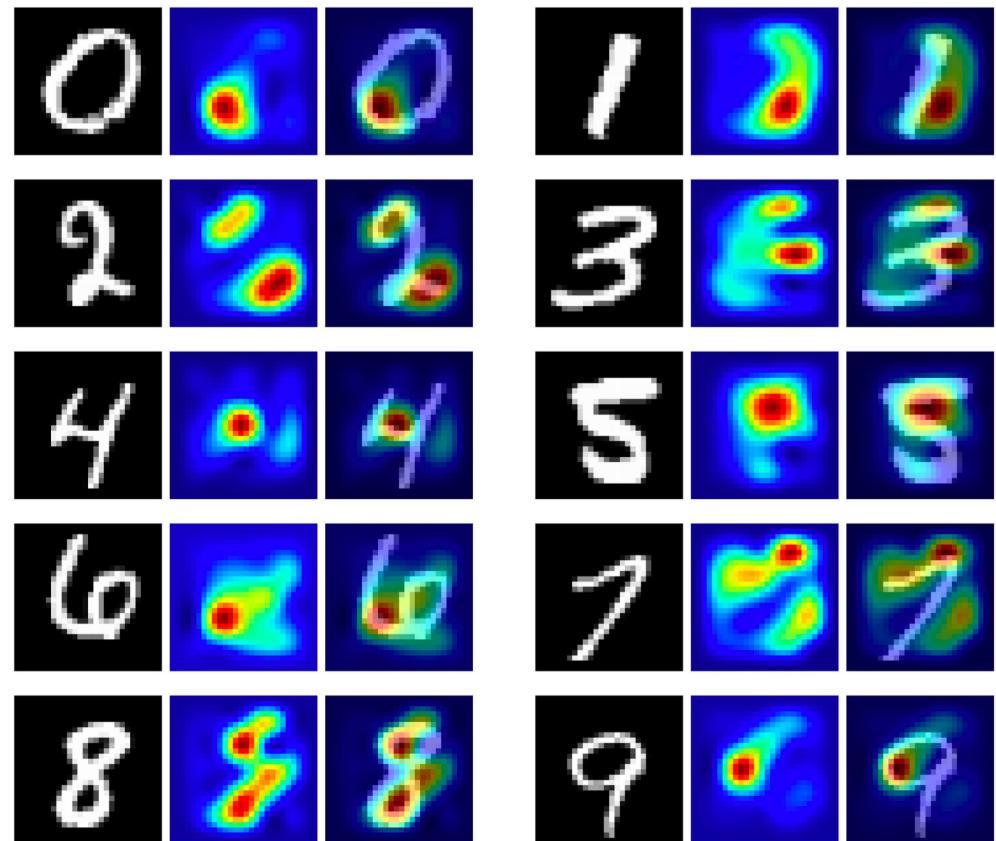
[5] Y. LeCun, et al. "Gradient-based learning applied to document recognition." Proceedings of the IEEE, 1998

Pre-trained attention



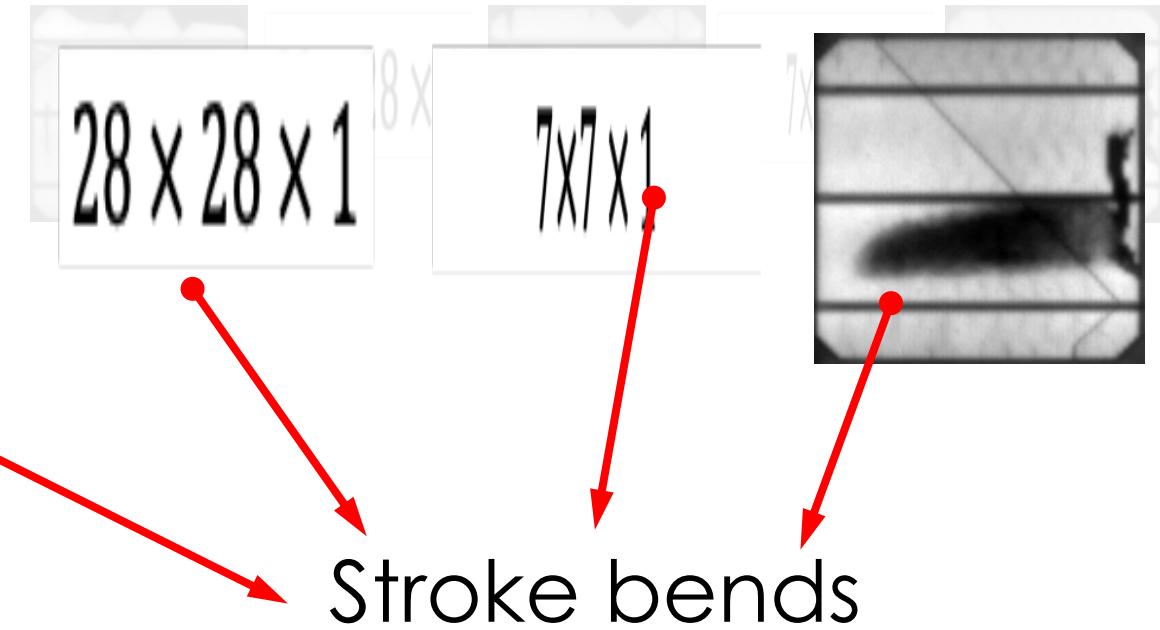
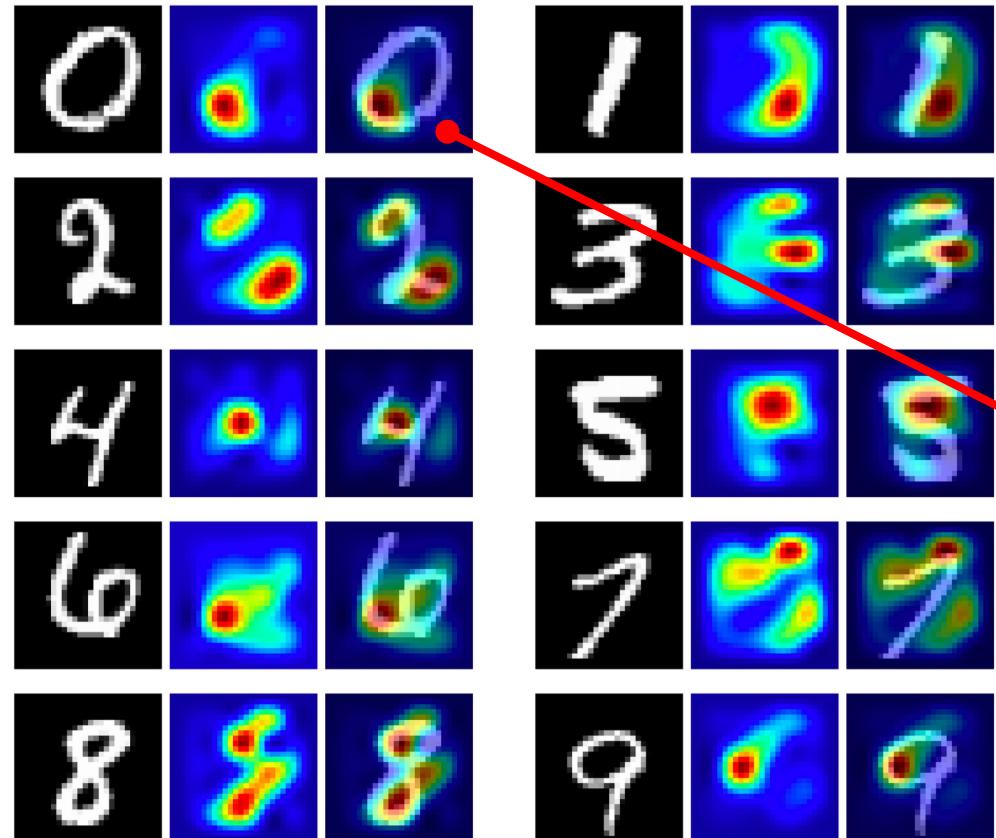
[5] Y. LeCun, et al. "Gradient-based learning applied to document recognition." Proceedings of the IEEE, 1998

Pre-trained attention



[5] Y. LeCun, et al. "Gradient-based learning applied to document recognition." Proceedings of the IEEE, 1998

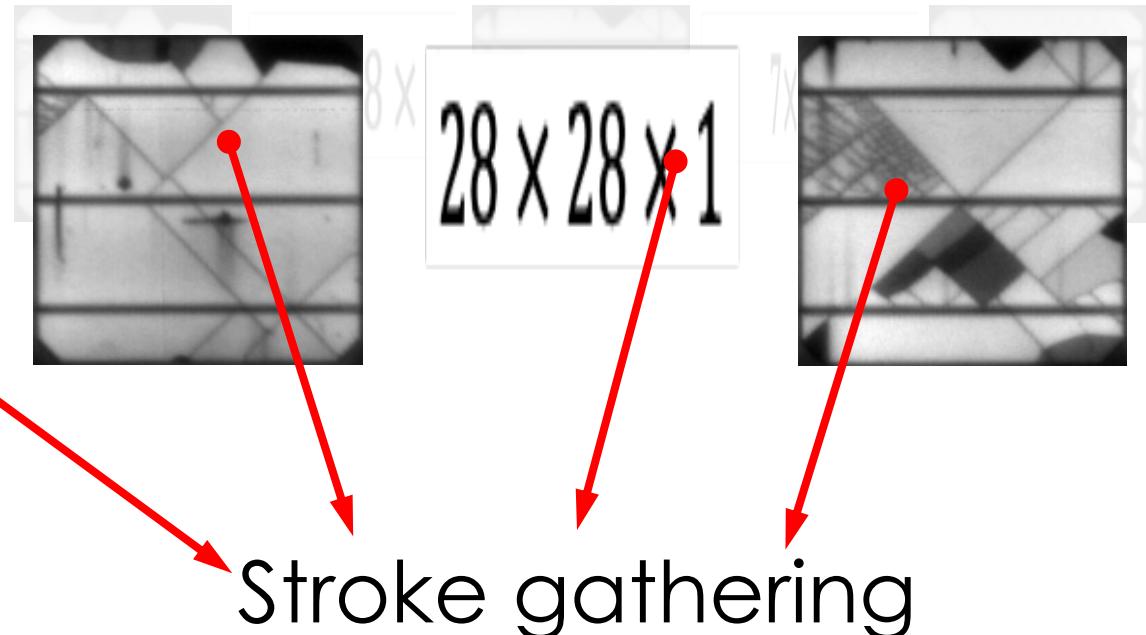
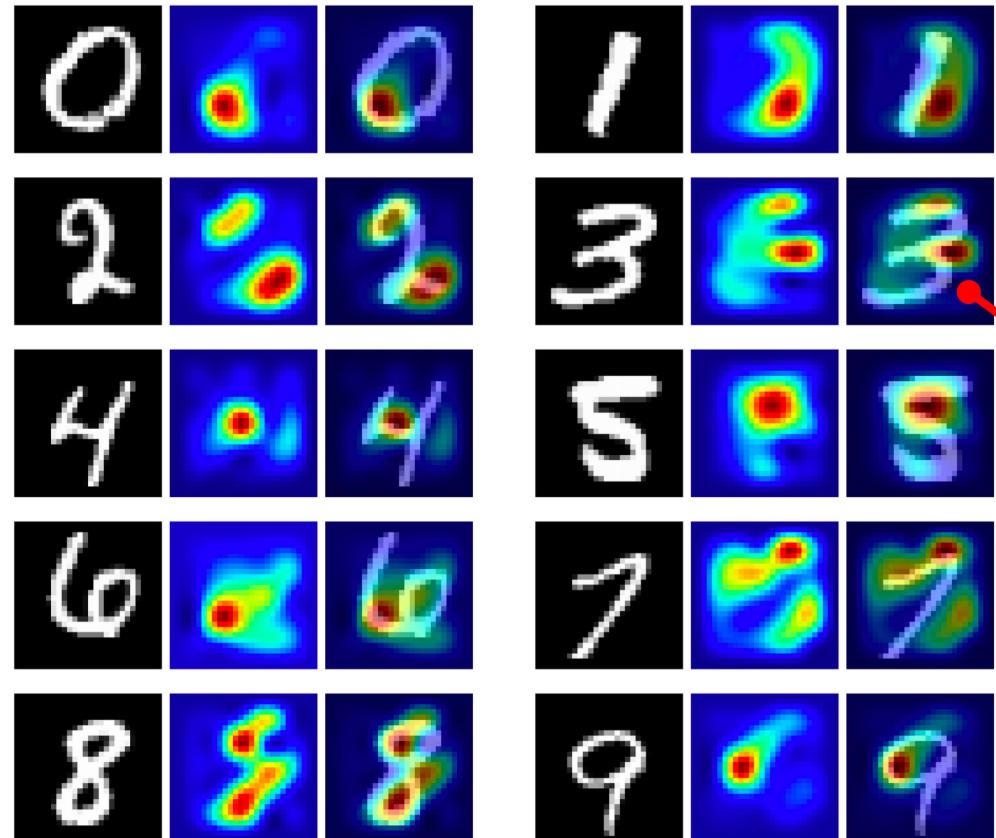
Pre-trained attention



Stroke bends

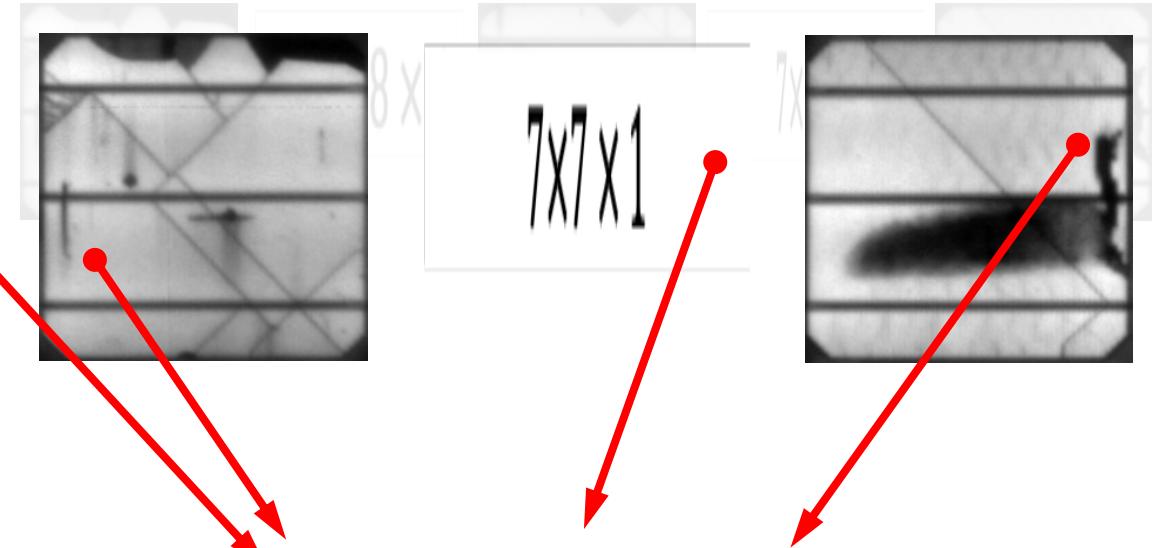
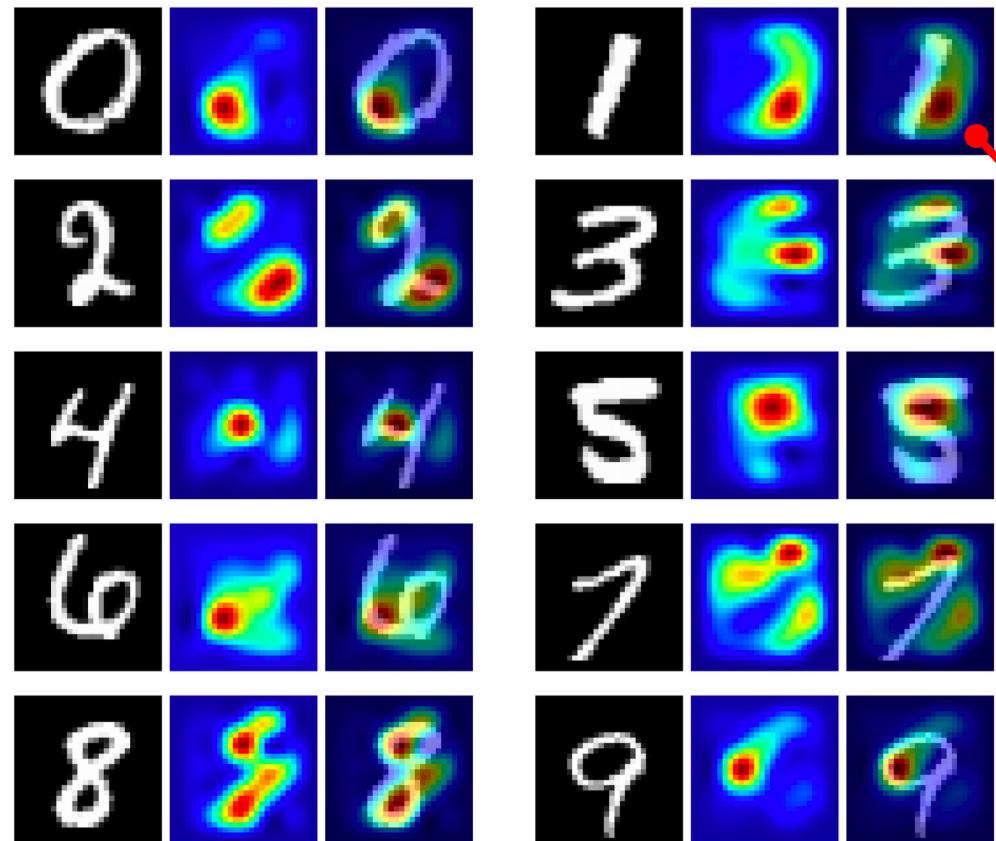
[5] Y. LeCun, et al. "Gradient-based learning applied to document recognition." Proceedings of the IEEE, 1998

Pre-trained attention



[5] Y. LeCun, et al. "Gradient-based learning applied to document recognition." Proceedings of the IEEE, 1998

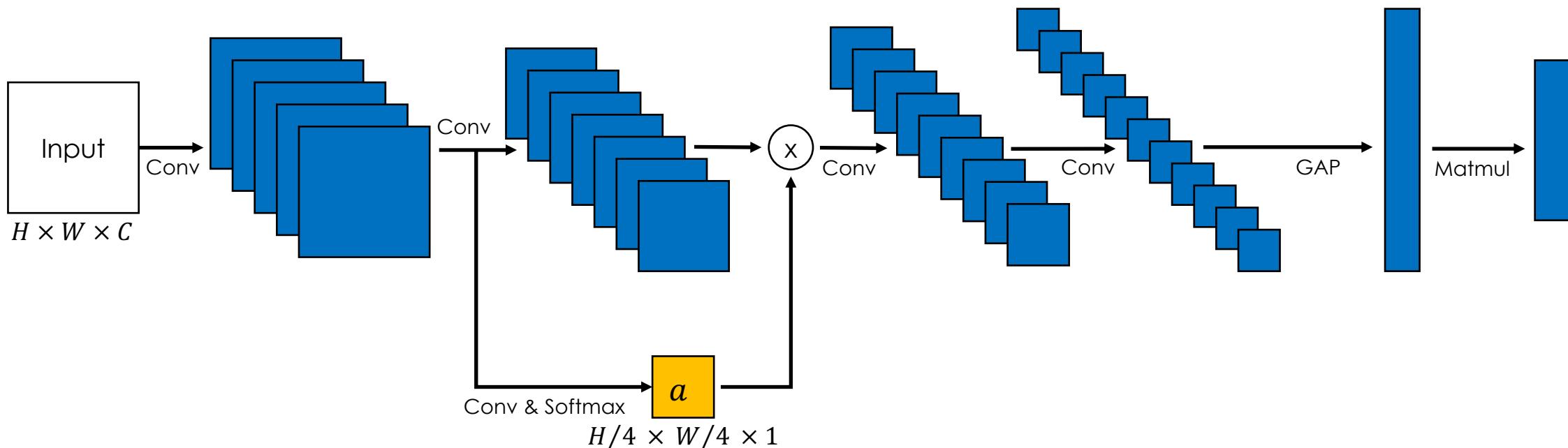
Pre-trained attention



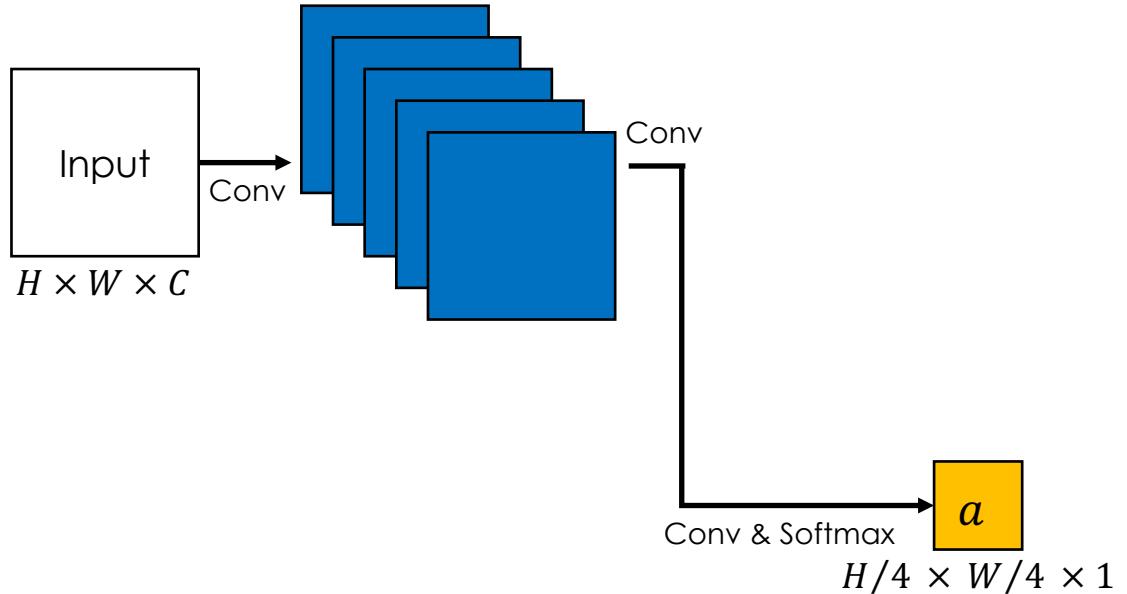
Stroke ends

[5] Y. LeCun, et al. "Gradient-based learning applied to document recognition." Proceedings of the IEEE, 1998

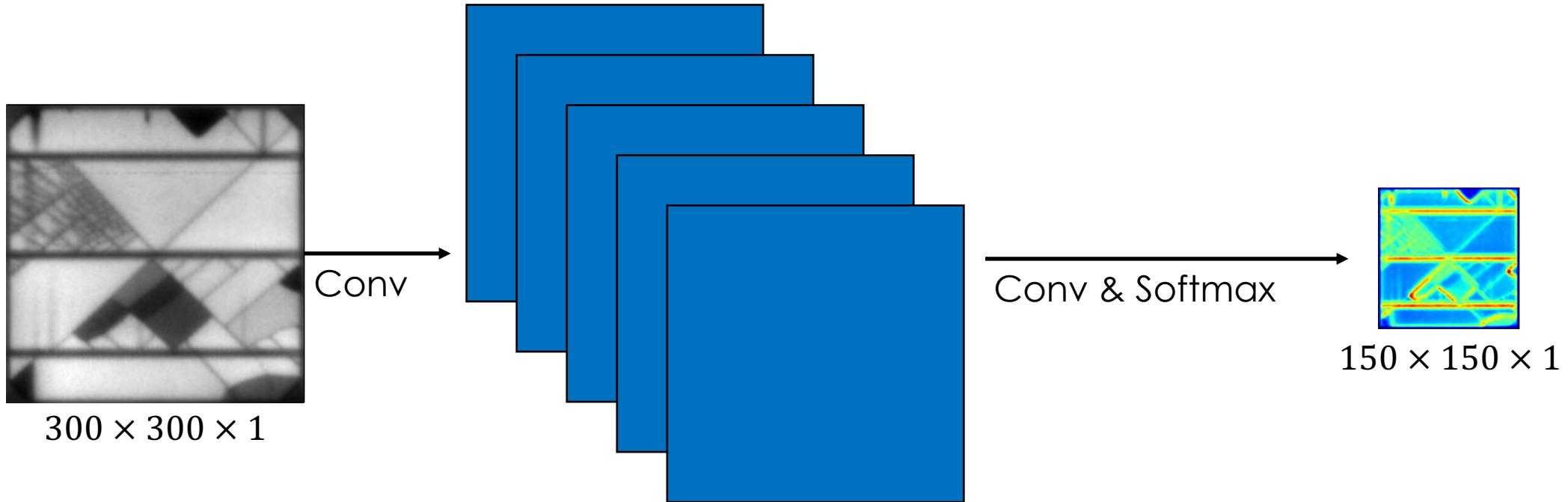
Pre-trained attention



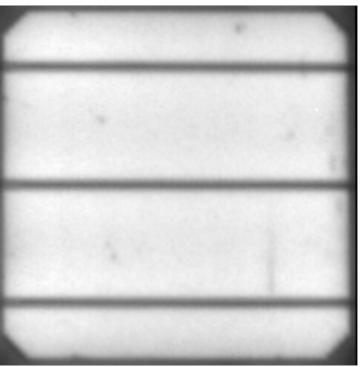
Pre-trained attention



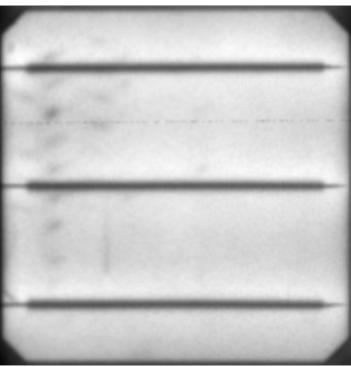
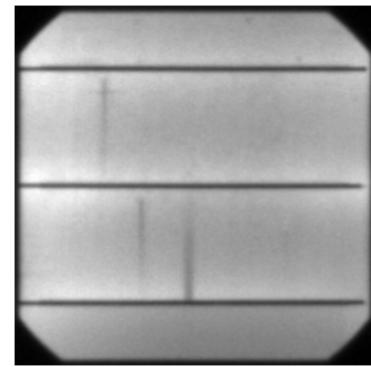
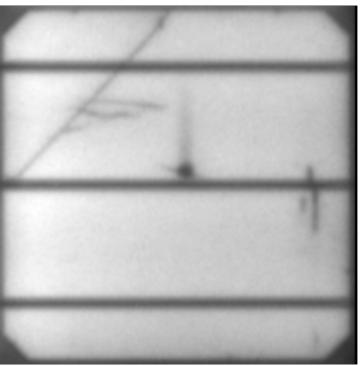
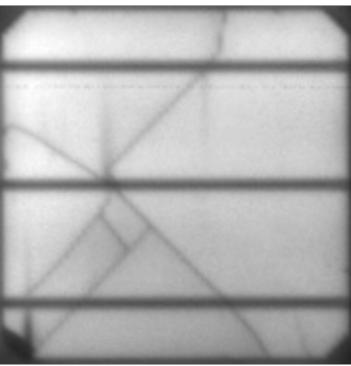
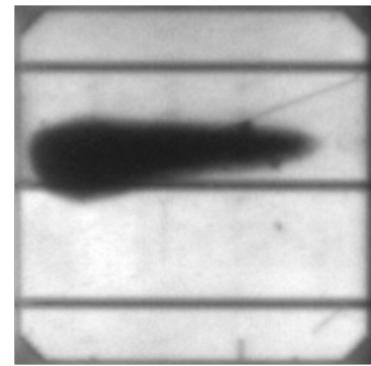
Pre-trained attention



Experiments: dataset [4]

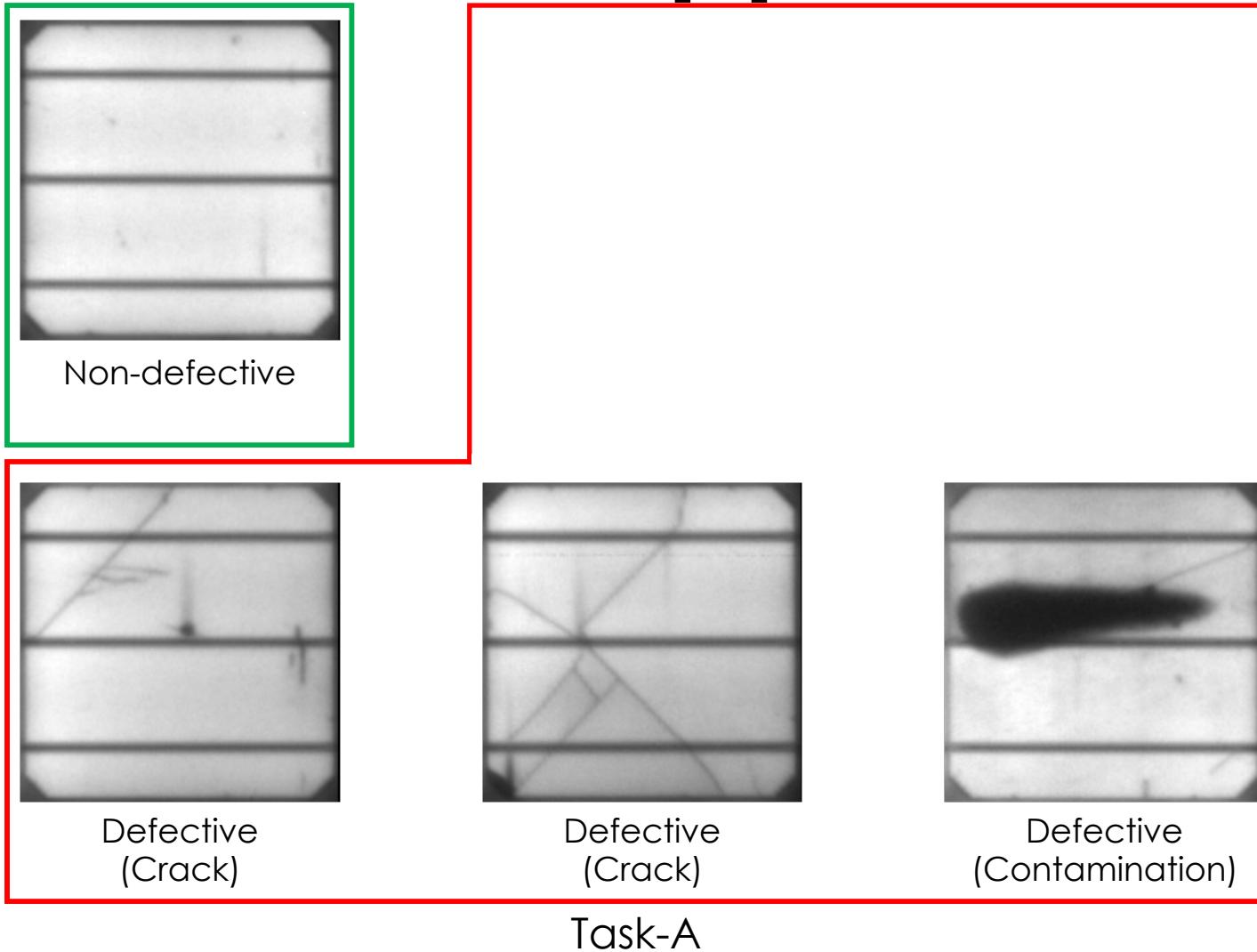


Non-defective

Defective
(1/3-level)Defective
(2/3-level)Defective
(Crack)Defective
(Crack)Defective
(Contamination)

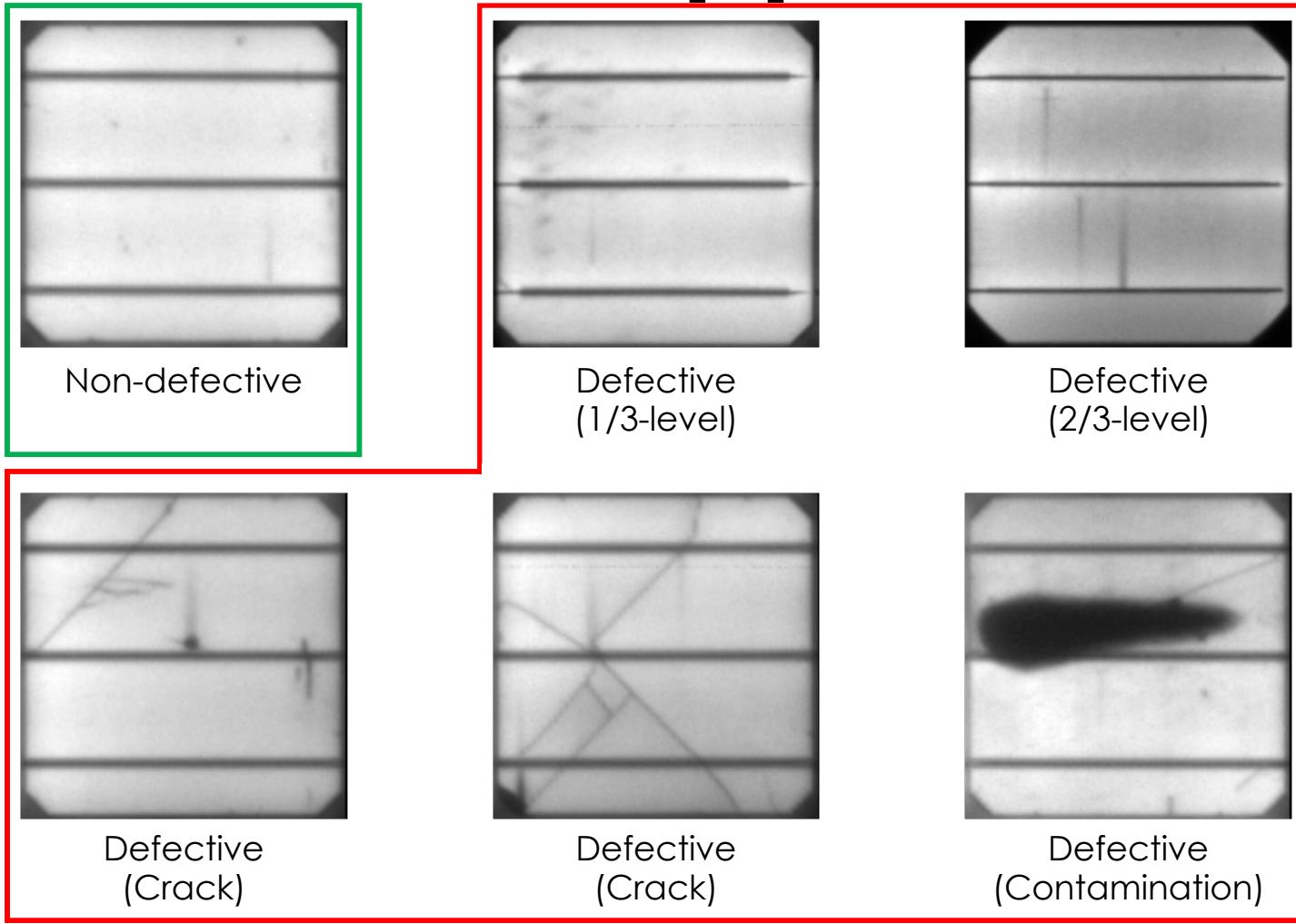
[4] C. Buerhop-Lutz et al., "A benchmark for visual identification of defective solar cells in electroluminescence imagery," EU PVSEC, 2018

Experiments: dataset [4]



[4] C. Buerhop-Lutz et al., "A benchmark for visual identification of defective solar cells in electroluminescence imagery," EU PVSEC, 2018

Experiments: dataset [4]



Task-B

[4] C. Buerhop-Lutz et al., "A benchmark for visual identification of defective solar cells in electroluminescence imagery," EU PVSEC, 2018

Experiments: models

Thresholding model

- Rule: thresholding of extracted feature values

Machine learning models

- Decision Tree (DT) [6]
- Random Forest (RF) [7]
- eXtreme Gradient Boosting (XGB) [8]
- Light Gradient Boosting Machine (LGBM) [9]
- Support Vector Machine (SVM) [10]

Deep learning model

- EfficientNet-B0 (EffNetB0) [11]: end-to-end SOTA classification model

[6] B. Li et al., "Classification and regression trees," *Biometrics*, 1984

[7] T. K. Ho, "Random decision forests," *ICDAR*, 1995

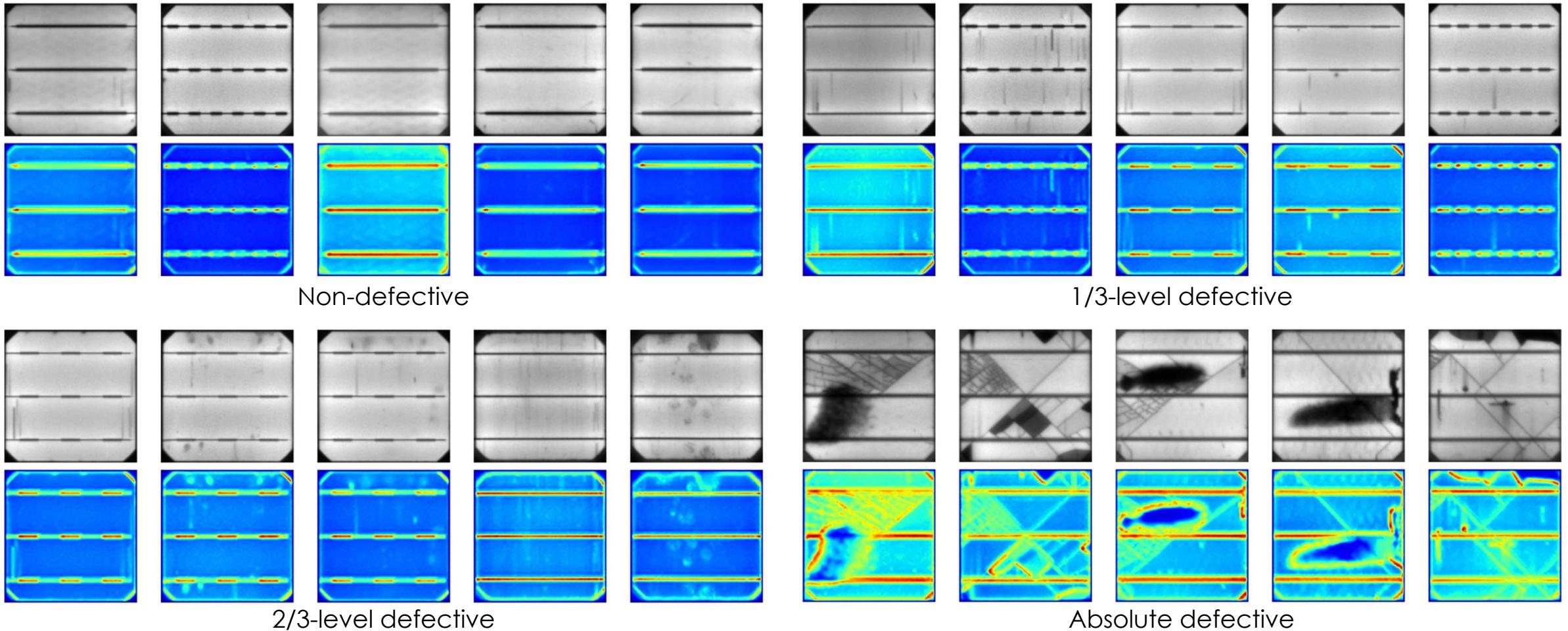
[8] T. Chen et al., "XGBoost: A scalable tree boosting system," *KDD*, 2016

[9] G. Ke et al., "LightGBM: A highly efficient gradient boosting decision tree," *NeurIPS*, 2017

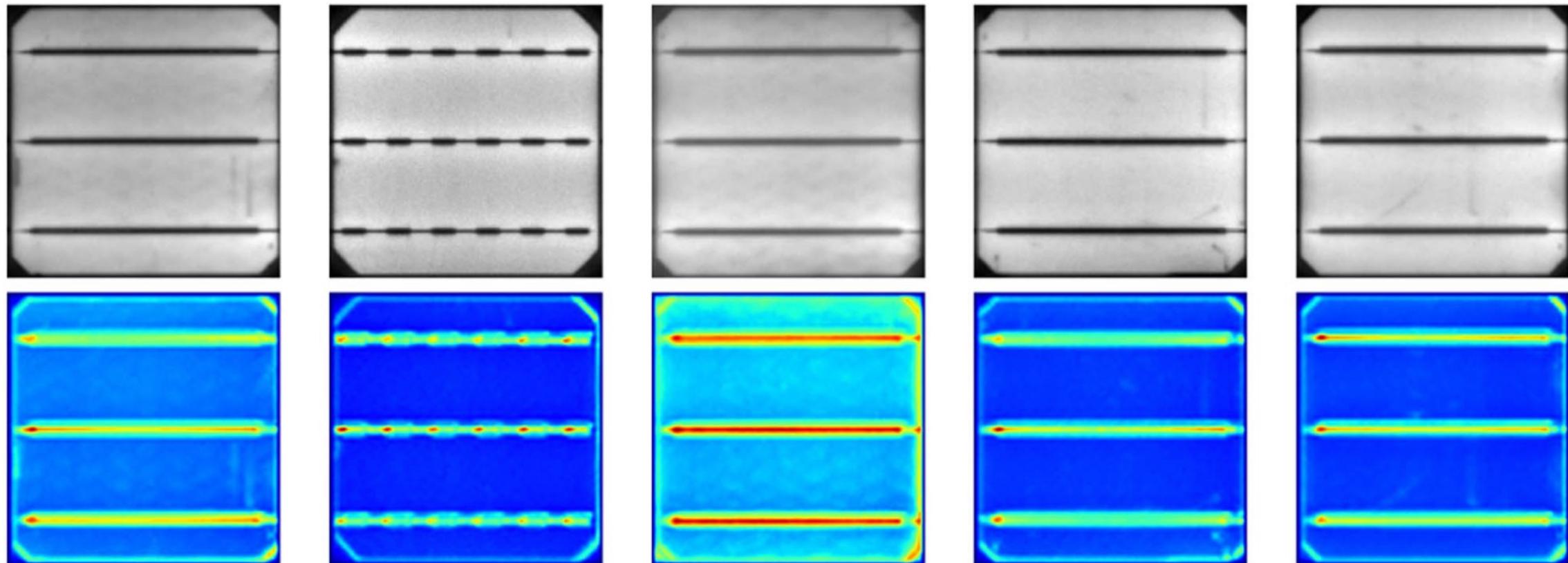
[10] C. Cortes et al., "Support-vector networks," *Machine Learning*, 1995

[11] M. Tan et al., "EfficientNet: Rethinking model scaling for convolutional neural networks," *ICML*, 2019

Attention on solar panels

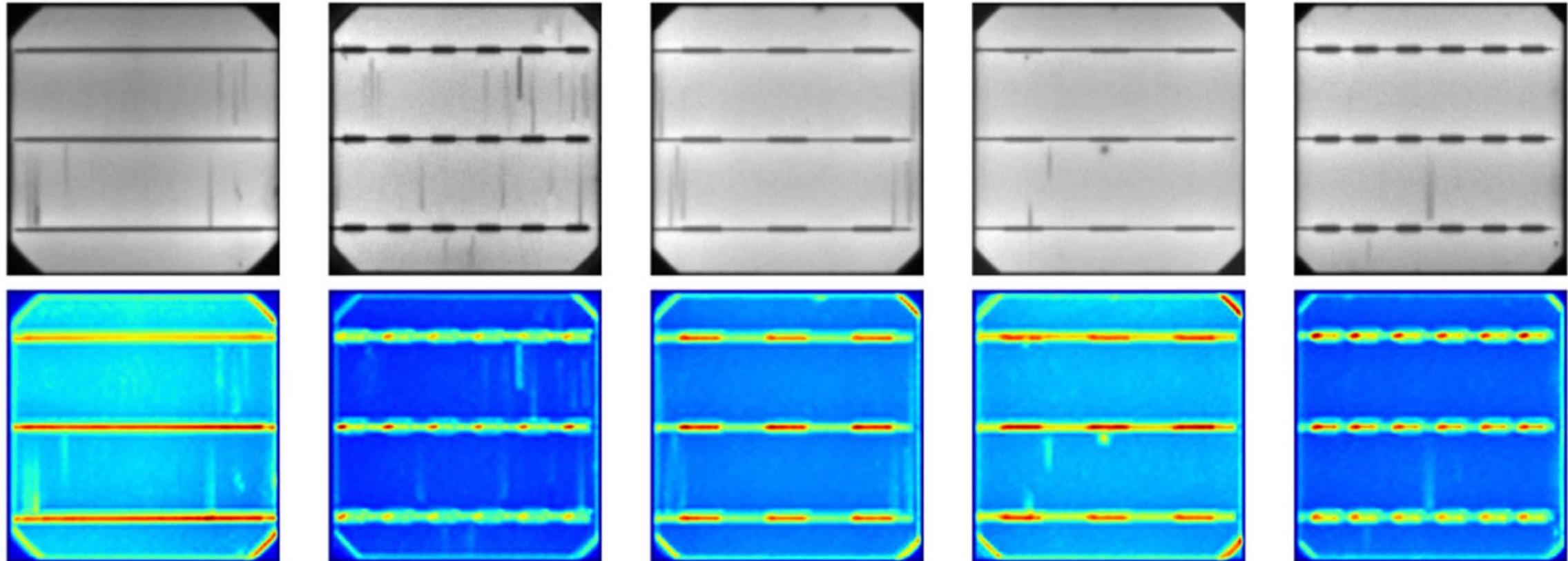


Attention on solar panels



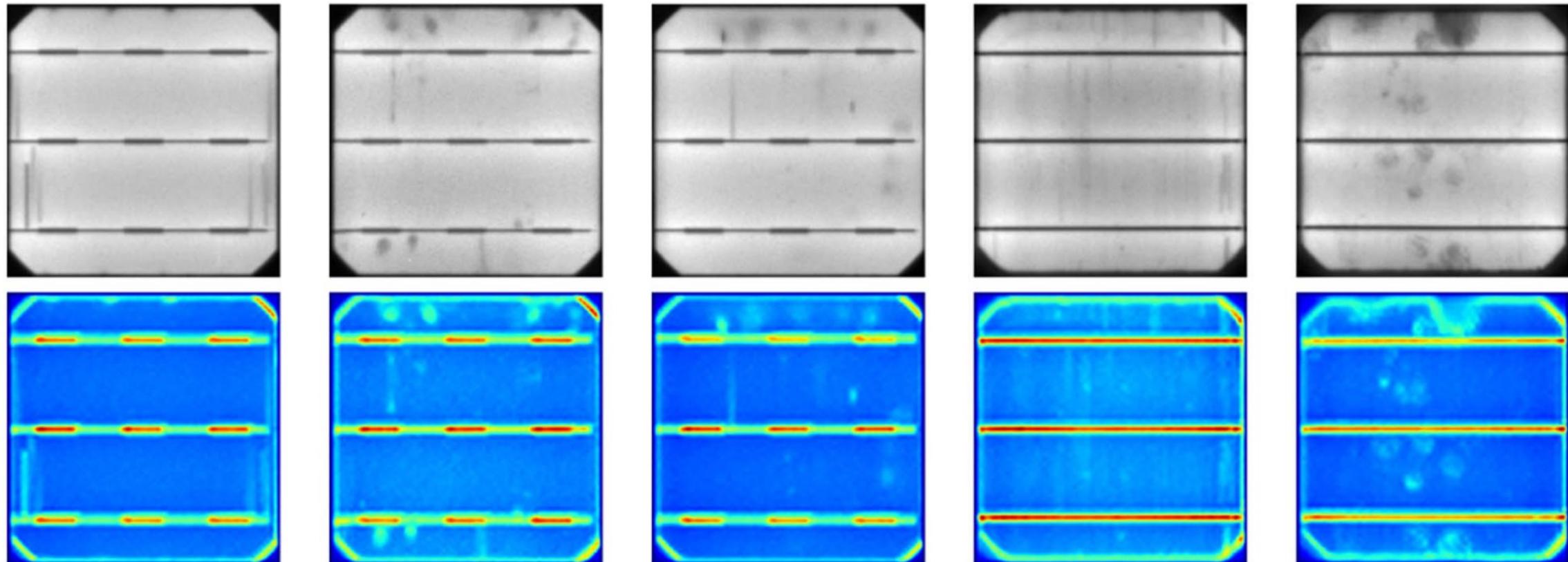
Non-defective

Attention on solar panels



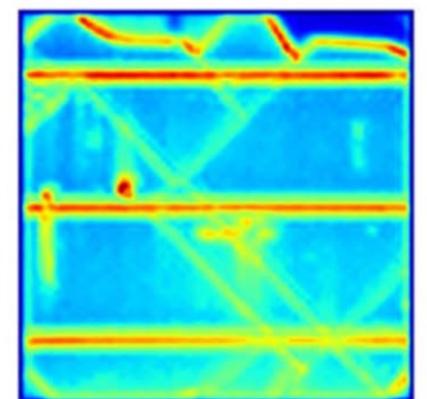
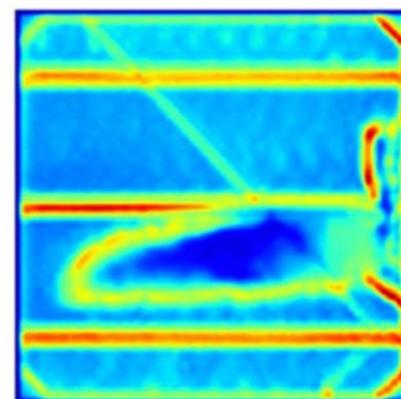
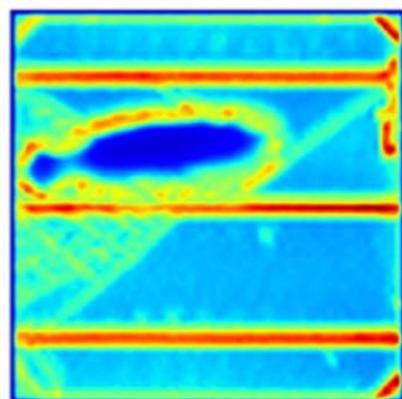
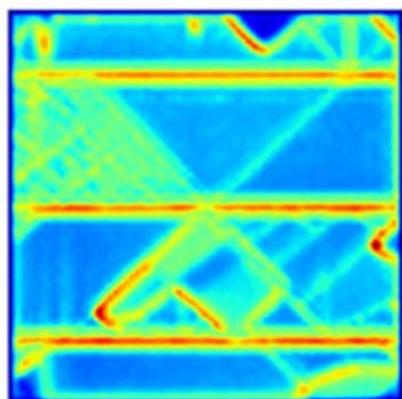
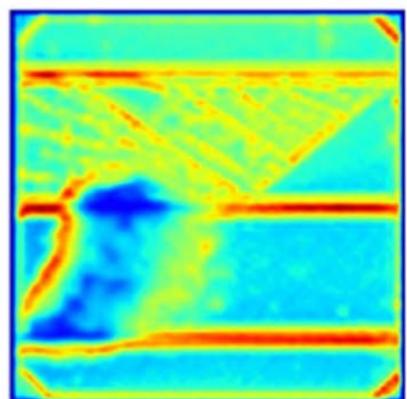
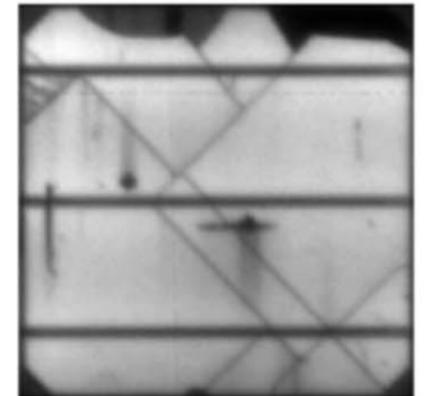
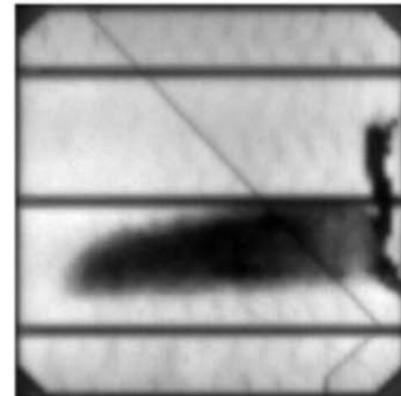
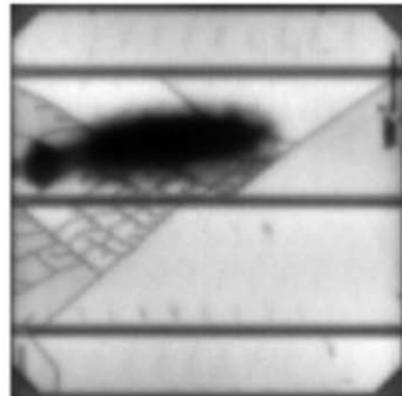
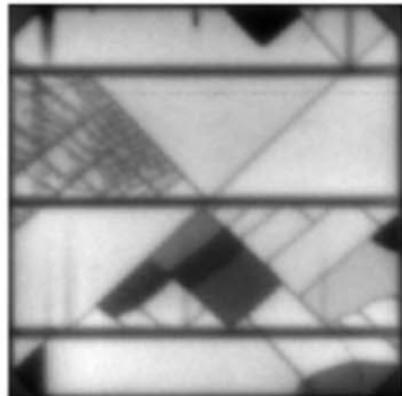
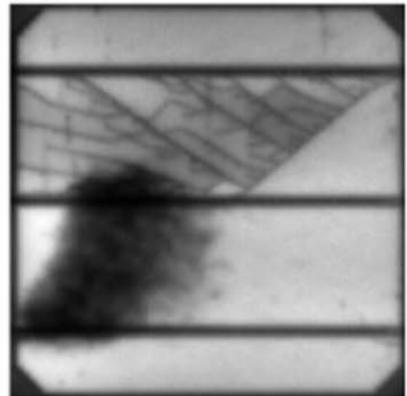
1/3-level defective

Attention on solar panels



2/3-level defective

Attention on solar panels



Absolute defective

Detection performance

5-fold cross validation
Average Maximum

Task	Model	Precision	Recall	F1-score	AUROC
Task-A	Rule	0.836 (0.982) → 0.871 (1.000)	0.954 (0.983) → 0.959 (1.000)	0.888 (0.956) → 0.906 (0.969)	0.874 (0.984) → 0.880 (0.990)
	DT	0.743 (1.000) → 0.726 (0.919)	0.753 (0.935) → 0.842 (1.000)	0.726 (0.840) → 0.774 (0.919)	0.815 (0.901) → 0.820 (0.951)
	RF	0.752 (0.938) → 0.796 (0.922)	0.909 (0.984) → 0.874 (1.000)	0.817 (0.961) → 0.820 (0.939)	0.831 (0.997) → 0.870 (0.973)
	XGB	0.694 (0.922) → 0.821 (0.978)	0.945 (0.983) → 0.771 (1.000)	0.790 (0.937) → 0.788 (0.925)	0.817 (0.992) → 0.844 (0.964)
	LGBM	0.688 (0.827) → 0.754 (0.979)	0.945 (1.000) → 0.900 (1.000)	0.789 (0.905) → 0.784 (0.887)	0.813 (0.981) → 0.829 (0.949)
	SVM	0.557 (0.915) → 0.624 (0.873)	0.958 (1.000) → 0.916 (1.000)	0.675 (0.893) → 0.733 (0.932)	0.748 (0.914) → 0.728 (0.962)
	EffNetB0	0.614 (0.846)	0.869 (0.984)	0.694 (0.846)	0.796 (0.937)
	L-CNN [17]	- (0.904)	- (0.954)	- (0.929)	- (0.934)
	DFB-SVM [17]	- (0.948)	- (0.974)	- (0.961)	- (0.979)
Task-B	Rule	0.816 (0.946) → 0.803 (0.929)	0.870 (0.967) → 0.881 (0.992)	0.839 (0.926) → 0.825 (0.905)	0.851 (0.963) → 0.855 (0.929)
	DT	0.696 (0.843) → 0.719 (0.752)	0.888 (0.991) → 0.931 (0.974)	0.771 (0.835) → 0.811 (0.848)	0.748 (0.866) → 0.794 (0.891)
	RF	0.751 (0.954) → 0.752 (0.888)	0.833 (0.957) → 0.908 (1.000)	0.773 (0.875) → 0.815 (0.892)	0.790 (0.872) → 0.812 (0.943)
	XGB	0.691 (0.810) → 0.752 (0.859)	0.916 (1.000) → 0.889 (0.991)	0.781 (0.888) → 0.809 (0.875)	0.778 (0.899) → 0.801 (0.914)
	LGBM	0.662 (0.801) → 0.708 (0.861)	0.951 (1.000) → 0.929 (1.000)	0.775 (0.890) → 0.797 (0.879)	0.789 (0.905) → 0.791 (0.927)
	SVM	0.684 (0.871) → 0.563 (0.784)	0.811 (1.000) → 0.984 (1.000)	0.715 (0.761) → 0.707 (0.854)	0.656 (0.774) → 0.666 (0.862)
	EffNetB0	0.594 (0.745)	0.938 (1.000)	0.719 (0.752)	0.642 (0.831)
	L-CNN [17]	- (0.807)	- (0.916)	- (0.858)	- (0.935)
	DFB-SVM [17]	- (0.879)	- (0.966)	- (0.916)	- (0.970)

Detection performance

Task	Model				Original image	Attention map
		Precision	Recall	F1-score		
Task-A	Rule	0.836 (0.982) → 0.871 (1.000)	0.954 (0.983) → 0.959 (1.000)	0.888 (0.956) → 0.906 (0.969)	0.874 (0.984)	→ 0.880 (0.990)
	DT	0.743 (1.000) → 0.726 (0.919)	0.753 (0.935) → 0.842 (1.000)	0.726 (0.840) → 0.774 (0.919)	0.815 (0.901)	→ 0.820 (0.951)
	RF	0.752 (0.938) → 0.796 (0.922)	0.909 (0.984) → 0.874 (1.000)	0.817 (0.961) → 0.820 (0.939)	0.831 (0.997)	→ 0.870 (0.973)
	XGB	0.694 (0.922) → 0.821 (0.978)	0.945 (0.983) → 0.771 (1.000)	0.790 (0.937) → 0.788 (0.925)	0.817 (0.992)	→ 0.844 (0.964)
	LGBM	0.688 (0.827) → 0.754 (0.979)	0.945 (1.000) → 0.900 (1.000)	0.789 (0.905) → 0.784 (0.887)	0.813 (0.981)	→ 0.829 (0.949)
	SVM	0.557 (0.915) → 0.624 (0.873)	0.958 (1.000) → 0.916 (1.000)	0.675 (0.893) → 0.733 (0.932)	0.748 (0.914)	→ 0.728 (0.962)
	EffNetB0	0.614 (0.846)	0.869 (0.984)	0.694 (0.846)		0.796 (0.937)
	L-CNN [17]	- (0.904)	- (0.954)	- (0.929)		- (0.934)
	DFB-SVM [17]	- (0.948)	- (0.974)	- (0.961)		- (0.979)
Task-B	Rule	0.816 (0.946) → 0.803 (0.929)	0.870 (0.967) → 0.881 (0.992)	0.839 (0.926) → 0.825 (0.905)	0.851 (0.963)	→ 0.855 (0.929)
	DT	0.696 (0.843) → 0.719 (0.752)	0.888 (0.991) → 0.931 (0.974)	0.771 (0.835) → 0.811 (0.848)	0.748 (0.866)	→ 0.794 (0.891)
	RF	0.751 (0.954) → 0.752 (0.888)	0.833 (0.957) → 0.908 (1.000)	0.773 (0.875) → 0.815 (0.892)	0.790 (0.872)	→ 0.812 (0.943)
	XGB	0.691 (0.810) → 0.752 (0.859)	0.916 (1.000) → 0.889 (0.991)	0.781 (0.888) → 0.809 (0.875)	0.778 (0.899)	→ 0.801 (0.914)
	LGBM	0.662 (0.801) → 0.708 (0.861)	0.951 (1.000) → 0.929 (1.000)	0.775 (0.890) → 0.797 (0.879)	0.789 (0.905)	→ 0.791 (0.927)
	SVM	0.684 (0.871) → 0.563 (0.784)	0.811 (1.000) → 0.984 (1.000)	0.715 (0.761) → 0.707 (0.854)	0.656 (0.774)	→ 0.666 (0.862)
	EffNetB0	0.594 (0.745)	0.938 (1.000)	0.719 (0.752)		0.642 (0.831)
	L-CNN [17]	- (0.807)	- (0.916)	- (0.858)		- (0.935)
	DFB-SVM [17]	- (0.879)	- (0.966)	- (0.916)		- (0.970)

Detection performance

Task	Model	Precision	Recall	F1-score	AUROC
Task-A	Rule	0.836 (0.982) → 0.871 (1.000)	0.954 (0.983) → 0.959 (1.000)	0.888 (0.956) → 0.906 (0.969)	0.874 (0.984) → 0.880 (0.990)
	DT	0.743 (1.000) → 0.726 (0.919)	0.753 (0.935) → 0.842 (1.000)	0.726 (0.840) → 0.774 (0.919)	0.815 (0.901) → 0.820 (0.951)
	RF	0.752 (0.938) → 0.796 (0.922)	0.909 (0.984) → 0.874 (1.000)	0.817 (0.961) → 0.820 (0.939)	0.831 (0.997) → 0.870 (0.973)
	XGB	0.694 (0.922) → 0.771 (0.970)	0.771 (0.970) → 0.771 (0.970)	0.738 (0.925)	0.817 (0.992) → 0.844 (0.964)
	LGBM	0.688 (0.827) → 0.720 (0.922)	0.720 (0.922) → 0.720 (0.922)	0.714 (0.887)	0.813 (0.981) → 0.829 (0.949)
	SVM	0.557 (0.915) → 0.614 (0.937)	0.614 (0.937) → 0.614 (0.937)	0.583 (0.932)	0.748 (0.914) → 0.728 (0.962)
	EffNetB0	- (0.914)	- (0.914)	- (0.914)	0.796 (0.937)
	L-CNN [17]	- (0.934)	- (0.934)	- (0.934)	- (0.934)
	DFB-SVM [17]	- (0.979)	- (0.979)	- (0.979)	- (0.979)
Task-B	Rule	0.816 (0.946) → 0.816 (0.946)	0.825 (0.905)	0.851 (0.963) → 0.855 (0.929)	
	DT	0.696 (0.843) → 0.696 (0.843)	0.711 (0.848)	0.748 (0.866) → 0.794 (0.891)	
	RF	0.751 (0.954) → 0.751 (0.954)	0.715 (0.892)	0.790 (0.872) → 0.812 (0.943)	
	XGB	0.691 (0.810) → 0.771 (0.970)	0.709 (0.875)	0.778 (0.899) → 0.801 (0.914)	
	LGBM	0.662 (0.801) → 0.720 (0.922)	0.707 (0.879)	0.789 (0.905) → 0.791 (0.927)	
	SVM	0.684 (0.871) → 0.694 (0.874)	0.707 (0.854)	0.656 (0.774) → 0.666 (0.862)	
	EffNetB0	0.594 (0.745)	0.938 (1.000)	0.719 (0.752)	0.642 (0.831)
	L-CNN [17]	- (0.807)	- (0.916)	- (0.858)	- (0.935)
	DFB-SVM [17]	- (0.879)	- (0.966)	- (0.916)	- (0.970)

Detection performance

Task	Model	Precision	Recall	F1-score	AUROC
Task-A	Rule	0.836 (0.982) → 0.871 (1.000)	0.954 (0.983) → 0.959 (1.000)	0.888 (0.956) → 0.906 (0.969)	0.874 (0.984) → 0.880 (0.990)
	DT	0.743 (1.000) → 0.726 (0.919)	0.753 (0.935) → 0.842 (1.000)	0.726 (0.840) → 0.774 (0.919)	0.815 (0.901) → 0.820 (0.951)
	RF	0.752 (0.938)	- (0.939)	- (0.939)	0.831 (0.997) → 0.870 (0.973)
	XGB	0.694 (0.922) → 0.748 (0.866)	0.88 (0.925)	0.817 (0.992) → 0.844 (0.964)	
	LGBM	0.688 (0.827) → 0.748 (0.866)	0.84 (0.887)	0.813 (0.981) → 0.829 (0.949)	
	SVM	0.557 (0.915) → 0.614 (0.914)	0.83 (0.932)	0.748 (0.914) → 0.728 (0.962)	
	EffNetB0	- (0.934)	- (0.934)	- (0.934)	0.796 (0.937)
	L-CNN [17]	- (0.934)	- (0.934)	- (0.934)	- (0.934)
	DFB-SVM [17]	- (0.979)	- (0.979)	- (0.979)	- (0.979)
Task-B	Rule	0.816 (0.946) → 0.851 (0.963)	0.825 (0.905) → 0.855 (0.929)	0.825 (0.905) → 0.855 (0.929)	0.851 (0.963) → 0.880 (0.990)
	DT	0.696 (0.843) → 0.778 (0.899)	0.711 (0.848) → 0.801 (0.914)	0.711 (0.848) → 0.801 (0.914)	0.748 (0.866) → 0.794 (0.891)
	RF	0.751 (0.954) → 0.789 (0.905)	0.755 (0.892) → 0.791 (0.927)	0.755 (0.892) → 0.791 (0.927)	0.790 (0.872) → 0.812 (0.943)
	XGB	0.691 (0.810) → 0.656 (0.774)	0.699 (0.875) → 0.666 (0.862)	0.699 (0.875) → 0.666 (0.862)	0.778 (0.899) → 0.801 (0.914)
	LGBM	0.662 (0.801) → 0.594 (0.774)	0.697 (0.879) → 0.666 (0.862)	0.697 (0.879) → 0.666 (0.862)	0.789 (0.905) → 0.791 (0.927)
	SVM	0.684 (0.871) → 0.594 (0.774)	0.697 (0.854) → 0.666 (0.862)	0.697 (0.854) → 0.666 (0.862)	0.656 (0.774) → 0.666 (0.862)
	EffNetB0	- (0.807)	- (0.858)	- (0.858)	0.642 (0.831)
	L-CNN [17]	- (0.879)	- (0.916)	- (0.916)	- (0.935)
	DFB-SVM [17]	- (0.879)	- (0.966)	- (0.966)	- (0.970)

Detection performance

Task	Model		Recall	F1-score	AUROC
Task-A	Rule DT RF XGB LGBM SVM EffNetB0 L-CNN [17] DFB-SVM [17]		Decision Tree (Simple machine learning)	Needs large-scale dataset EfficientNet-B0	(SOTA deep learning)
Task-B	Task-A		0.820 (0.951)	>	0.796 (0.937)
	Task-B		0.794 (0.891)	>	0.642 (0.831)

Training and inference efficiency

2.696×10^{-3} sec for feature extraction

Second	Model	Training	Inference
CPU	Rule	8.367×10^{-2}	3.609×10^{-6}
	DT	3.897×10^{-3}	1.071×10^{-5}
	RF	1.568×10^{-1}	8.135×10^{-5}
	XGB	3.643×10^{-2}	1.935×10^{-5}
	LGBM	4.681×10^{-2}	2.917×10^{-6}
	SVM	6.197×10^{-2}	2.796×10^{-5}
	EffNetB0	8.960×10^2	2.863×10^{-2}
GPU	EffNetB0	6.654×10^1	9.086×10^{-3}

Training and inference efficiency

2.696×10^{-3} sec for feature extraction

Second	Model	Training	Inference
CPU	Rule	8.367×10^{-2}	3.609×10^{-6}
	DT	3.897×10^{-3}	1.071×10^{-5}
	RF	1.568×10^{-1}	8.135×10^{-5}
	XGB	3.643×10^{-2}	1.935×10^{-5}
	LGBM	4.681×10^{-2}	2.917×10^{-6}
	SVM	6.197×10^{-2}	2.796×10^{-5}
	EffNetB0	8.960×10^2	2.863×10^{-2}
GPU	EffNetB0	6.654×10^1	9.086×10^{-3}

$\times 230k$



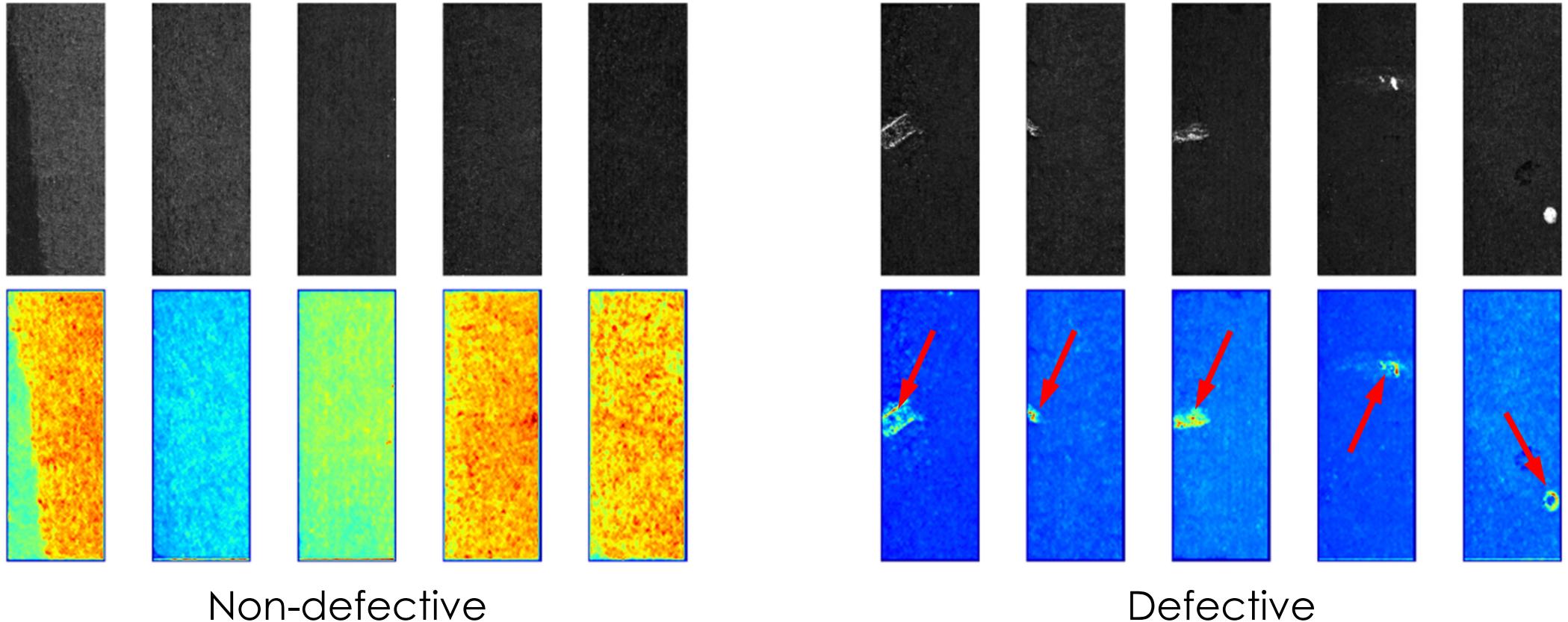
Training and inference efficiency

2.696×10^{-3} sec for feature extraction

Second	Model	Training	Inference
CPU	Rule	8.367×10^{-2}	3.609×10^{-6}
	DT	3.897×10^{-3}	1.071×10^{-5}
	RF	1.568×10^{-1}	8.135×10^{-5}
	XGB	3.643×10^{-2}	1.935×10^{-5}
	LGBM	4.681×10^{-2}	2.917×10^{-6}
	SVM	6.197×10^{-2}	2.796×10^{-5}
	EffNetB0	8.960×10^2	2.863×10^{-2}
GPU	EffNetB0	6.654×10^1	9.086×10^{-3}

$\times 3k$

Experiment on another dataset [12]

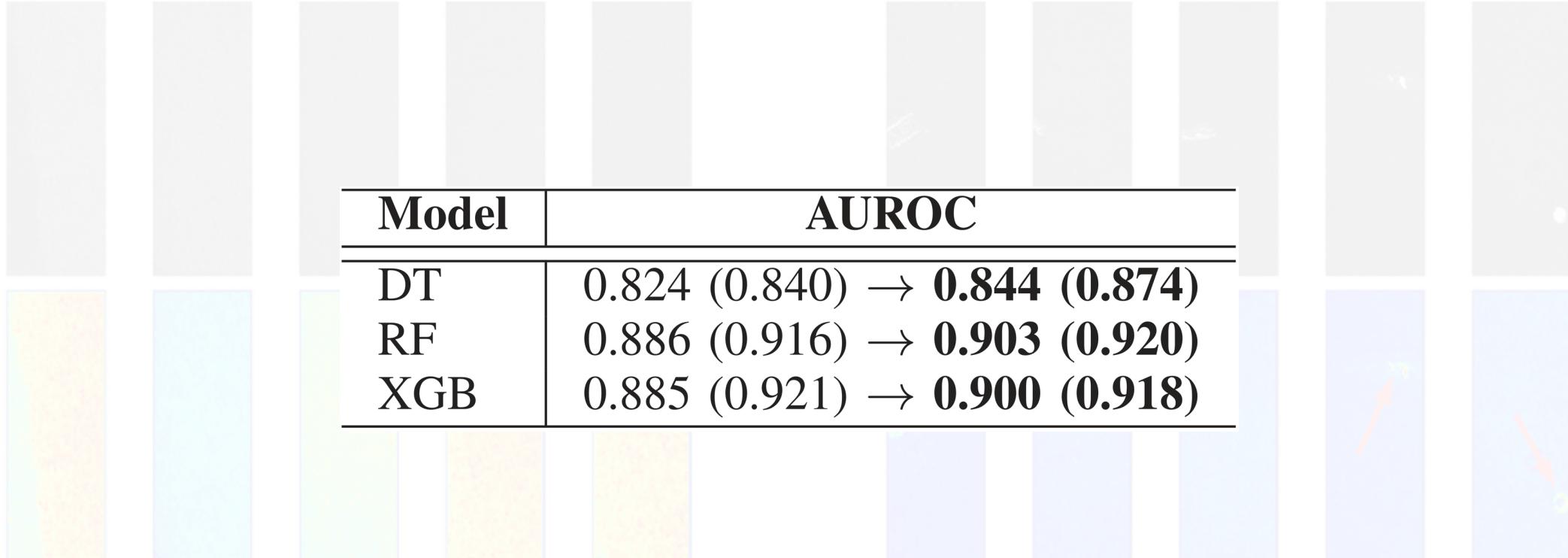


[12] J. Božič et al., "Mixed supervision for surface- defect detection: From weakly to fully supervised learning," *Computers in Industry*, 2021.

Experiment on another dataset [12]

Model	AUROC
DT	0.824 (0.840) → 0.844 (0.874)
RF	0.886 (0.916) → 0.903 (0.920)
XGB	0.885 (0.921) → 0.900 (0.918)

Non-defective Defective



[12] J. Božič et al., "Mixed supervision for surface- defect detection: From weakly to fully supervised learning," *Computers in Industry*, 2021.

Conclusions

- Simple yet powerful method for a real world problem
 - Attention mechanism recycling with 13 statistical features
 - Outperforms SOTA defect detection
 - Serves the purpose of sustainable green energy
- Applicable to other visual inspections
 - Surface defect detection in steel, film manufacturing, etc.

Future works

- Analysis of attention dependency on
 - Training dataset and
 - Neural network structure
- Attention recycling combined anomaly detection
 - Unsupervised anomaly detection
 - Cost-effective training strategy