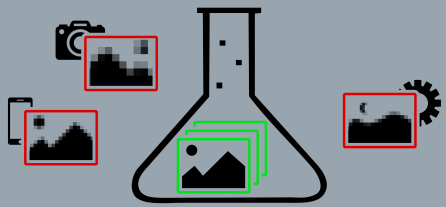
FRIEDRICH-ALEXANDER
UNIVERSITÄT
ERLANGEN-NÜRNBERG
TECHNISCHE FAKULTÄT

Reliable Camera Model Identification Using Sparse Gaussian Processes

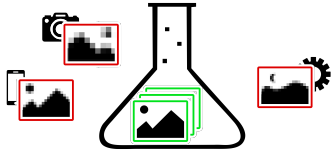
Benedikt Lorch, Franziska Schirrmacher, Anatol Maier, Christian Riess
IT Security Infrastructures Lab
Friedrich-Alexander University Erlangen-Nuremberg
December 10, 2021



Training-test mismatches in multimedia forensics

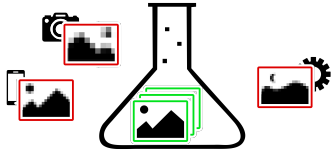
Training-test mismatches in multimedia forensics

Training-test mismatches in multimedia forensics



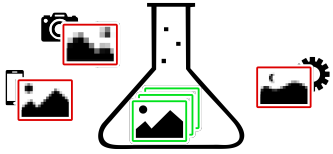
- ML methods are **sensitive to training-test mismatches**

Training-test mismatches in multimedia forensics



- ML methods are **sensitive to training-test mismatches**
- Forensic tools are particularly exposed to **images from unknown origins**
  $\Rightarrow$ High risk of silent failure

Training-test mismatches in multimedia forensics

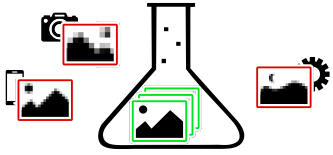


- ML methods are **sensitive to training-test mismatches**
- Forensic tools are particularly exposed to **images from unknown origins**
  $\Rightarrow$ High risk of silent failure

Mitigating training-test mismatches

- Generalization (open challenge): data augmentation, domain-invariant features, ...

Training-test mismatches in multimedia forensics



- ML methods are **sensitive to training-test mismatches**
- Forensic tools are particularly exposed to **images from unknown origins**
  $\Rightarrow$ High risk of silent failure

Mitigating training-test mismatches

- Generalization (open challenge): data augmentation, domain-invariant features, ...
- Proposal: **Create reliable detectors that express uncertainty in unfamiliar situations**
  $\Rightarrow$ Uncertainty allows analyst to quantify when to trust in the method's prediction

Reliable camera model identification with Gaussian process classifier

- WIFS'20: Reliable Bayesian detector for JPEG double compression[1]

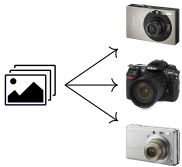
[1] B. Lorch, A. Maier, and C. Riess, "Reliable JPEG forensics via model uncertainty," in *IEEE International Workshop on Information Forensics and Security*, Dec. 2020.

Reliable camera model identification with Gaussian process classifier

- WIFS'20: Reliable Bayesian detector for JPEG double compression[1]

This work

- Task: Camera model identification



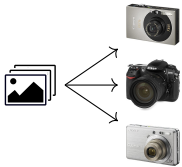
[1] B. Lorch, A. Maier, and C. Riess, "Reliable JPEG forensics via model uncertainty," in *IEEE International Workshop on Information Forensics and Security*, Dec. 2020.

Reliable camera model identification with Gaussian process classifier

- WIFS'20: Reliable Bayesian detector for JPEG double compression[1]

This work

- Task: Camera model identification
- Gaussian process provides predictive distribution (mean + uncertainty)



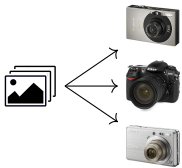
[1] B. Lorch, A. Maier, and C. Riess, "Reliable JPEG forensics via model uncertainty," in *IEEE International Workshop on Information Forensics and Security*, Dec. 2020.

Reliable camera model identification with Gaussian process classifier

- WIFS'20: Reliable Bayesian detector for JPEG double compression[1]

This work

- Task: Camera model identification
- Gaussian process provides predictive distribution (mean + uncertainty)
- Research question: **Can we identify out-of-distribution images based on uncertainty?**
  - Unknown camera models
  - Unseen post-processing



[1] B. Lorch, A. Maier, and C. Riess, "Reliable JPEG forensics via model uncertainty," in *IEEE International Workshop on Information Forensics and Security*, Dec. 2020.

Gaussian Processes

Basic properties of Gaussian processes (GPs)

- GP defines a **probability distribution over functions**
- Non-parametric
- Optimize hyper-parameters by maximizing the marginal likelihood

Basic properties of Gaussian processes (GPs)

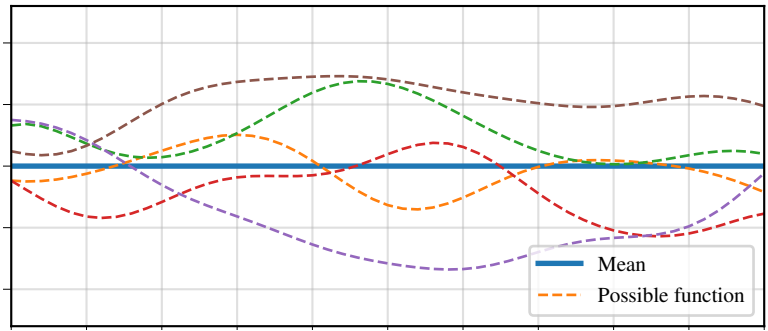
- GP defines a **probability distribution over functions**
- Non-parametric
- Optimize hyper-parameters by maximizing the marginal likelihood
- Goal: Obtain **predictive distribution** instead of point estimate
  $\Rightarrow$ Mean gives prediction, variance indicates uncertainty

Basic properties of Gaussian processes (GPs)

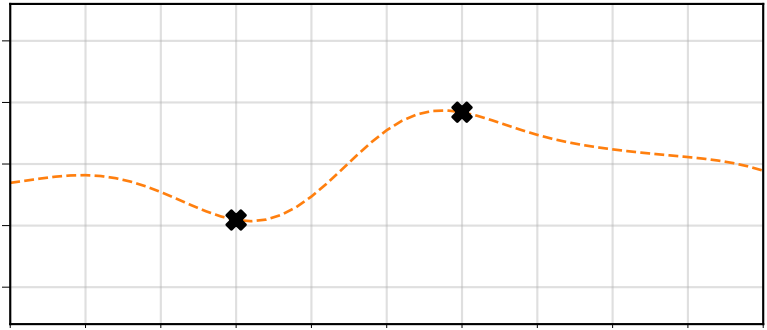
- GP defines a **probability distribution over functions**
- Non-parametric
- Optimize hyper-parameters by maximizing the marginal likelihood
- Goal: Obtain **predictive distribution** instead of point estimate
  $\Rightarrow$ Mean gives prediction, variance indicates uncertainty
- Drawback: Exact inference requires $\mathcal{O}(n^3)$ for n training points

Regression toy example: GP prior

- Specify function prior: mean, covariance (through kernel)
- Kernel defines smoothness, periodicity, ...

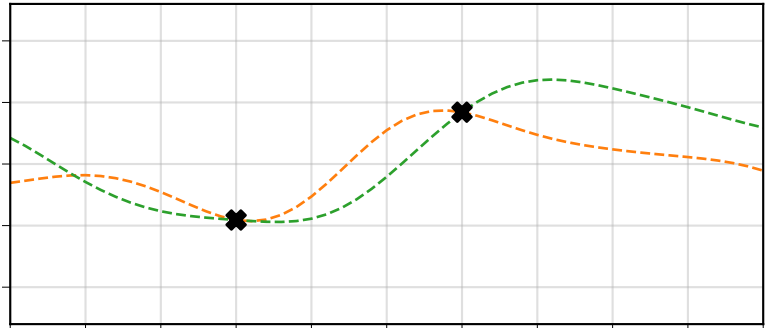


Regression toy example: GP posterior



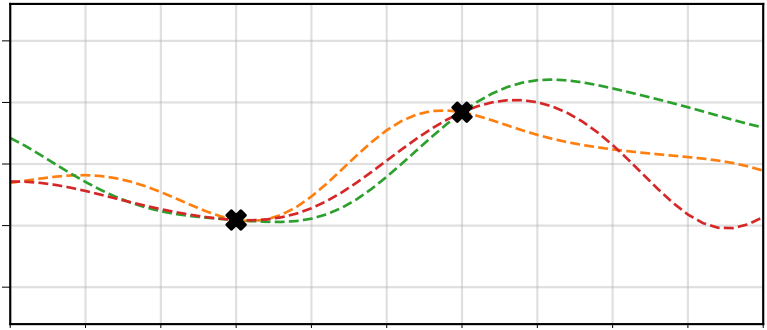
- Condition possible functions on observations

Regression toy example: GP posterior



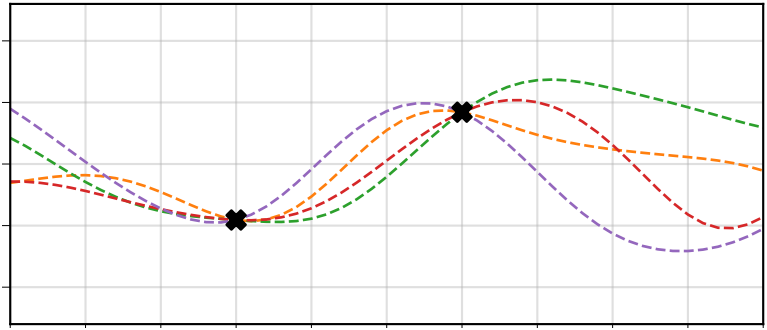
- Condition possible functions on observations

Regression toy example: GP posterior



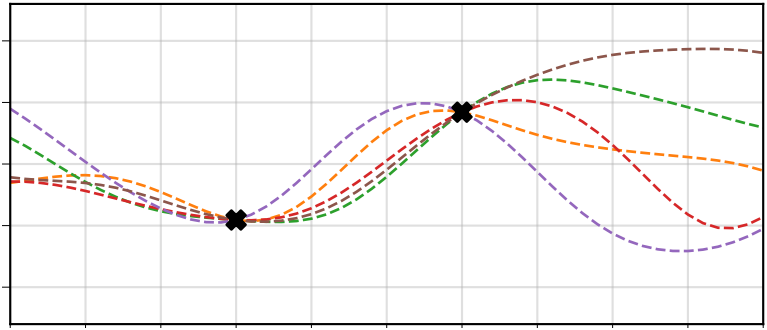
- Condition possible functions on observations

Regression toy example: GP posterior



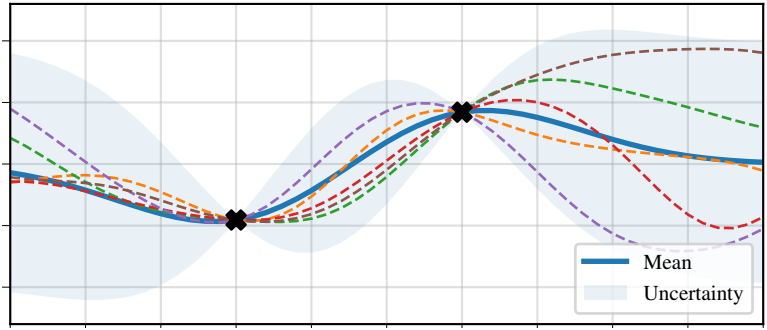
- Condition possible functions on observations

Regression toy example: GP posterior



- Condition possible functions on observations

Regression toy example: GP posterior

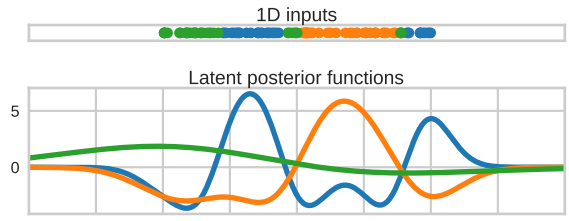


- Condition possible functions on observations
- Close to observations: Functions agree $\Rightarrow$ low uncertainty
- Far from observations: Functions disagree $\Rightarrow$ high uncertainty

From regression to multi-class classification: Another toy example

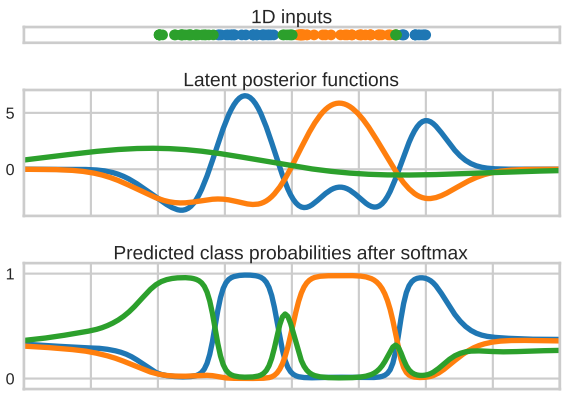
From regression to multi-class classification: Another toy example

- One latent function per class



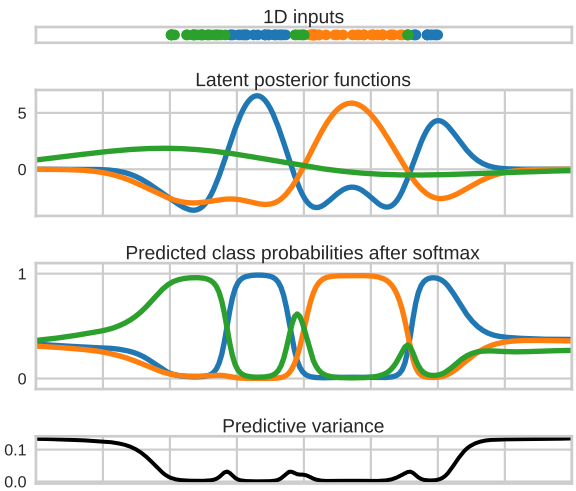
From regression to multi-class classification: Another toy example

- One latent function per class
- Rescale latent functions through softmax to obtain class probabilities



From regression to multi-class classification: Another toy example

- One latent function per class
- Rescale latent functions through softmax to obtain class probabilities



Proposed Gaussian process classifier for camera model identification

- One latent function with RBF kernel per class
- Approximate latent posterior by variational distribution[2]
- Tune hyper-parameters by maximizing the evidence lower bound (ELBO)[3]
- Implementation with GPyTorch[4]

[2] M. Titsias, "Variational learning of inducing variables in sparse Gaussian processes," in *International Conference on Artificial Intelligence and Statistics*, ser. Proceedings of Machine Learning Research, vol. 5, Apr. 2009, pp. 567–574.

[3] J. Hensman, A. G. de G. Matthews, and Z. Ghahramani, "Scalable variational Gaussian process classification," in *International Conference on Artificial Intelligence and Statistics*, vol. 38, May 2015, pp. 351–360.

[4] J. R. Gardner, G. Pleiss, D. Bindel, *et al.*, "GPyTorch: Blackbox matrix-matrix Gaussian process inference with GPU acceleration," in *Advances in Neural Information Processing Systems*, Dec. 2018, pp. 7587–7597.

Experiments and Results

Experimental Setup

- Known: 10 randomly selected camera models from Dresden database
- Unknown: 18 remaining camera models

Experimental Setup

- Known: 10 randomly selected camera models from Dresden database
- Unknown: 18 remaining camera models
- Split camera devices and scenes into disjoint training and test sets

Experimental Setup

- Known: 10 randomly selected camera models from Dresden database
- Unknown: 18 remaining camera models
- Split camera devices and scenes into disjoint training and test sets
- Extract SPAM features from full-resolution images

Experimental Setup

- Known: 10 randomly selected camera models from Dresden database
- Unknown: 18 remaining camera models
- Split camera devices and scenes into disjoint training and test sets
- Extract SPAM features from full-resolution images
- Repeat 5x with different training devices and scenes

Experimental Setup

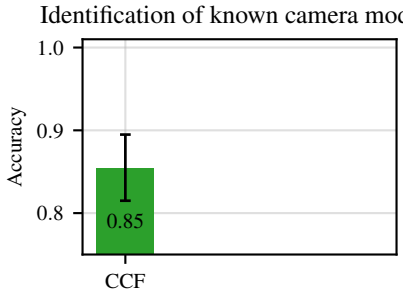
- Known: 10 randomly selected camera models from Dresden database
- Unknown: 18 remaining camera models
- Split camera devices and scenes into disjoint training and test sets
- Extract SPAM features from full-resolution images
- Repeat 5x with different training devices and scenes

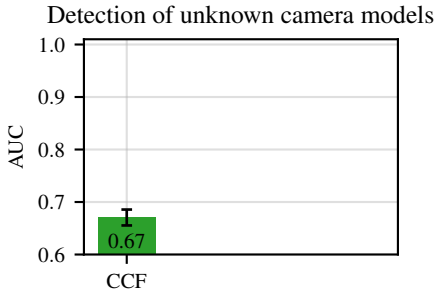
Evaluation metrics

- Classification accuracy for known camera models
- Ability to reject images from unknown camera models (AUC)

Comparison of Gaussian process classifier (GPC) to related methods

- CCF[5]: Combines one-class SVMs with multi-class SVM



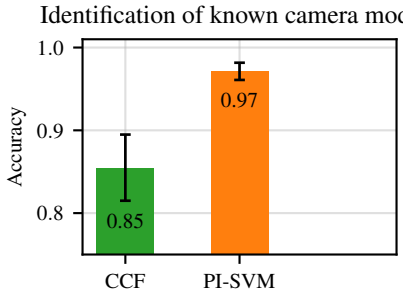


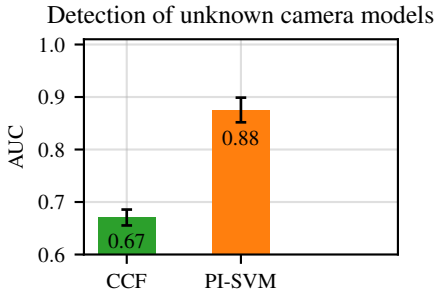
[5] B. Wang, X. Kong, and X. You, "Source camera identification using support vector machines," in *Advances in Digital Forensics V*, Sep. 2009, pp. 107–118

[6] P. R. Mendes Júnior, L. Bondi, P. Bestagini, *et al.*, "An in-depth study on open-set camera model identification," *IEEE Access*, vol. 7, pp. 180 713–180 726, Jun. 2019

Comparison of Gaussian process classifier (GPC) to related methods

- CCF[5]: Combines one-class SVMs with multi-class SVM
- PI-SVM[6]: Calibrates posterior scores using extreme value theory



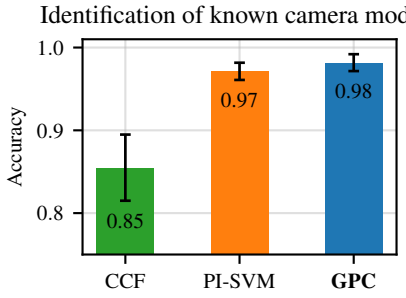


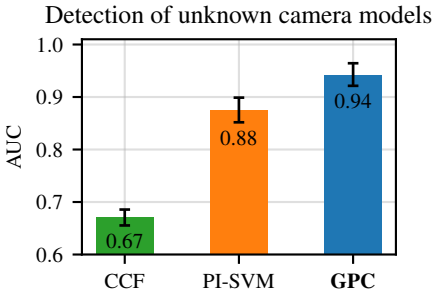
[5] B. Wang, X. Kong, and X. You, "Source camera identification using support vector machines," in *Advances in Digital Forensics V*, Sep. 2009, pp. 107–118

[6] P. R. Mendes Júnior, L. Bondi, P. Bestagini, *et al.*, "An in-depth study on open-set camera model identification," *IEEE Access*, vol. 7, pp. 180 713–180 726, Jun. 2019

Comparison of Gaussian process classifier (GPC) to related methods

- CCF[5]: Combines one-class SVMs with multi-class SVM
- PI-SVM[6]: Calibrates posterior scores using extreme value theory





[5] B. Wang, X. Kong, and X. You, "Source camera identification using support vector machines," in *Advances in Digital Forensics V*, Sep. 2009, pp. 107–118

[6] P. R. Mendes Júnior, L. Bondi, P. Bestagini, *et al.*, "An in-depth study on open-set camera model identification," *IEEE Access*, vol. 7, pp. 180 713–180 726, Jun. 2019

Sparse GP: Reduce cubic complexity using *inducing points*

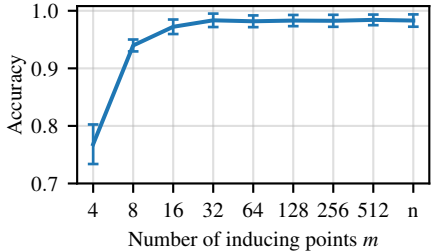
- Naive inference requires $\mathcal{O}(n^3)$ (here $n = 1350$ training samples)

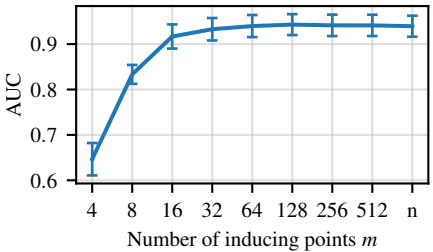
Sparse GP: Reduce cubic complexity using *inducing points*

- Naive inference requires $\mathcal{O}(n^3)$ (here $n = 1350$ training samples)
- Reduce complexity to $\mathcal{O}(n \cdot m^2)$ by compressing n observations into m *inducing points*

Sparse GP: Reduce cubic complexity using *inducing points*

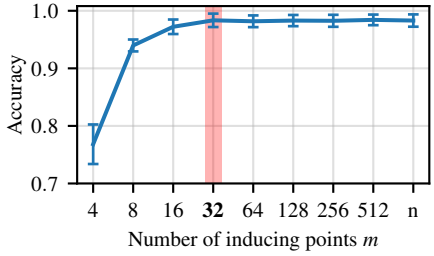
- Naive inference requires $\mathcal{O}(n^3)$ (here $n = 1350$ training samples)
- Reduce complexity to $\mathcal{O}(n \cdot m^2)$ by compressing n observations into m *inducing points*

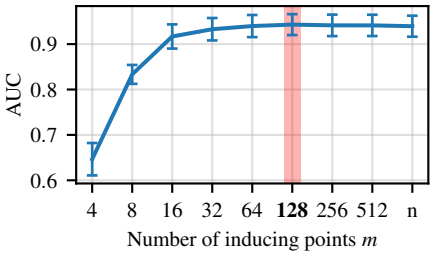




Sparse GP: Reduce cubic complexity using *inducing points*

- Naive inference requires $\mathcal{O}(n^3)$ (here $n = 1350$ training samples)
- Reduce complexity to $\mathcal{O}(n \cdot m^2)$ by compressing n observations into m *inducing points*

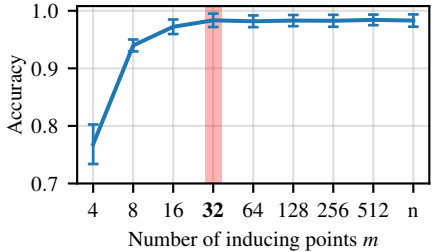


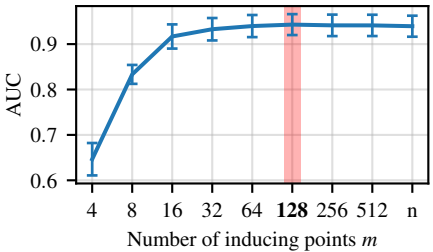


- Accuracy saturates with $m = 32$, AUC with $m = 128$

Sparse GP: Reduce cubic complexity using *inducing points*

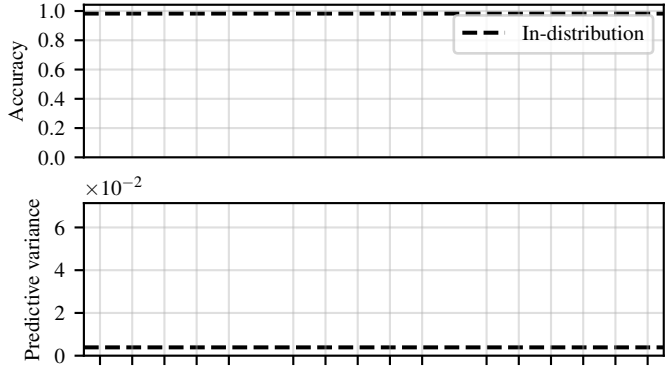
- Naive inference requires $\mathcal{O}(n^3)$ (here $n = 1350$ training samples)
- Reduce complexity to $\mathcal{O}(n \cdot m^2)$ by compressing n observations into m *inducing points*



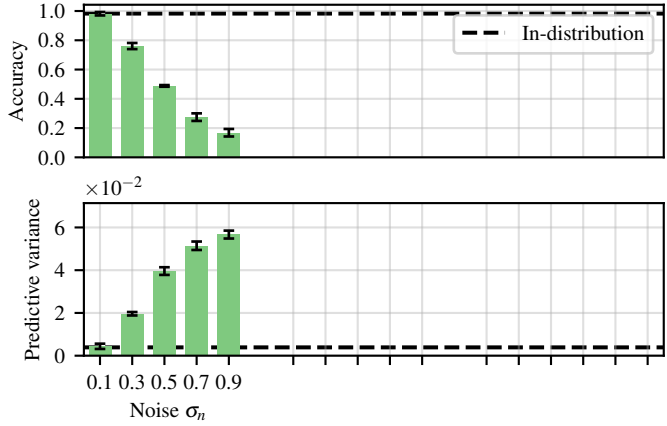


- Accuracy saturates with $m = 32$, AUC with $m = 128$
- $\Rightarrow$ Sparse GP could be scaled to more cameras/training samples

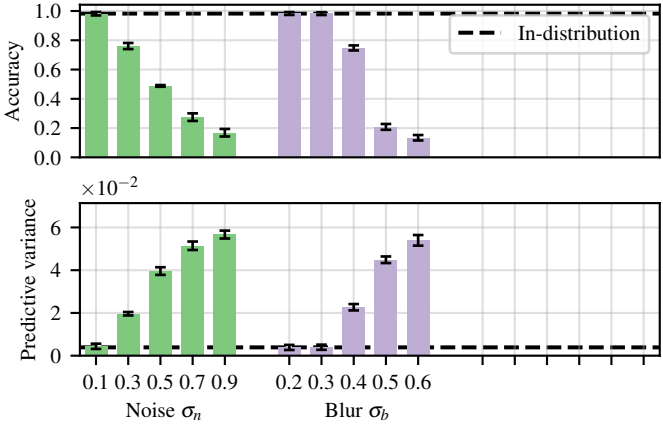
Uncertainty allows detecting images with unseen post-processing



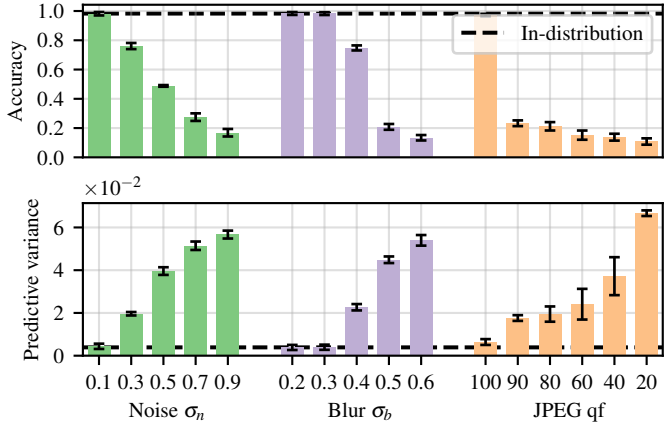
Uncertainty allows detecting images with unseen post-processing



Uncertainty allows detecting images with unseen post-processing



Uncertainty allows detecting images with unseen post-processing

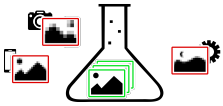


⇒ Accuracy decreases with unseen post-processing, but uncertainty allows detecting these images

Conclusion

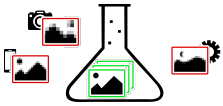
Reliable forensic detectors through predictive uncertainty

- Forensic tools often applied to images from unknown origins
- High risk of silent failure due to training-test mismatch



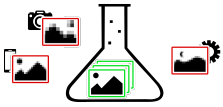
Reliable forensic detectors through predictive uncertainty

- Forensic tools often applied to images from unknown origins
- High risk of silent failure due to training-test mismatch
- Bayesian detectors enable detecting out-of-distribution images
- Gaussian process classifier for camera model identification
  - High classification accuracy
  - Enables rejecting images from unknown cameras and unseen post-processing



Reliable forensic detectors through predictive uncertainty

- Forensic tools often applied to images from unknown origins
- High risk of silent failure due to training-test mismatch
- Bayesian detectors enable detecting out-of-distribution images
- Gaussian process classifier for camera model identification
  - High classification accuracy
  - Enables rejecting images from unknown cameras and unseen post-processing



Long term goal

- Foster research on **reliable** and **trustworthy** learning-based methods

`github.com/btlorch/gaussian-processes-for-camera-model-identification`

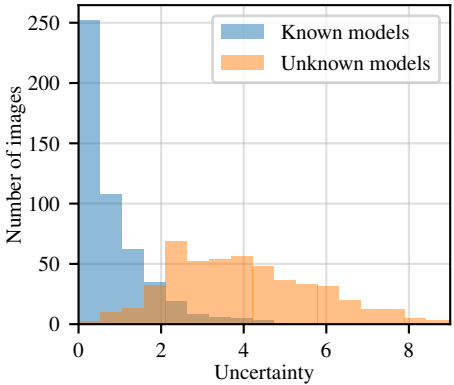
Thank you

Image sources

- Canon Ixus 70
- Nikon D200
- Sony DSC-W170
- PyTorch logo
- GitHub logo

Detection of unknown camera models AUC

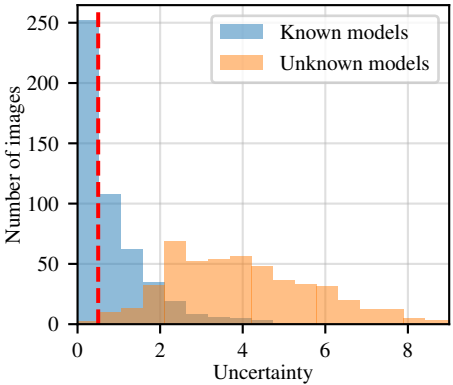
Negative class: Known models Positive class: Unknown models

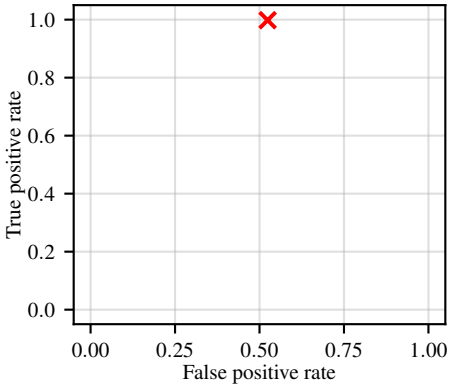


Synthetic example

Detection of unknown camera models AUC

Negative class: Known models Positive class: Unknown models

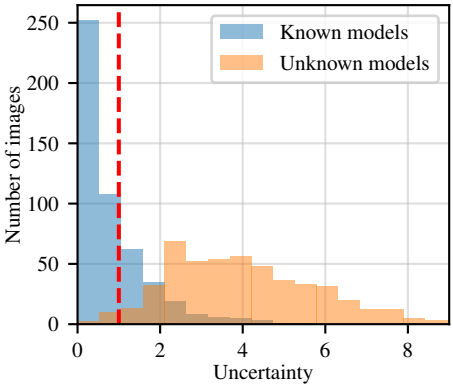


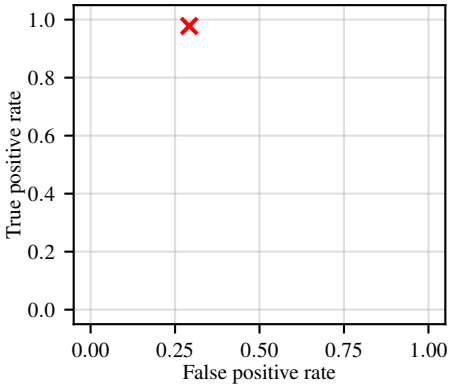


Synthetic example

Detection of unknown camera models AUC

Negative class: Known models Positive class: Unknown models

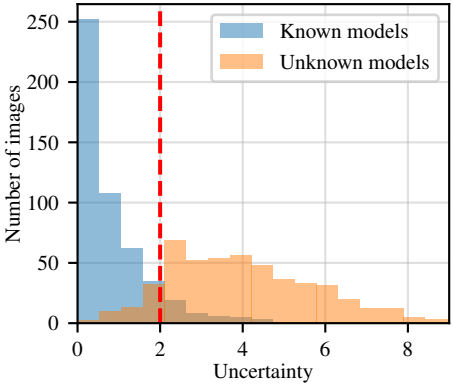


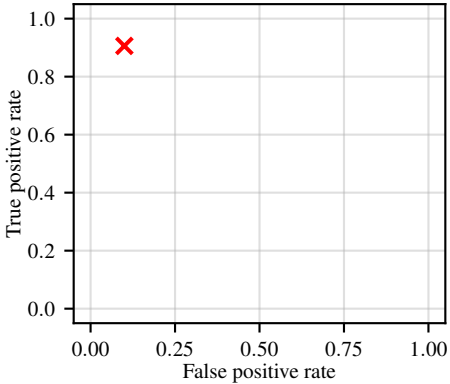


Synthetic example

Detection of unknown camera models AUC

Negative class: Known models Positive class: Unknown models

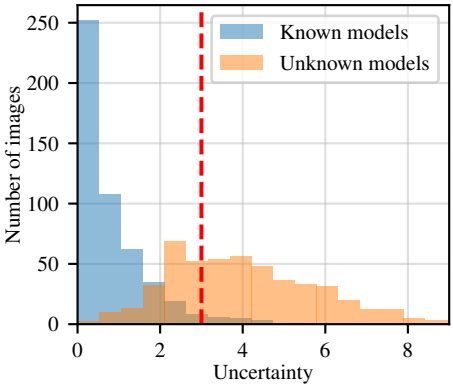


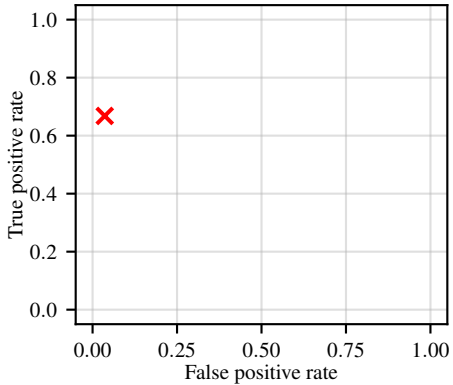


Synthetic example

Detection of unknown camera models AUC

Negative class: Known models Positive class: Unknown models

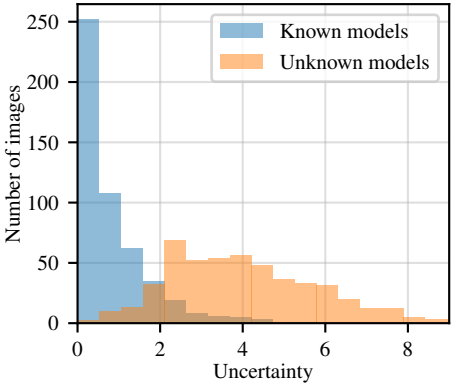


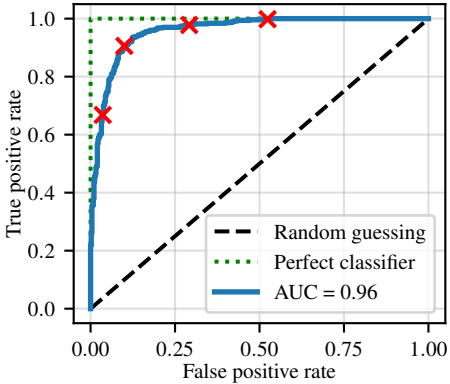


Synthetic example

Detection of unknown camera models AUC

Negative class: Known models Positive class: Unknown models

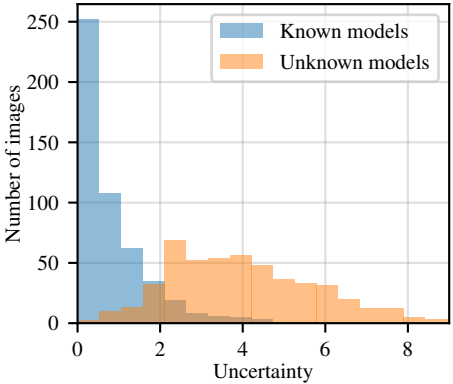


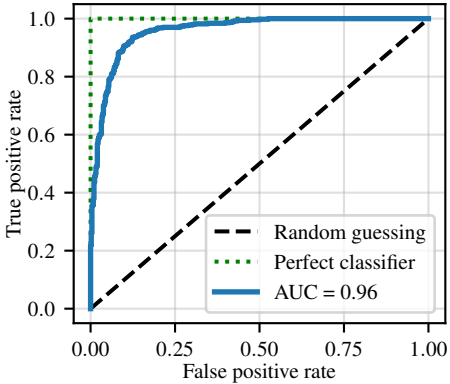


Synthetic example

Detection of unknown camera models AUC

Negative class: Known models Positive class: Unknown models





Synthetic example