НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ

Laboratory of Algorithms and Technologies for Network Analysis

PATTERN RECOGNITION AND IMAGE UNDERSTANDING

A.V. Savchenko

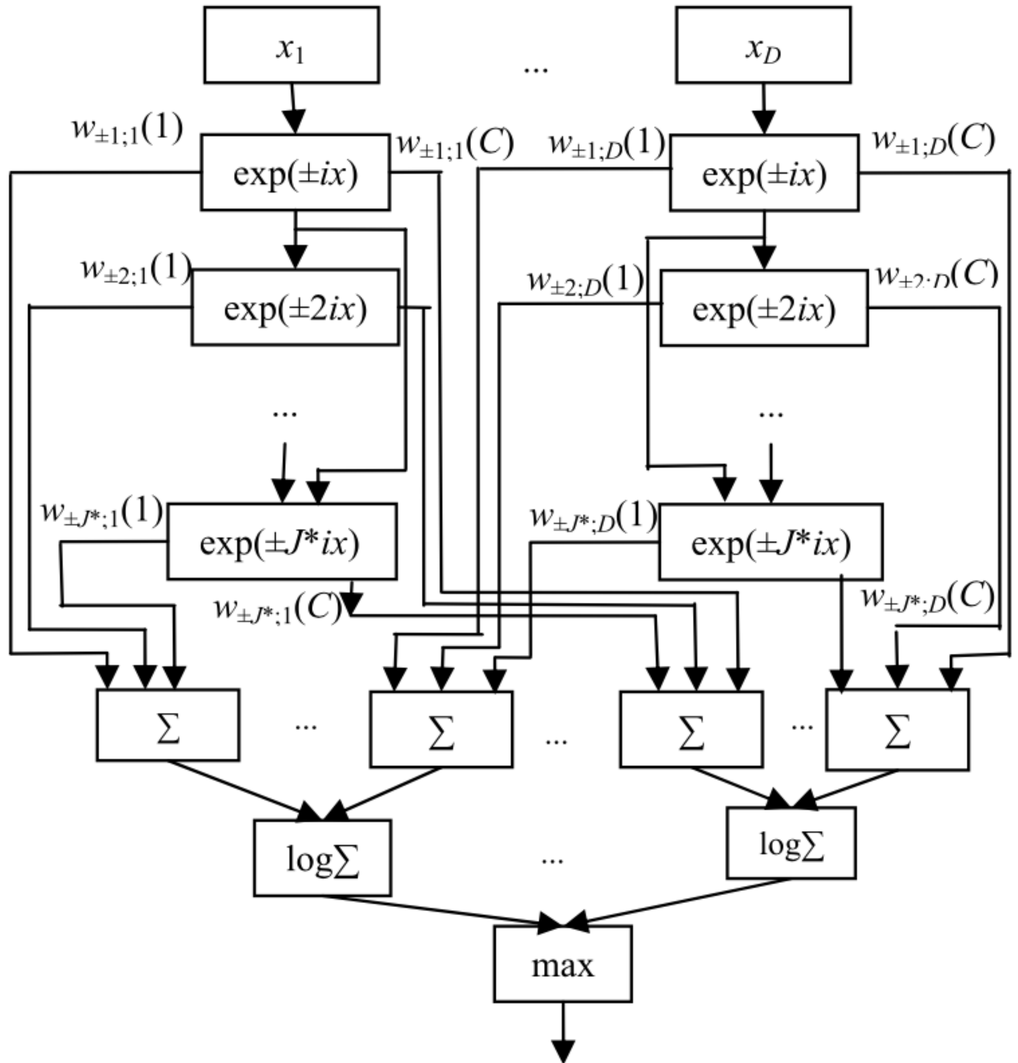
Nizhny Novgorod, 2020

RECENT RESULTS

PROBABILISTIC NEURAL NETWORK (PNN) WITH EXPONENTIAL ACTIVATIONS

Orthogonal density expansions to speed up classification:

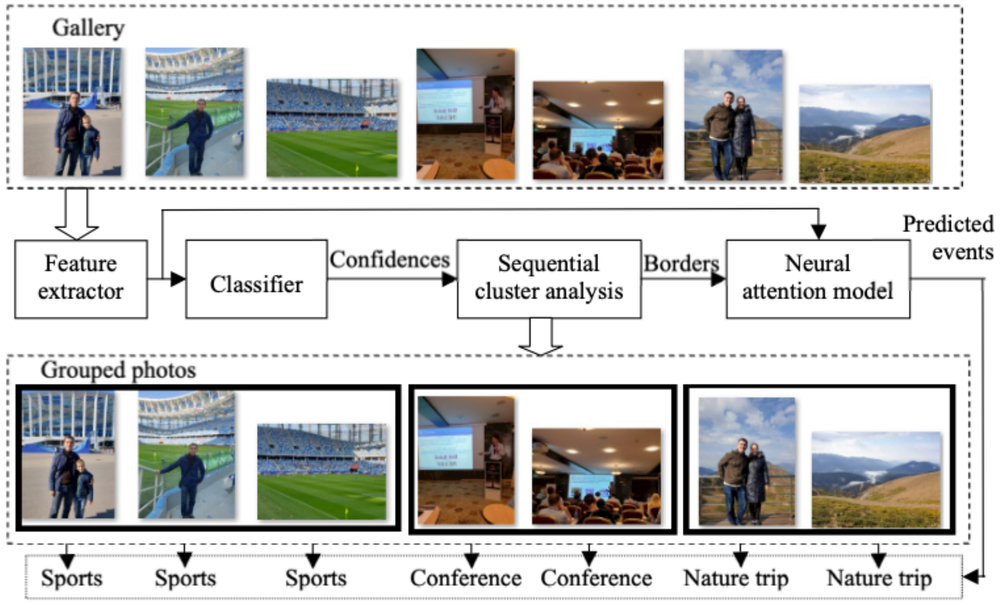
- Converges to Bayesian solution
- Very high training speed
- Classification is $(N/C)^{2/3}$–times faster than instance-based learning (N is the number of training instances, C is the number of classes).

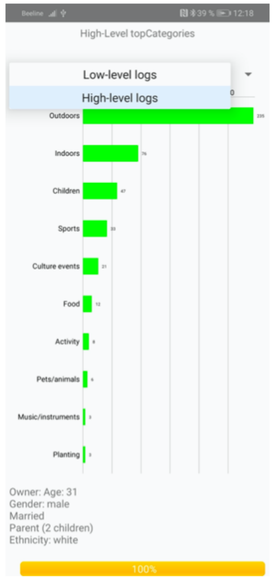


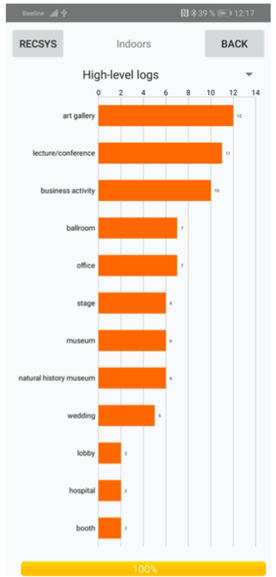
Savchenko A.V., IEEE Transactions of Neural Networks and Learning Systems (IF 8.793), 2020

ANALYSIS OF PHOTOS ON MOBILE DEVICE

- Prior art requires the input image set to represent one album. Our invention automatically assigns album's borders from a gallery by computing similarities between confidences of each pair of consecutive photos in a gallery
- LSTM and estimation of importance scores is replaced by attention network: invention is faster and does not need training dataset with given importance scores of each photo



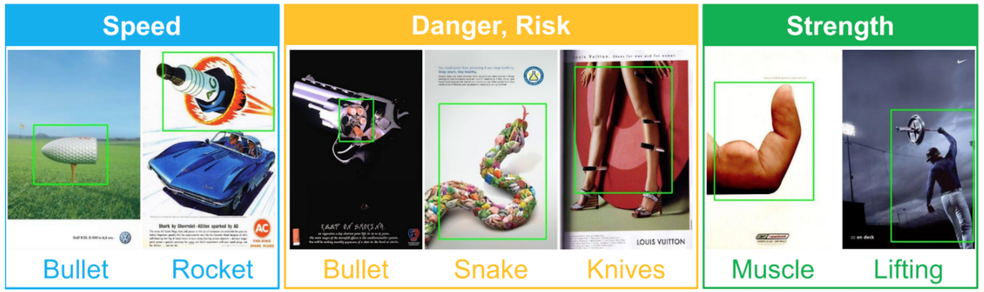


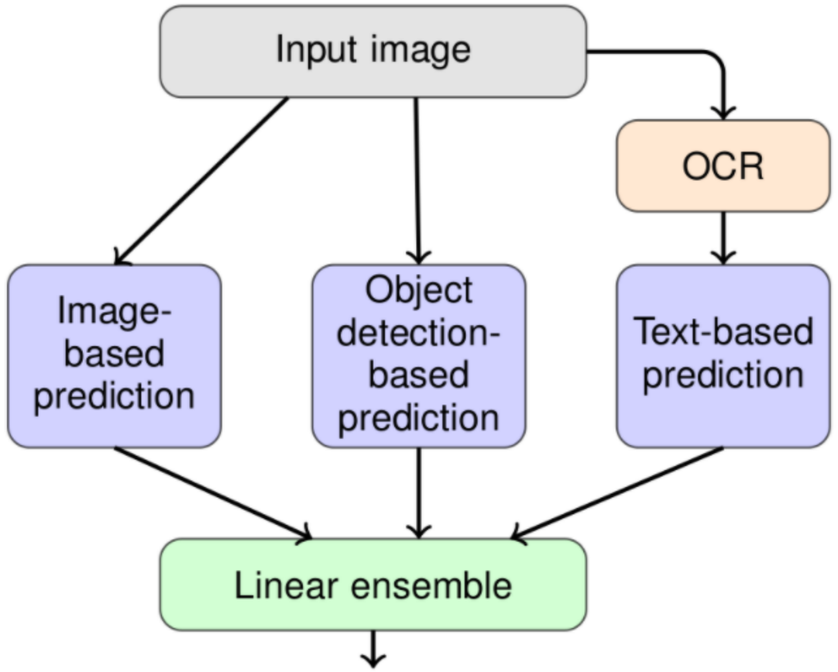


Savchenko A.V., Proc. of IJCNN (Core A), 2020
Submitted for patent application by Samsung

SYMBOLISM PREDICTION IN ADS

- Ads are all about symbols, hard to parse even for a human sometimes
- We developed a multimodal pipeline that combines NLP (reading text on ads with OCR) and computer vision and outperforms state of the art



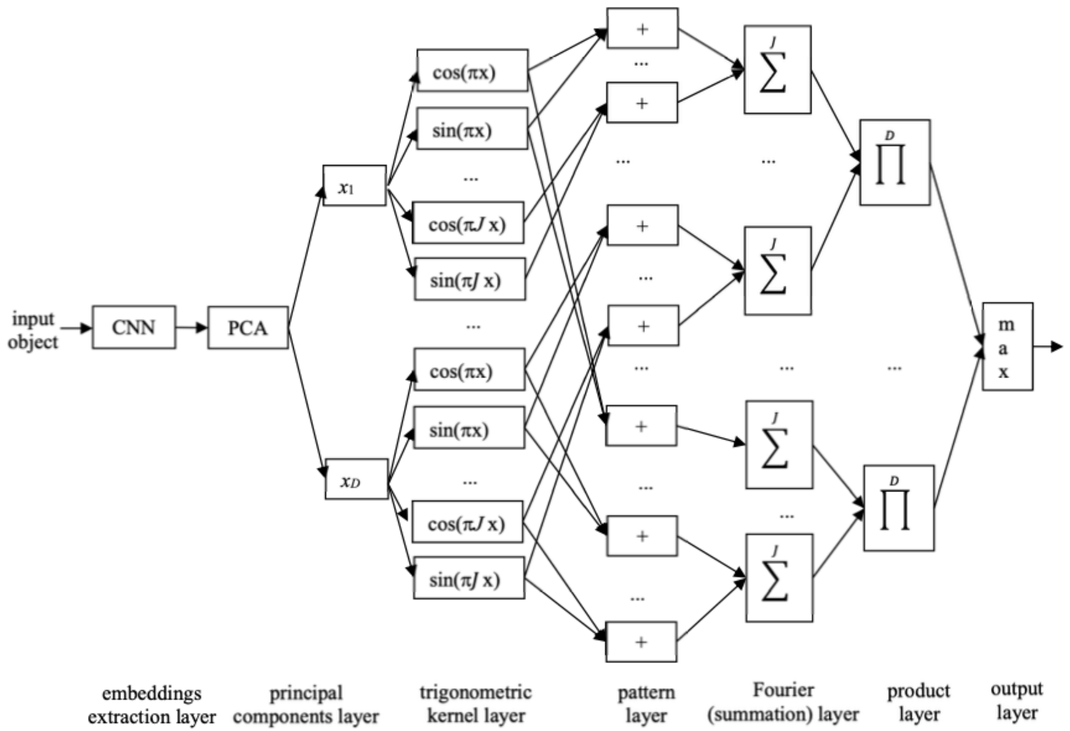


Savchenko A.V. et al, Proc. of COLING (Core A), 2020 – joint work with
- PDMI RAS, St. Petersburgh (A. Alexeev, S. Nikolenko),
- Kazan Federal University (E. Tutubalina),
- Samara University (E. Myasnikov)

CURRENT WORK

SEQUENTIAL ANALYSIS IN FOURIER PNN

Fourier PNN



Submitted to Neural Networks journal

Transform L_2-normed observed point $\mathbf{x}$ into principal component scores scaled by Eq. 6.

$LogF[c] = \log \frac{N(c)}{N}$, $Cand[c] = 1, c = \{1, ..., C\}$
for $l \in \{1, ..., L\}$ **do**
  for $d \in \{d_{l-1} + 1, ..., d_l\}$ **do**
    $Cos[1] = \cos(\pi \overline{x}^{(d)}), Sin[1] = \sin(\pi \overline{x}^{(d)})$
    for $j \in \{1, ..., J-1\}$ **do**
      $Cos[j+1] = Cos[j]Cos[1] - Sin[j]Sin[1]$
      $Sin[j+1] = Cos[j]Sin[1] + Sin[j]Cos[1]$
    end for
    for $c \in \{1, ..., C\}$ **do**
      if $Cand[c] == 1$ **then**
        $f = 0.5$
        for $j \in \{1, ..., J\}$ **do**
          $f = f + \check{a}_{j;d}(c)Cos[j] + \check{b}_{j;d}(c)Sin[j]$
        end for
        $LogF[c] = LogF[c] + \log f$
      end if
    end for
  end for
  $LogF^* = \max_{c \in \{1,...,C | Cand[c]=1\}} LogF[c]$
  for $c \in \{1, ..., C\}$ **do**
    if $LogF[c] < (LogF^* + d_l \cdot \log \delta)$ **then**
      $Cand[c] = 0$
    end if
  end for
  if $(Cand[1] + ... + Cand[C]) == 1$ **then**
    break
  end if
end for
Return the class label $c^* = \operatorname{argmax}_{c \in \{1,...,C | Cand[c]=1\}} LogF[c]$

Best runtime complexity (m=D/L):
$O(mN^{1/3}C^{2/3})$

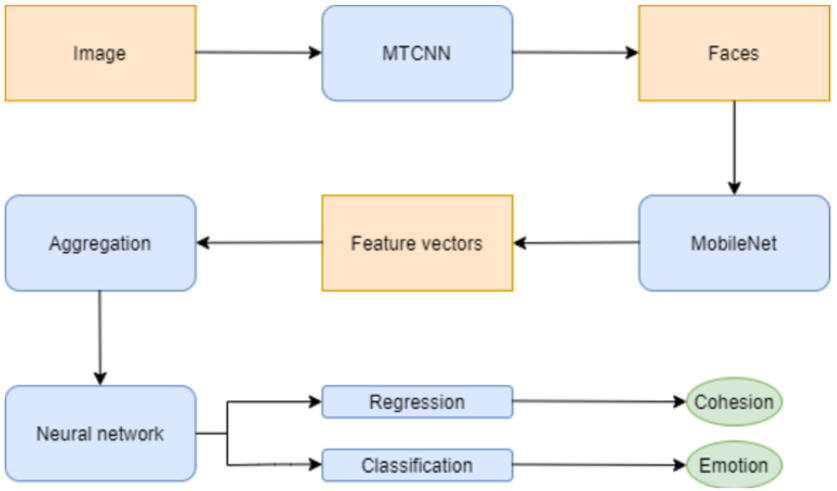
Average runtime complexity:

$$O\left(N^{1/3}C^{2/3}m\sum_{l=1}^{L}\gamma^{l-1}\right)$$

EMOTION RECOGNITION IN VIDEO

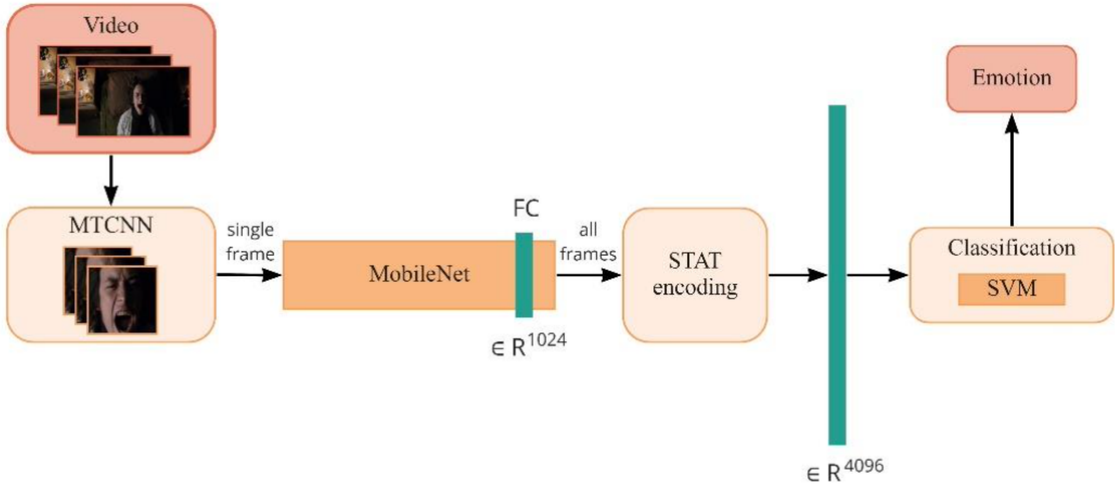
Efficient Group-based Cohesion Prediction in Images using Facial Descriptors

(a) Strongly disagree (b) Disagree (c) Agree (d) Strongly agree



Supported by Russian Science Foundation

MobileEmotiFace: Efficient Facial Image Representations in Video-based Emotion Recognition on Mobile Devices



COMPUTER-ASSISTED PRONUNCIATION TOOL

Gain optimization in spectral distortions

$$\rho_{go-COSH}(\hat{G}_{\mathbf{x}}, \hat{G}_r) = \min_{\lambda>0} \rho_{COSH}(\hat{G}_{\mathbf{x}}, \lambda \hat{G}_r),$$

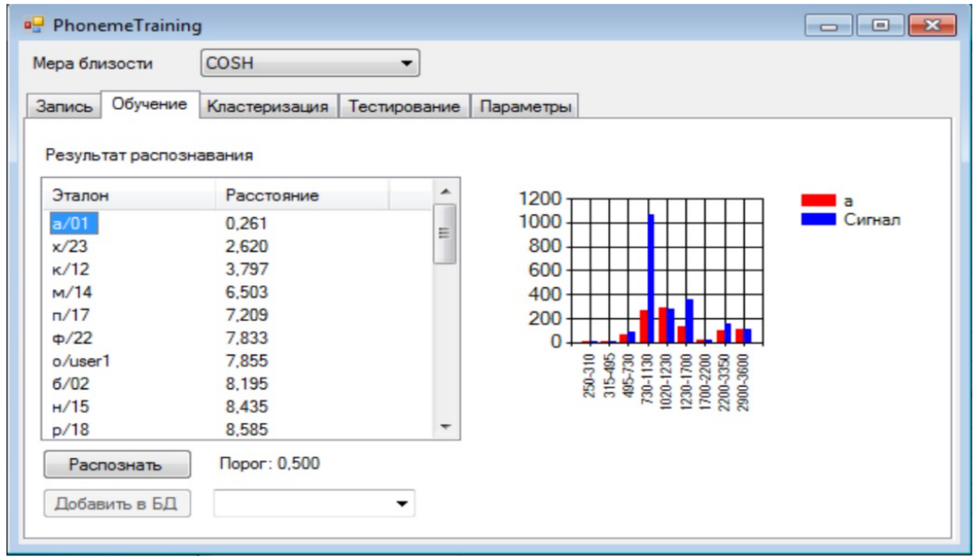
$$\rho_{go-COSH}(\hat{G}_{\mathbf{x}}, \hat{G}_r) = \frac{1}{F}\sqrt{\left(\sum_{f=1}^{F} \frac{\hat{G}_{\mathbf{x}}(f)}{\hat{G}_r(f)}\right)\left(\sum_{f=1}^{F} \frac{\hat{G}_r(f)}{\hat{G}_{\mathbf{x}}(f)}\right)} - 1.$$

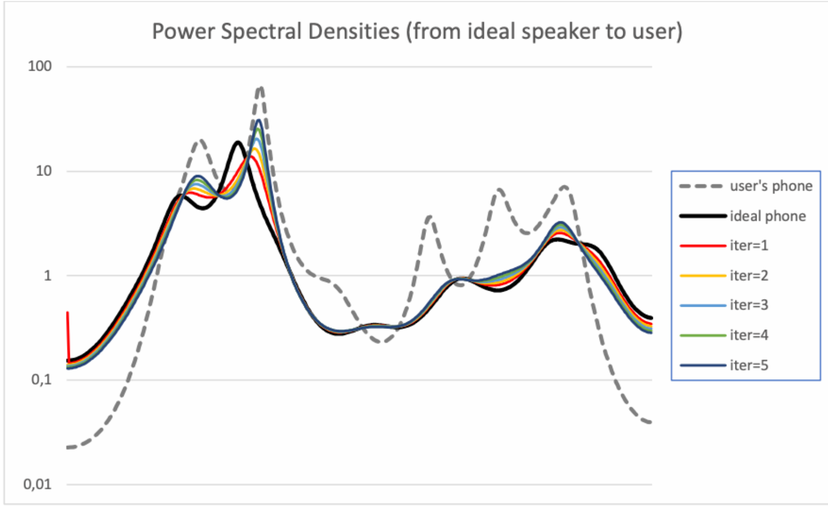
Adaptation of LPC (linear prediction coding) coefficients

$$a^{t+1}(m) = a^t(m) - \gamma \frac{\partial \rho_{go-COSH}(\hat{G}_{\mathbf{a}}, \hat{G}_r)}{\partial a_m}.$$

$$x^*_{r;n} = \begin{cases} x_n, & n < p \\ e_n(\mathbf{x}) - \sum_{m=1}^{p} \tilde{a}^*_r(m) \cdot x^*_{r;n-m}, & n \geq p. \end{cases}$$

Extended version of MOTOR's paper – submitted to Optimization Letters





ONGOING INDUSTRIAL PROJECTS

- 2020 – Mura defect detection (phase I)

Future projects – under discussion

- Mura defect detection (phase II) – collaboration with ViCueSoft (Nizhny Novgorod)
- AutoML for image classification on mobile devices – collaboration with HSE-Moscow (Ilya Makarov)
- Multimodal analysis of images and texts – collaboration with PDMI RAS (Sergey Nikolenko, Elena Tutubalina)
- Pet (cows) recognition in video

НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ