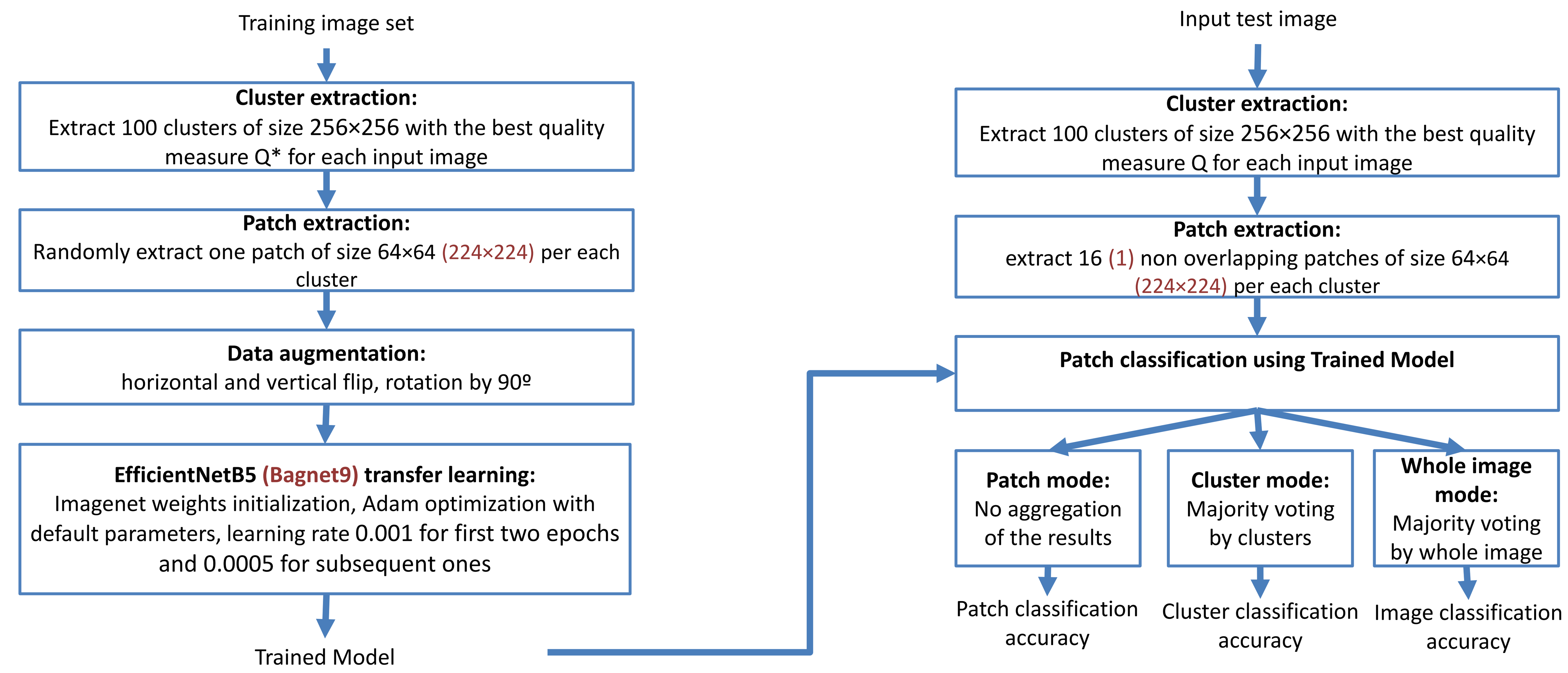
Source camera identification using neural networks

A Y Denisova

Problem statement

Research goal:
Source camera identification is a forensic problem used for image authentication. The identification goal is to determine the camera model by digital image. At present, the most prosperous approach to source camera identification applies neural networks to classify camera models. The aim of the current research is to provide verification and modification of the source camera identification method based on the EfficientNetB5 neural network proposed by Hadwiger and Riess.
The original method is very simple in implementation and it is reported to be very efficient in camera model classification. However, the current research proved that the original method's performance is overestimated. Therefore, a modification of the original method is proposed using the BagNet9 network. The experimental results with Forcheim Image Dataset show that modified method gives significantly better camera identification accuracy than the original method. Thus, BagNet9 is more effective in terms of camera identification than EfficientNetB5.

Original (Modified) method:



*Quality measure for cluster: $Q=\frac{1}{3}\sum_{c\in\{R,G,B\}}\alpha\beta\left(\mu_c-\mu_c^2\right)+(1-\alpha)\left(1-e^{\gamma\sigma_c}\right)$, μ_C – mean brightness in channel C; σ_C – standart deviation of brightness in channel C; $\alpha=0.7, \beta=4, \gamma=\ln(0.01)$

Dataset:

Characteristics	Forchiem Image Dataset
Device number	27
Camera model number	25
Scene number per camera	143

In the current research 25 non repetitive cameras were used.

Scenes from 1 to 97 are used as training. Scenes from 98 to 116 were applied as validation data. Scenes from 117 to 143 are testing data.

Fig. 1. The example of scene 1 for all devices in Forchiem Image Dataset

Experimental results

Backbone model selection:

In first experiment, to substitute EfficientNetB5 in the original method 230 network architectures were considered. The network architectures used are obtained from Tensorflow 2.0 and TF2CV packages. The networks were tested for three classes to reduce the computational time. The clusters of 128×128 were applied. The classification accuracy is obtained only at the patches level. The patch classification accuracy is reported in Table 1 for the top 15 neural networks and EfficientNetB5.

Table 1: camera source identification accuracy for 3 classes at the patch level

Network Architechture	Epoch number for training	Accuracy		
		Training	Validation	Testing
BagNet9	22	0.9791	0.9828	**0.9550**
BagNet33	19	0.9943	0.9852	**0.9465**
BagNet	13	0.9804	0.9804	**0.9365**
SePreResNetbc26b	23	0.9335	0.9158	**0.9158**
PreResNetbc14b	54	0.9655	0.9508	**0.9087**
resnext14_32x4d	23	0.9784	0.9497	**0.9048**
densenet161	10	0.956	0.9431	**0.9003**
resnext14_16x4d	15	0.9716	0.9502	**0.8983**
senet16	9	0.9691	0.9502	**0.8963**
resnext14_32x2d	9	0.9626	0.9575	**0.8936**
densenet169	26	0.9509	0.9519	**0.8895**
darknet_ref	13	0.9395	0.9373	**0.8866**
squeezenet_v1_0	25	0.8834	0.8814	**0.8814**
densenet201	11	0.9493	0.9383	**0.879**
resnetbc14b	15	0.9523	0.9469	**0.8782**
regnetx160	19	0.9531	0.9385	**0.8779**
pyramidnet101_a360	19	0.9372	0.9371	**0.8763**
EfficientNetB5	18	0.825	0.7535	**0.6579**

Modified and original method comparison:

In second experiment, the accuracy of the BagNet9 and EfficientNetB5 using 25 classes and the simple data augmentation are compared. In the original method, patch size is 64×64. In modified method, patches of size 224×224 are used. The patch size 224×224 allows eliminating interpolation of the data before training BagNet9 because this network requires input with size 224×224. The accuracies for patch, cluster, and image modes were computed. But for patch 224×224, the accuracy by patches is equal to the accuracy by clusters because each cluster contains only one patch. The results of the second experiment are presented in Tables 2 and 3.

Table 2: Camera classification Accuracy for EfficientNetB5 and BagNet9 for patch size 64×64 and 25 classes

Network Architechture	Test accuracy		
	Patch	Cluster	Whole image
BagNet9	0.5885	0.7261	**0.8750**
EfficientNetB5	0.2949	0.4584	**0.6752**

Table 3: Camera classification Accuracy for EfficientNetB5 and BagNet9 for patch size 224×224 and 25 classes

Network Architechture	Test accuracy		
	Patch	Cluster	Whole image
BagNet9	0.7779	0.7779	**0.9253**
EfficientNetB5	0.2013	0.2013	**0.5043**

Conclusion

In current research, different neural networks for source camera identification by photographic images are considered. A comparison of 230 neural networks is provided and the network is selected to modify the original method proposed by Hadwiger and Riess. The original method is based on the EfficientNetB5 network. In experiments, it is found that the BagNet9 network has better accuracy than EfficientNetB5. Moreover, the accuracy of EfficientNetB5 was significantly overestimated by Hadwiger and Riess. However, the best accuracy achieved by BagNet9 was 92% which is less than state-of-the-art accuracy. To further enhance the quality of identification additional research is needed.

Acknowledgement: The study was supported by RFBR project № 19-29-09045.

САМАРСКИЙ УНИВЕРСИТЕТ SAMARA UNIVERSITY
SAMARA NATIONAL RESEARCH UNIVERSITY
34, Moskovskoye shosse, Samara, Russia (www.ssau.ru)

ИСОИ РАН
FEDERAL SCIENTIFIC RESEARCH CENTER "CRYSTALLOGRAPHY AND PHOTONICS" OF RUSSIAN ACADEMY OF SCIENCES
151 Molodogvardeyskaya street, Samara, Russia, (https://ipsiras.ru/)