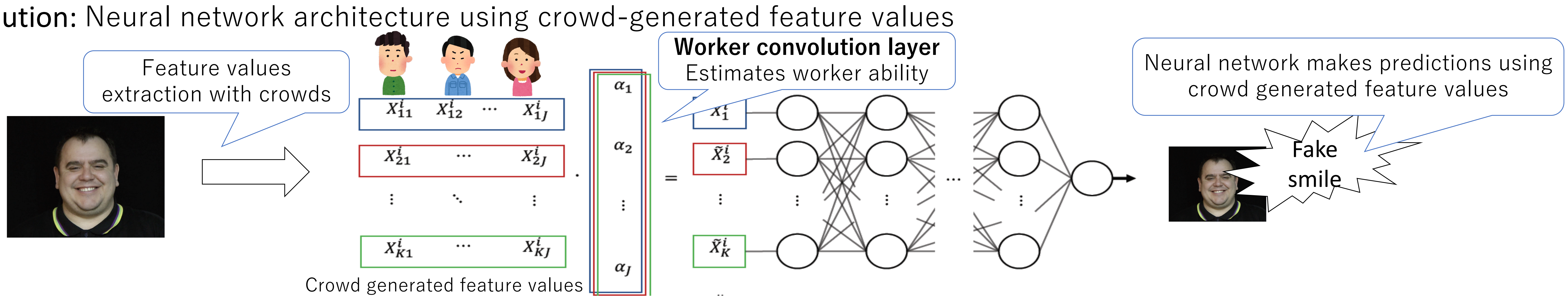
CrowNN: Human-in-the-loop Network with Crowd-generated Inputs

Yusuke Sakata (Kyoto University) Yukino Baba (Tsukuba University) Hisashi Kashima (Kyoto University/RIKEN AIP)

Summary

Goal: Classification problem using feature values given by crowdsourcing workers with different capabilities

Solution: Neural network architecture using crowd-generated feature values

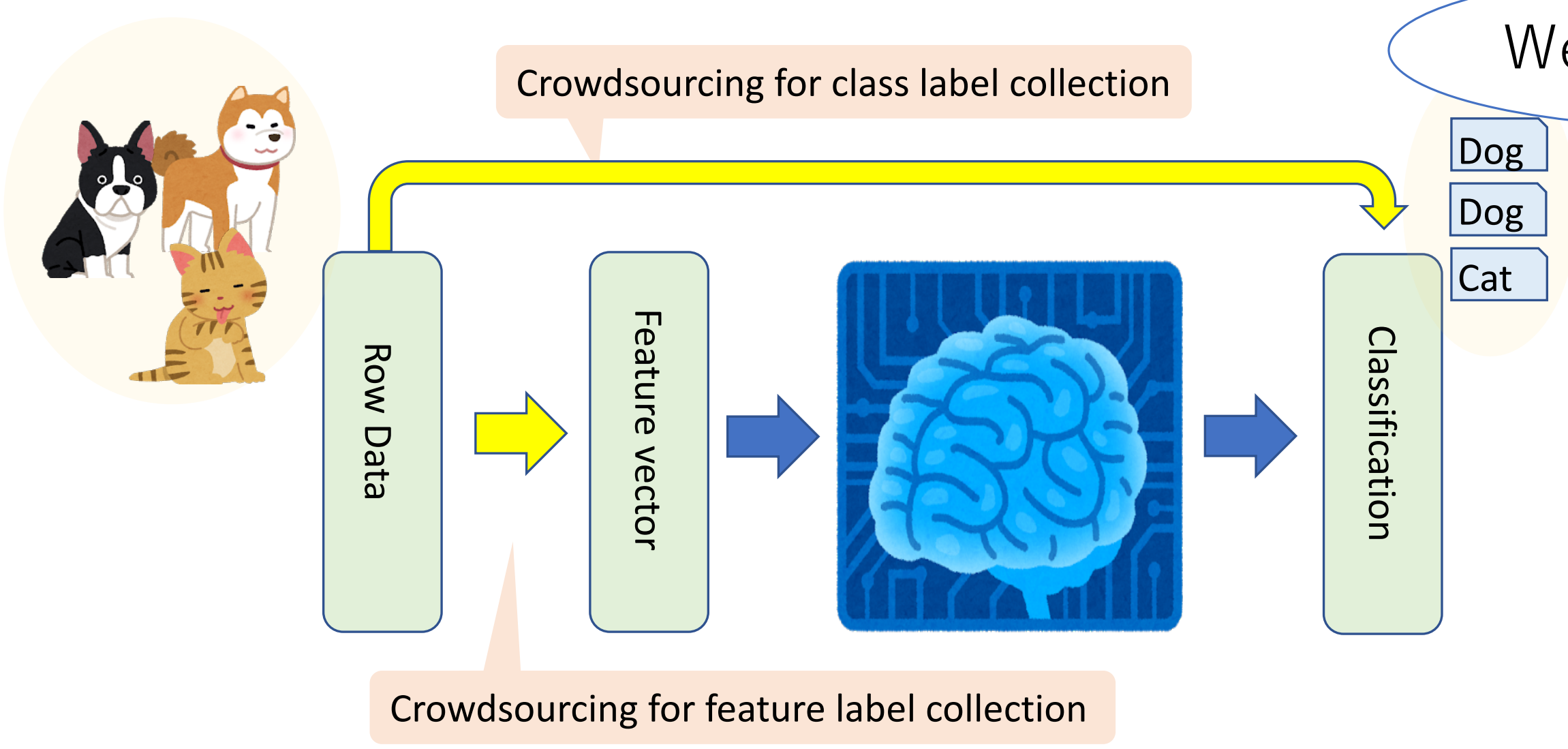

Feature values extraction with crowds
x^i_{11} x^i_{12} ⋯ x^i_{1J}
x^i_{21} ⋯ x^i_{2J}
x^i_{K1} ⋯ x^i_{KJ}
Crowd generated feature values
α_1 α_2 ⋮ α_J
=
$\tilde{x}^i_1$ $\tilde{x}^i_2$ ⋮ $\tilde{x}^i_K$
Worker convolution layer
Estimates worker ability
Neural network makes predictions using crowd generated feature values
Fake smile

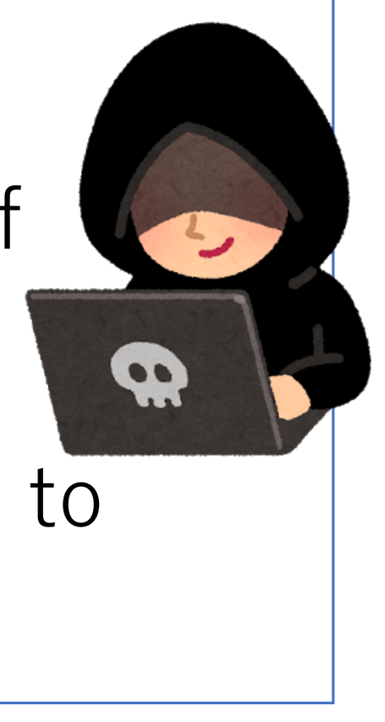
Background | Human-in-the-loop machine learning

Crowdsourcing is a system for outsourcing work to an unspecified number of workers via the Internet

Crowdsourcing is actively used in machine learning, especially for
① class label collection for supervised learning, and
② feature label extraction for data representation

We consider ②


Crowdsourcing for class label collection
Dog
Dog
Cat
Row Data
Feature vector
Classification
Crowdsourcing for feature label collection

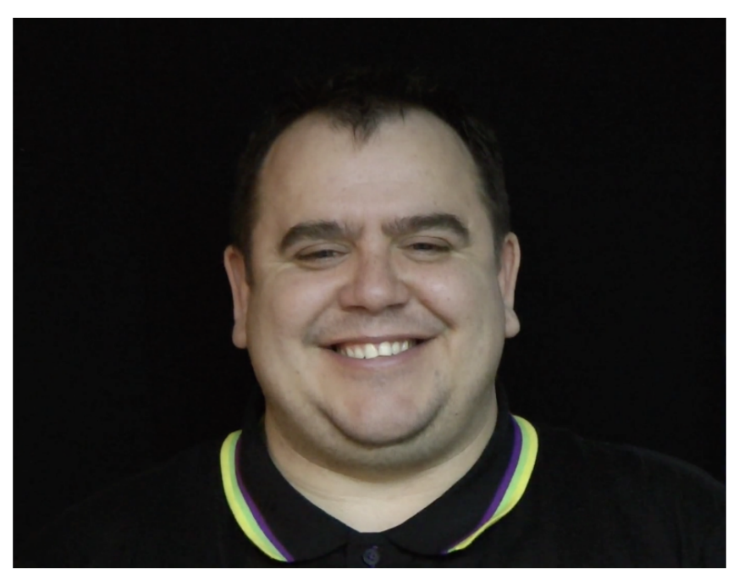


Challenge: Quality control of feature labels
- Quality of the provided feature labels is uneven because of different capability and diligence of crowd workers (Sometimes there are spam or malicious workers)
- We need to integrate feature labels from different workers to improve label quality

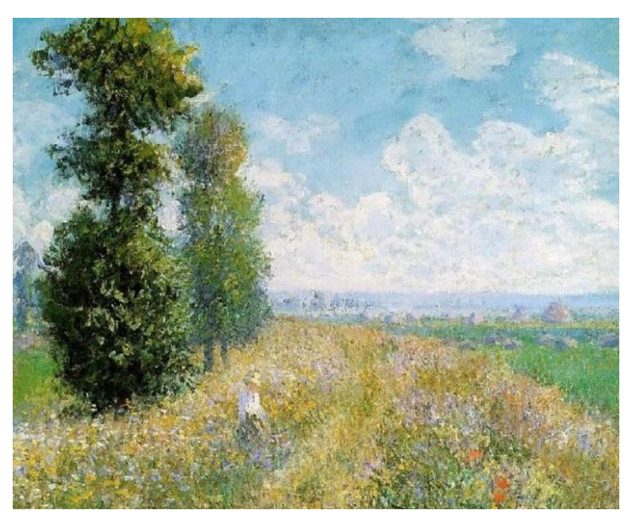
Motivation | Class labels are hard to give by non-experts

Real Smile

⟺



Fake Smile



Sisley's painting

Monet's painting

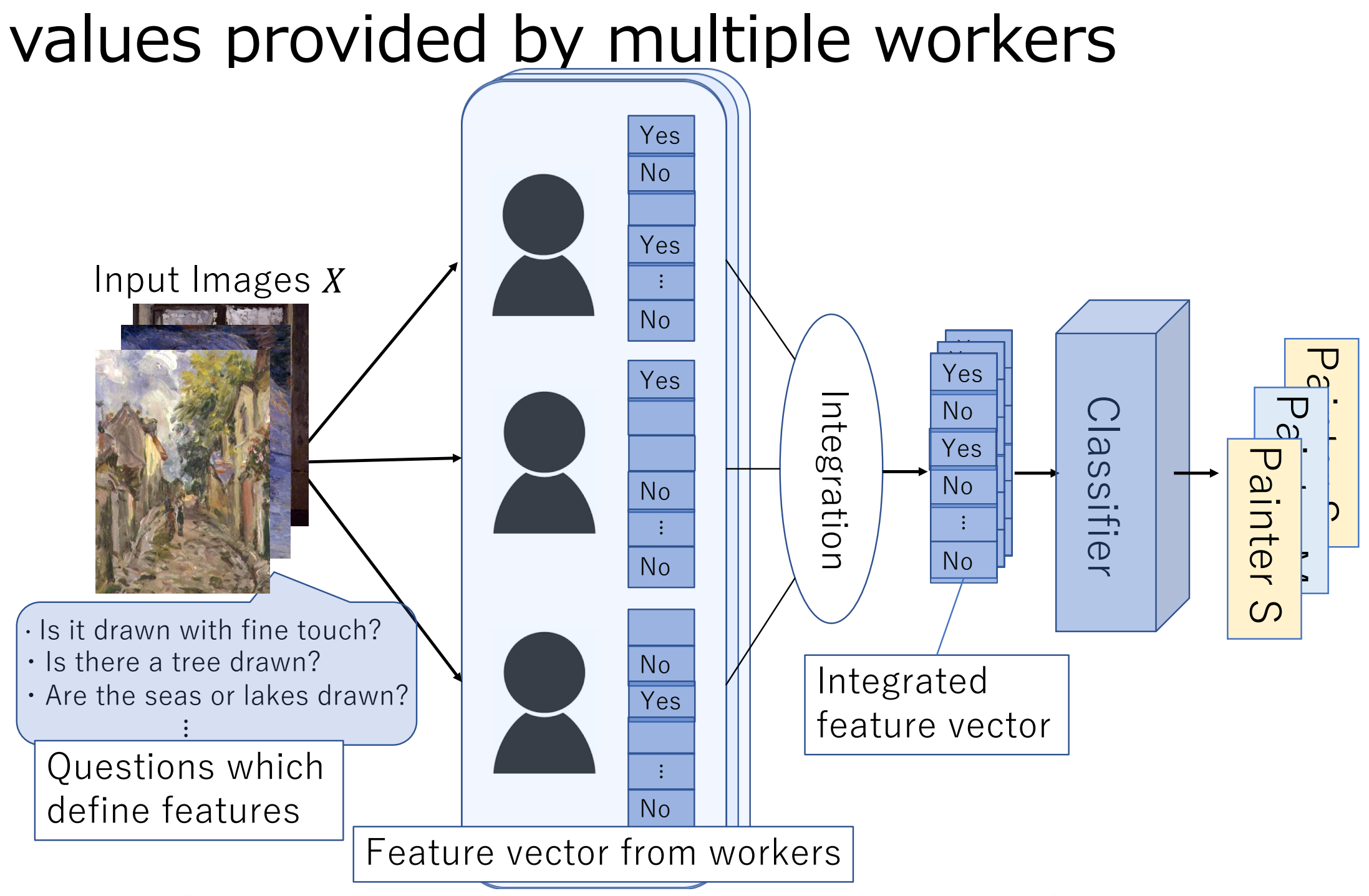
Hard for non-experts to give correct class labels.

Easier for non-experts to give feature labels

Ex) "Are the trees in this painting clearly drawn to the branches?"

Feature extraction using human beings is effective

Problem setting | Binary classification problem based on features values provided by multiple workers


Input Images X
· Is it drawn with fine touch?
· Is there a tree drawn?
· Are the seas or lakes drawn?
⋮
Questions which define features
Yes
No
Yes
⋮
No
Feature vector from workers
Integration
Integrated feature vector
Classifier
Painter S

- Perform binary classification from the feature values given by multiple workers.
- Feature labels are collected in the form of binary questions ("Yes" or "No").
  - 3 workers are assigned to each feature

Proposed method | Neural network simultaneously estimate workers' ability and the classifier

Workers with higher feature extraction ability contribute more to predictions
- Estimate the worker's ability based on the prediction result
- Integrate opinions based on the estimated ability

We propose **worker convolution layer**
- Express worker's ability as weights of one－dimensional filter α.
  - α_j corresponds to the ability of the j-th worker.
- Convolute with filters for each feature and generate integrated labels.

Experiments

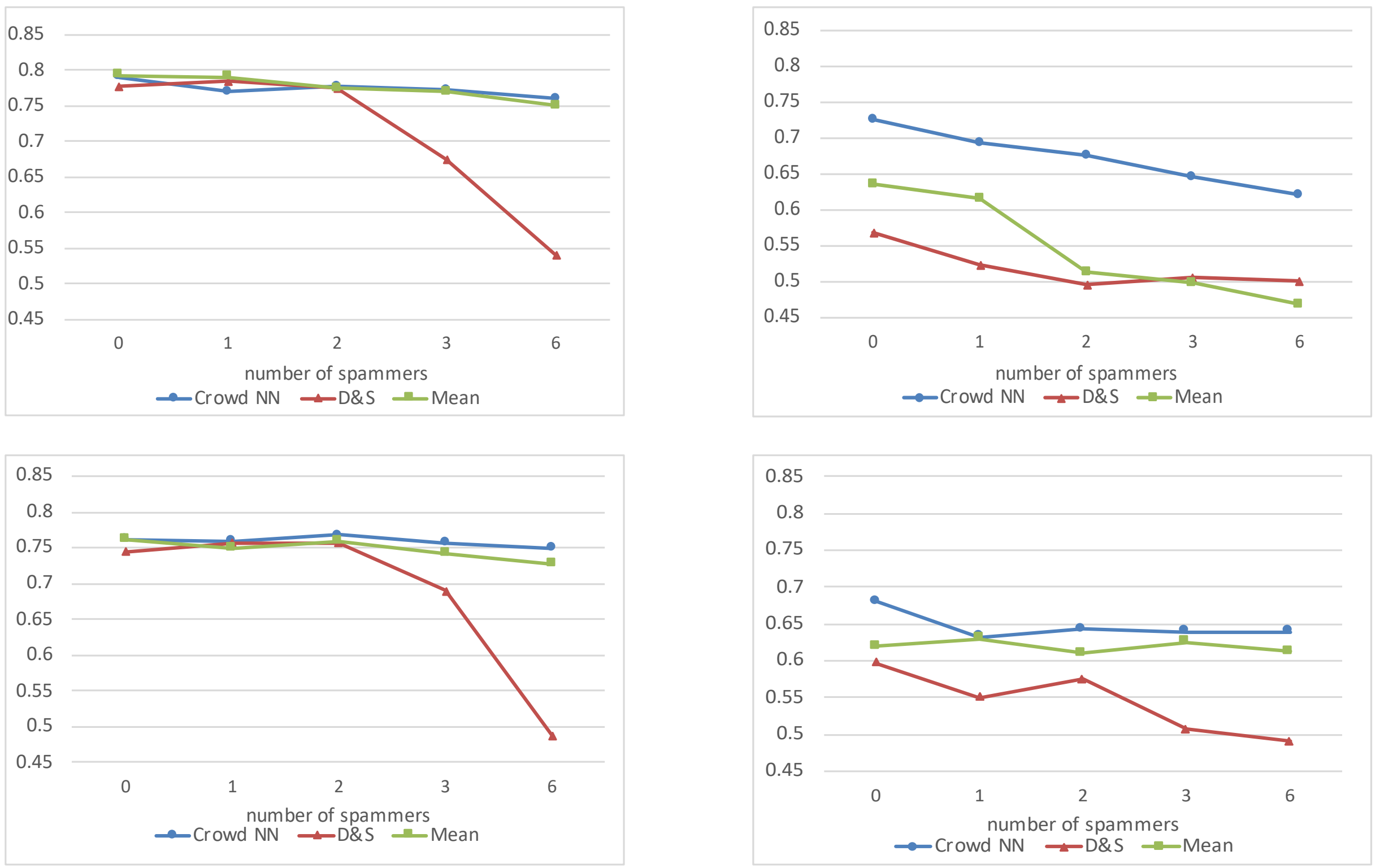
Three experiments with four datasets

① Model performance with original datasets

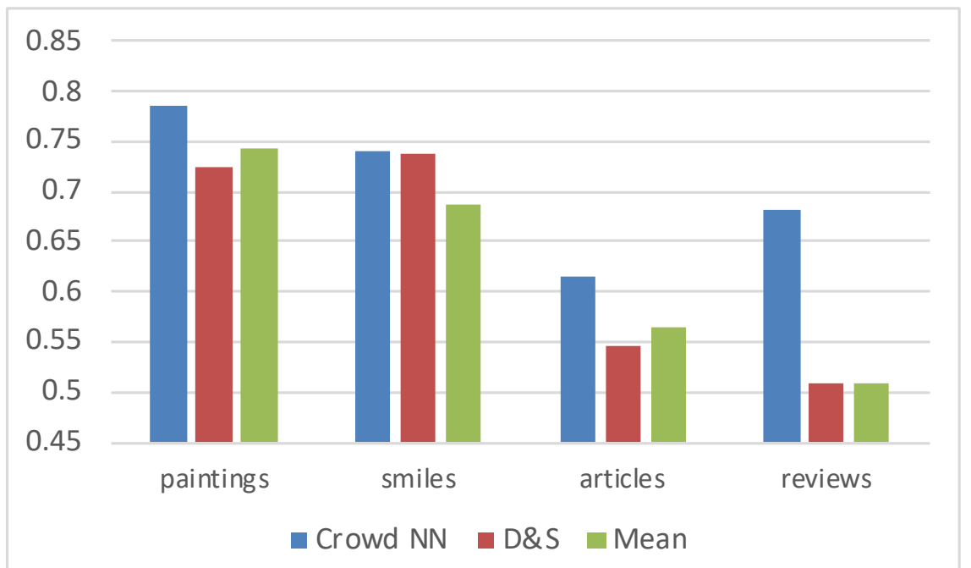
Dataset	Proposed method	Mean	Existing method
Paintings	0.790	<u>0.793</u>	0.778
Fake smiles	<u>0.763</u>	<u>0.763</u>	0.745
Fake reviews	<u>0.680</u>	0.620	0.598
Top news	<u>0.725</u>	0.635	0.568

②Experiments with simulated spam workers

Proposed method is robust against spam workers



③ Robustness against malicious workers



- Experiments with simulated malicious(giving reversed feature labels) workers.
- Proposed method exploits malicious workers to improve predictions.