The Usage of Grayscale or Color Images for Facial Expression Recognition with Deep Neural Networks

Dmitry A. Yudin[1], Alexandr V. Dolzhenko[2] and Ekaterina O. Kapustina[2]

[1] Moscow Institute of Physics and Technology (National Research University),
[2] Belgorod State Technological University named after V.G. Shukhov, Belgorod,
*yudin.da@mipt.ru

MIPT
MOSCOW INSTITUTE
OF PHYSICS AND TECHNOLOGY

Cognitive Dynamic Systems Laboratory MIPT

Datasets for facial expression recognition

Database details	CK	JAFFE	FER2013	FE	SoF	AffectNet
Image size	640x490 - 720x480	256x256	48x48	23x29 - 355x536	640x480	129x129 - 4706x4706
Image style	portrait	portrait	cropped face	cropped face	portrait	cropped face
Image type	grayscale, color	grayscale	grayscale	grayscale, color	color	color
Facial expression categories:						
Neutral	324	30	6194	6172	667	75374
Happy	138	31	8989	5693	1042	134915
Sad	56	31	6077	220	237 (sad/ anger/disgust)	25959
Surprise	166	30	4002	364	145 (surprise/ fear)	14590
Fear	50	32	5121	21	0	6878
Disgust	118	29	547	208	0	4303
Anger	90	30	4953	240	0	25382
Contempt	36	0	0	9	0	4250
Total:	978	213	35883	12927	2091	291651

Task Formulation

To solve the task it is necessary to develop various variants of deep neural network architectures and to test them on the available data set with **1-channel (grayscale)** and **3-channel (color)** image representation.

We must determine which image representation is best used for the task of facial expression recognition. Also, we need to select the best architecture that will provide best performance and the highest quality measures of image classification: accuracy, precision and recall

Examples of labeled images with facial expressions from AffectNet Dataset: 0 – Neutral, 1 – Happiness, 2– Sadness, 3 – Surprise, 4 – Fear, 5 – Disgust, 6 – Anger, 7 – Contempt

Dataset preparation

For image augmentation we have used 5 sequential steps:

1. Coarse Dropout – setting rectangular areas within images to zero. We have generated a dropout mask at 2 to 25 percent of image's size. In that mask, 0 to 2 percent of all pixels were dropped (random per image).
2. Affine transformation – image rotation on random degrees from -15 to 15.
3. Flipping of image along vertical axis with 0.9 probability.
4. Addition Gaussian noise to image with standard deviation of the normal distribution from 0 to 15.
5. Cropping away (cut off) random value of pixels on each side of the image from 0 to 10% of the image height/width.

| Errors in "Neutral" category | Errors in "Happiness" category | Errors in "Sadness" category | Errors in "Surprise" category | Errors in "Fear" category | Errors in "Disgust" category | Errors in "Anger" category | Errors in "Contempt" category |

Training and testing samples of used dataset

| Facial expression category	Training sample 1	Training sample 2 (balanced)	Testing sample 1	Testing sample 2 (cleaned)
0 - Neutral	74874	10000	500	490
1 - Happiness	134415	10000	500	451
2 - Sadness	25459	10000	500	473
3 - Surprise	14090	10000	500	453
4 - Fear	6348	10000	500	477
5 - Disgust	3803	10000	500	359
6 - Anger	24882	10000	500	351
7 - Contempt	3749	10000	500	156
Total:	287621	80000	4000	3210

Classification of Emotion Categories using Deep Convolutional Neural Networks

ResNetM architecture inspired from ResNet
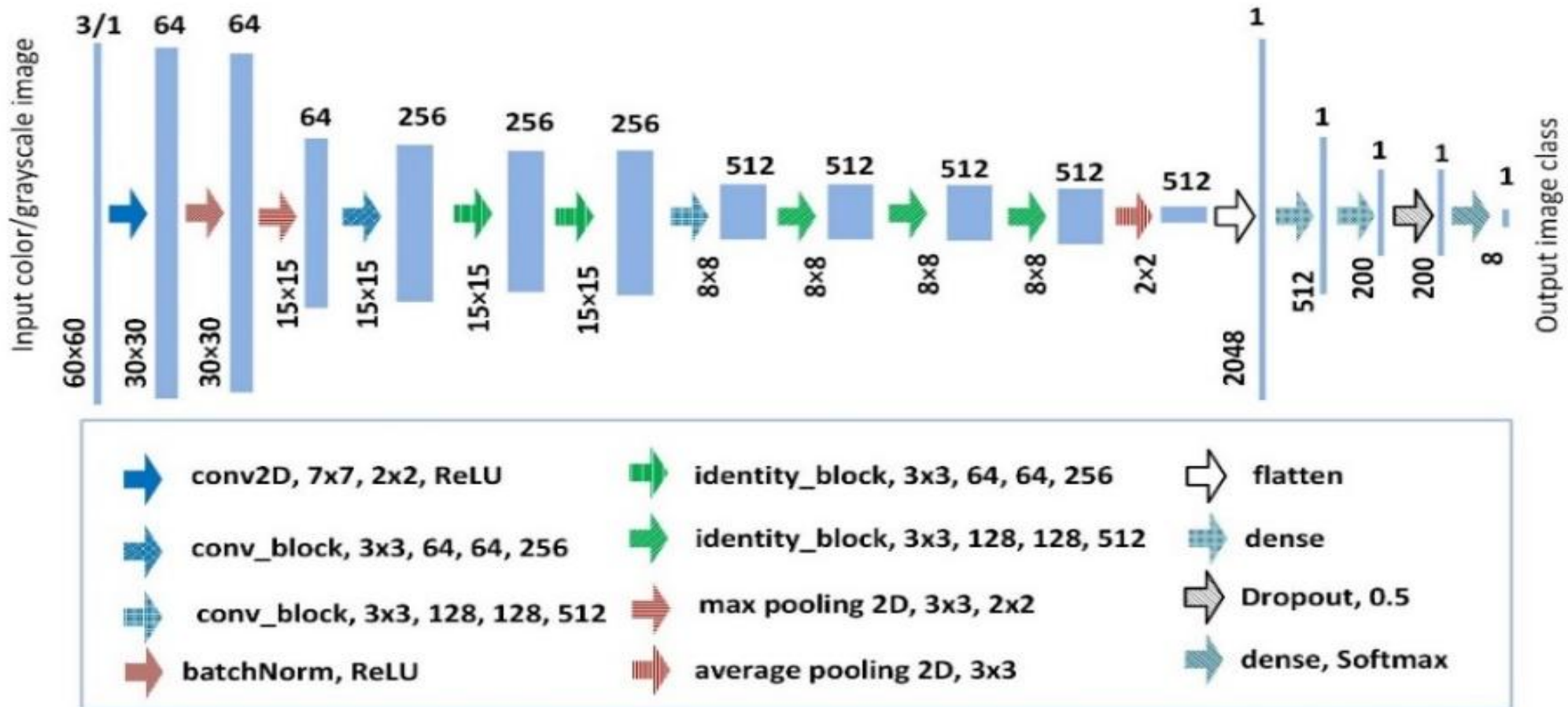
Classification of Emotion Categories using Deep Convolutional Neural Networks

DenseNet architecture is based on DenseNet169 model
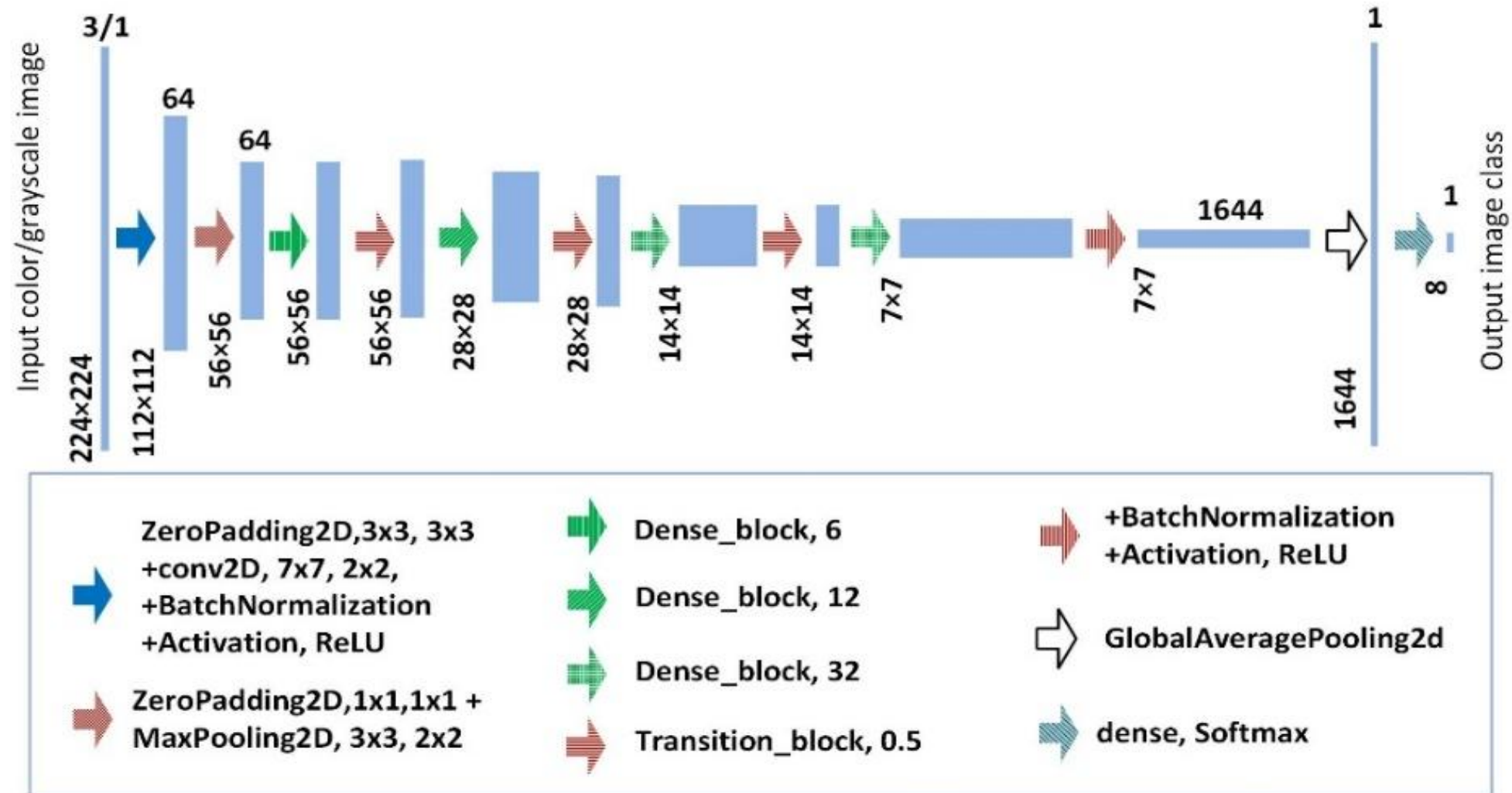
Classification of Emotion Categories using Deep Convolutional Neural Networks

Xception architecture with changed input tensor to 120x120x3 for color images and 120x120x1 for grayscale images
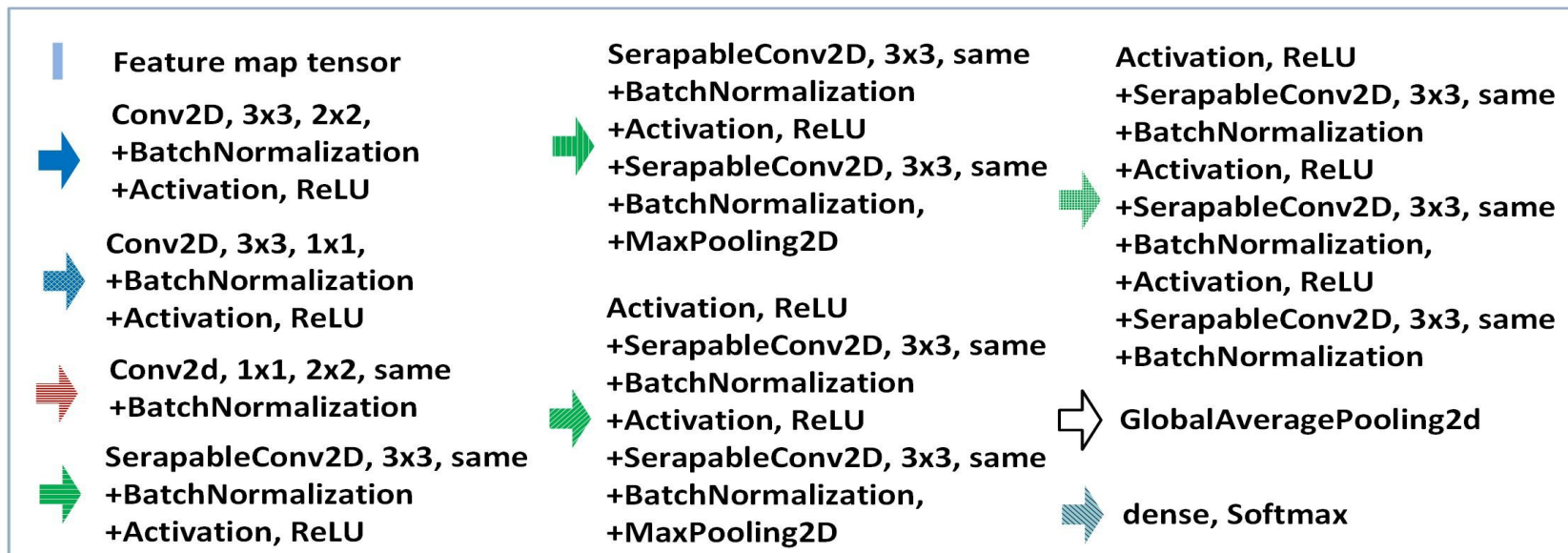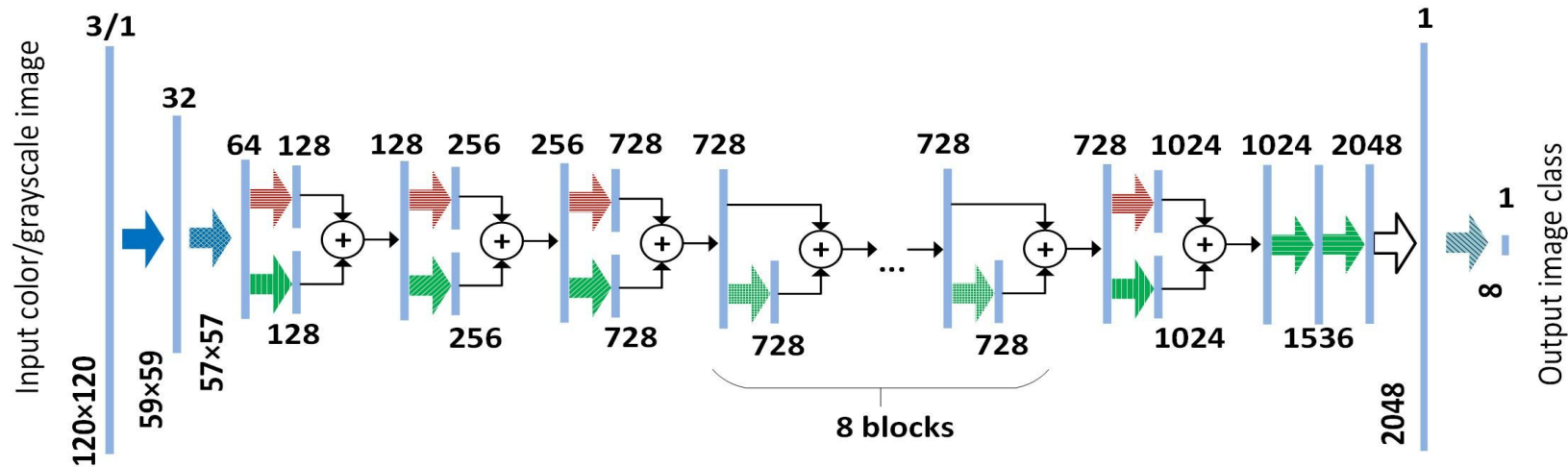
Training of deep neural networks with ResNetM, DenseNet and Xception architectures

ResNetM Training · DenseNet Training · Xception Training
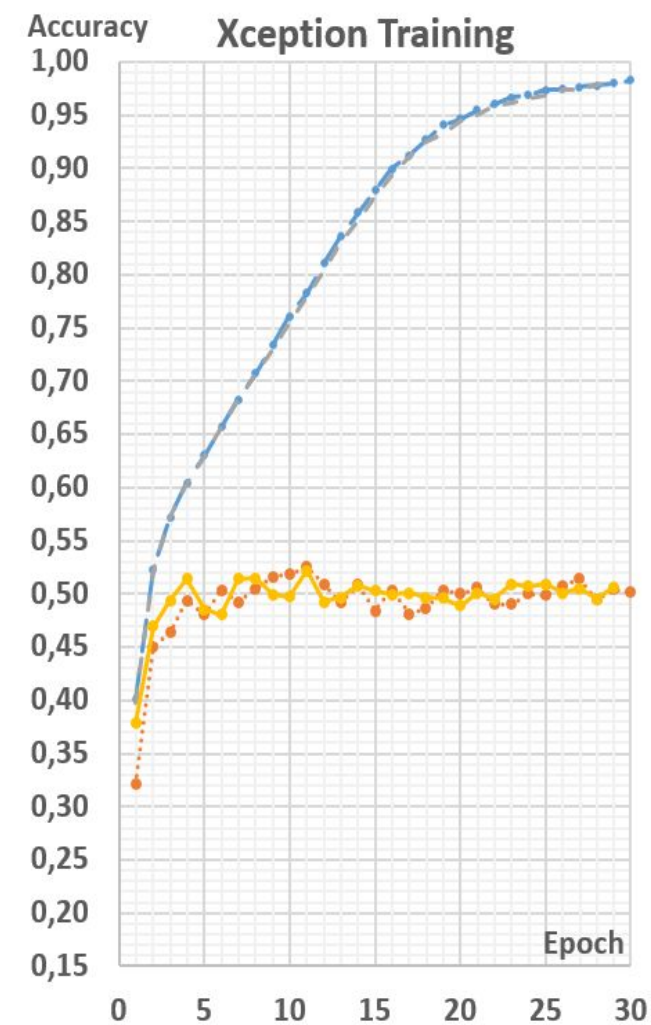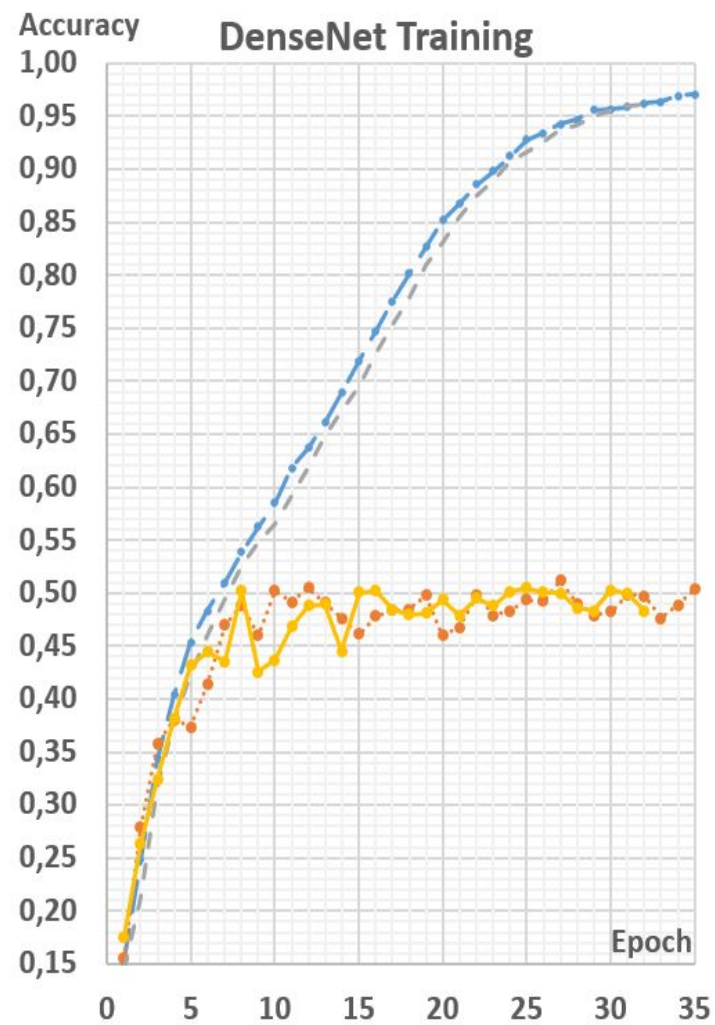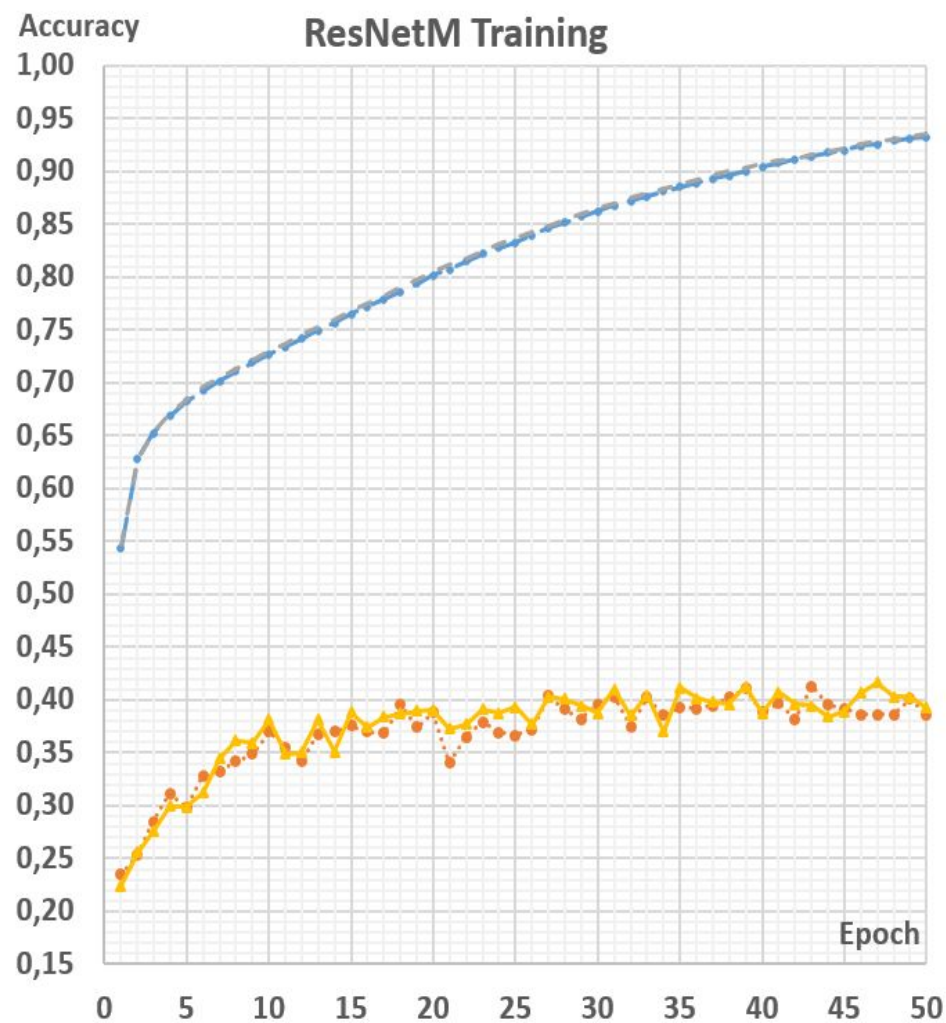
Accuracy on grayscale training sample · Accuracy on grayscale testing sample · Accuracy on color training sample · Accuracy on color testing sample

Quality of facial expression recognition on AffectNet Dataset

Metric	ResNetM		DenseNet		Xception	
	Color	Grayscale	Color	Grayscale	Color	Grayscale
Accuracy on train sample	0.9283	0.9139	0.9168	0.9428	0.9686	**0.9765**
Accuracy on test sample 2	0.4844	0.4781	0.5520	0.5427	0.5654	**0.5748**
Top-2 acc. on test sample 2	0.6748	0.6766	0.7467	0.7371	0.7355	**0.7670**
Classif. time per image, s	**0,0042**	0.0047	0.0305	0.0299	0.0120	0.0123
Weights number	2613392	**2607120**	12656200	12649928	20877872	20877296
Size of model, Mb	10654104	**10629016**	51933264	51908176	83826280	83823848

Quality of facial expression recognition on AffectNet Dataset

Metric	ResNetM		DenseNet		Xception	
	Color	Grayscale	Color	Grayscale	Color	Grayscale
Neutral (0): Precision on test sample	0.375	0.4083	0.4838	0.5644	0.5422	0.5223
Neutral (0): Recall on test sample	0.6061	0.5000	0.5490	0.3755	0.4592	0.5735
Happiness (1):Precision on test sample	0.5214	0.5325	0.7363	0.7701	0.7363	0.7973
Happiness (1):Recall on test sample	0.9468	0.9268	0.7428	0.7428	0.8049	0.7849
Sadness (2):Precision on test sample	0.5184	0.4103	0.6070	0.5617	0.5221	0.6099
Sadness (2):Recall on test sample	0.4165	0.5370	0.4735	0.4715	0.6490	0.4693
Surprise (3):Precision on test sample	0.4810	0.4708	0.4977	0.5177	0.5455	0.5000
Surprise (3):Recall on test sample	0.3907	0.3377	0.4966	0.5475	0.5033	0.6137
Fear (4):Precision on test sample	0.5880	0.5951	0.5867	0.5665	0.6181	0.5864
Fear (4):Recall on test sample	0.3501	0.3542	0.5744	0.5898	0.5597	0.5765
Disgust (5):Precision on test sample	0.6510	0.5679	0.5287	0.5600	0.5912	0.6655
Disgust (5):Recall on test sample	0.2702	0.3259	0.6156	0.5070	0.5599	0.5097
Anger (6):Recall on test sample	0.4645	0.4550	0.5552	0.4456	0.4941	0.4802
Anger (6):Recall on test sample	0.5413	0.4900	0.4587	0.5954	0.4758	0.5869
Contempt (7):Precision on test sample	0.3333	**0.5384**	0.3103	0.2827	0.3065	0.3407
Contempt (7):Recall on test sample	0.0192	0.0448	0.4038	**0.5128**	0.3654	0.2949

Contacts

Dmitry Yudin,
Senior researcher, MIPT

Yudin.da@mipt.ru

Cognitive Dynamic Systems Laboratory MIPT

- Lab website in russian:
  https://mipt.ru/science/labs/cognitive-dynamic-systems/
- Lab website in english:
  https://mipt.ru/english/research/labs/cds