HIGHER SCHOOL OF ECONOMICS
NATIONAL RESEARCH UNIVERSITY

Fast Nearest-Neighbor Classifier based on Sequential Analysis of Principal Components

Anastasiia D. Sokolova, Andrey V. Savchenko
Email: adsokolova96@mail.ru

AIST

Analysis of images, social networks, and texts 2019

Outline

- Image recognition problem
- Proposed recognition method based on sequential analysis of principal components
- Experimental results in face classification
- Concluding comments and future plans

AIST

Problem formulation

The problem is to assign an input set of T feature vectors $x(t) = [x_1(t), \dots, x_D(t)]$ into one of C classes. They are specified by a training set of $R \geq C$ points $x_r = [x_{r;1}, \dots, x_{r;D}], r \in \{1, \dots, R\}$, which class label $c(r) \in \{1, \dots, C\}$ is known. Dimensionality D is rather high

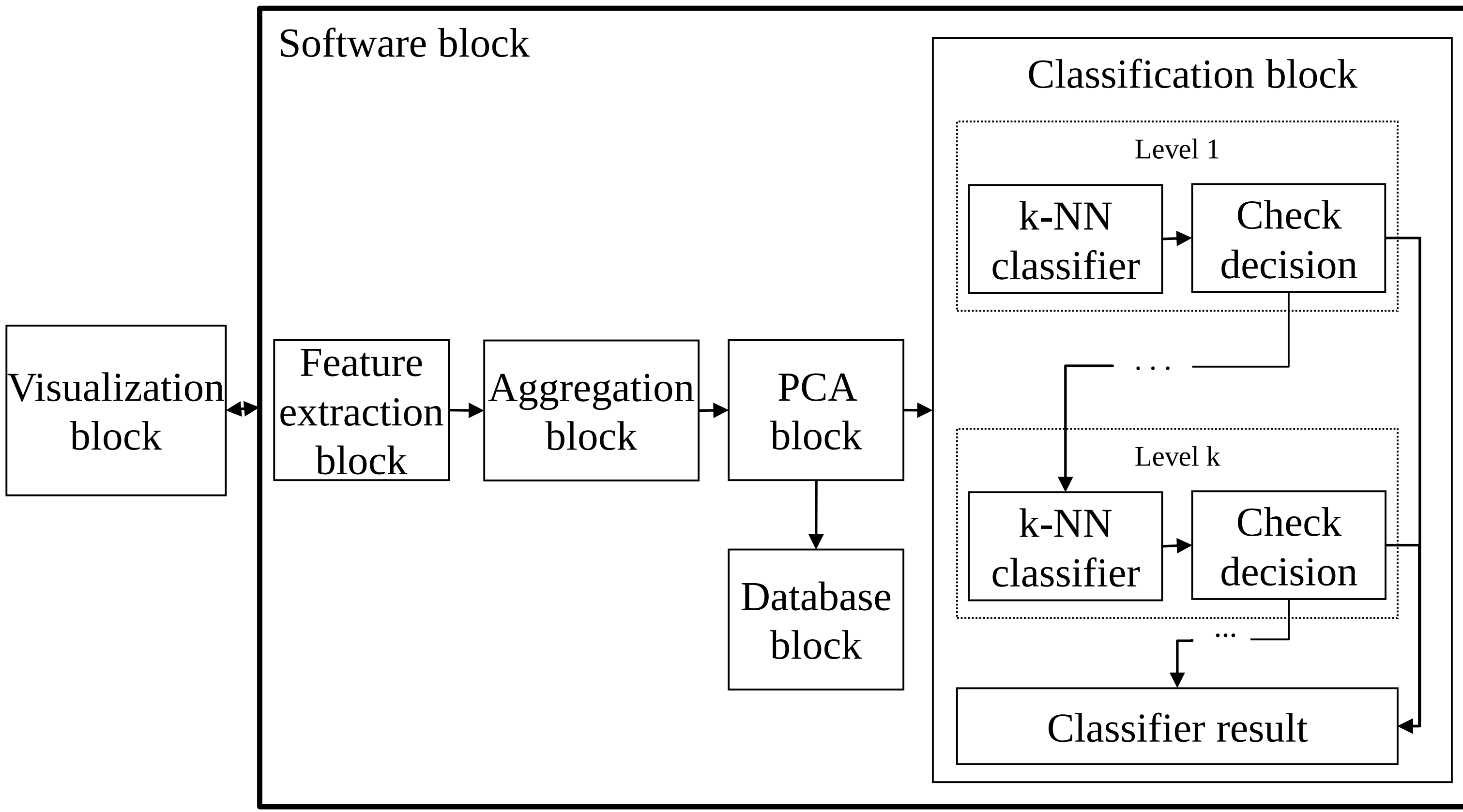
Instance-based learning approach

- k-nearest neighbor (k-NN)

As such feature vectors are high-dimensional, classification speed can be too low for many practical applications

AIST

Proposed algorithm



Proposed approach

Feature extraction is implemented using the deep CNNs trained with an external large dataset.

The distance between feature level to speed-up the matching:

$$\rho\left(\tilde{x}^{(l)}, \tilde{x}_r^{(l)}\right) = \rho\left(\tilde{x}^{(l-1)}, \tilde{x}_r^{(l-1)}\right) + \sum_{d=d_{(l-1)}+1}^{d_l} \rho(\tilde{x}_d, \tilde{x}_{r;d}) \quad (1)$$

Nearest neighbor class:

$$\mathrm{c}_l^* = \underset{c \in C_l}{\operatorname{argmin}}\, \rho_{\mathrm{c}}(\tilde{x}^{(l)}) \quad (2)$$

The set of candidates

$$C_{l+1} = \left\{ c \in C_l \,\middle|\, \frac{\rho_{\mathrm{c}}(\tilde{x}^{(l)})}{\rho_{\mathrm{c}_l^*}(\tilde{x}^{(l)})} \leq \delta \right\} \quad (3)$$

Aggregation techniques

Average features vector:

$$\overline{x} = \frac{1}{T}\sum_{t=1}^{T} x(t) \tag{4}$$

Medoid:

$$x^* = \underset{x(t)}{\operatorname{argmin}} \sum_{t'=1}^{T} \rho(x(t), x(t')) \tag{5}$$

Median:

$$x' = [x'_1, \dots, x'_D] \tag{6}$$

x'_i – the median of i-th component of all vectors $\{x(t), t = 1,2, \dots, T\}$

AIST

Datasets

- LFW (Labeled Faces in the Wild)
  - 1680 people
  - 13000 images
  - 1-10 images
- YTF (YouTube Faces)
  - 1595 people
  - 3425 videos
  - 48-6070 frames

Feature extraction

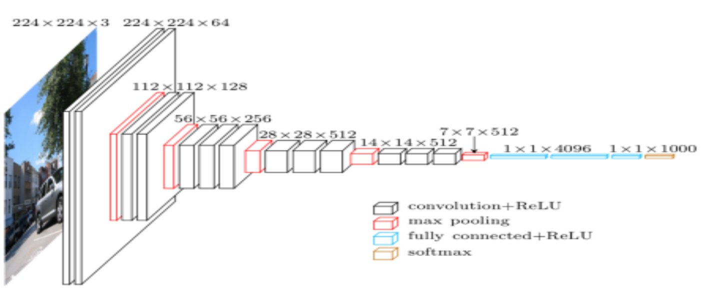
- Lightened CNN (version C)

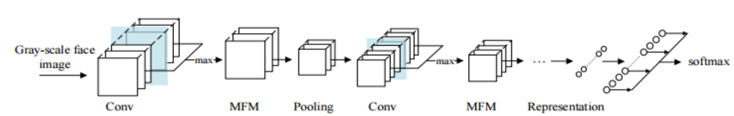
  Out: vector of 256 elements
- FaceNet

  Out: vector of 512 elements
- VggFace2

  Out: vector of 2048 elements
- VggFace

  Out: vector of 4096 elements



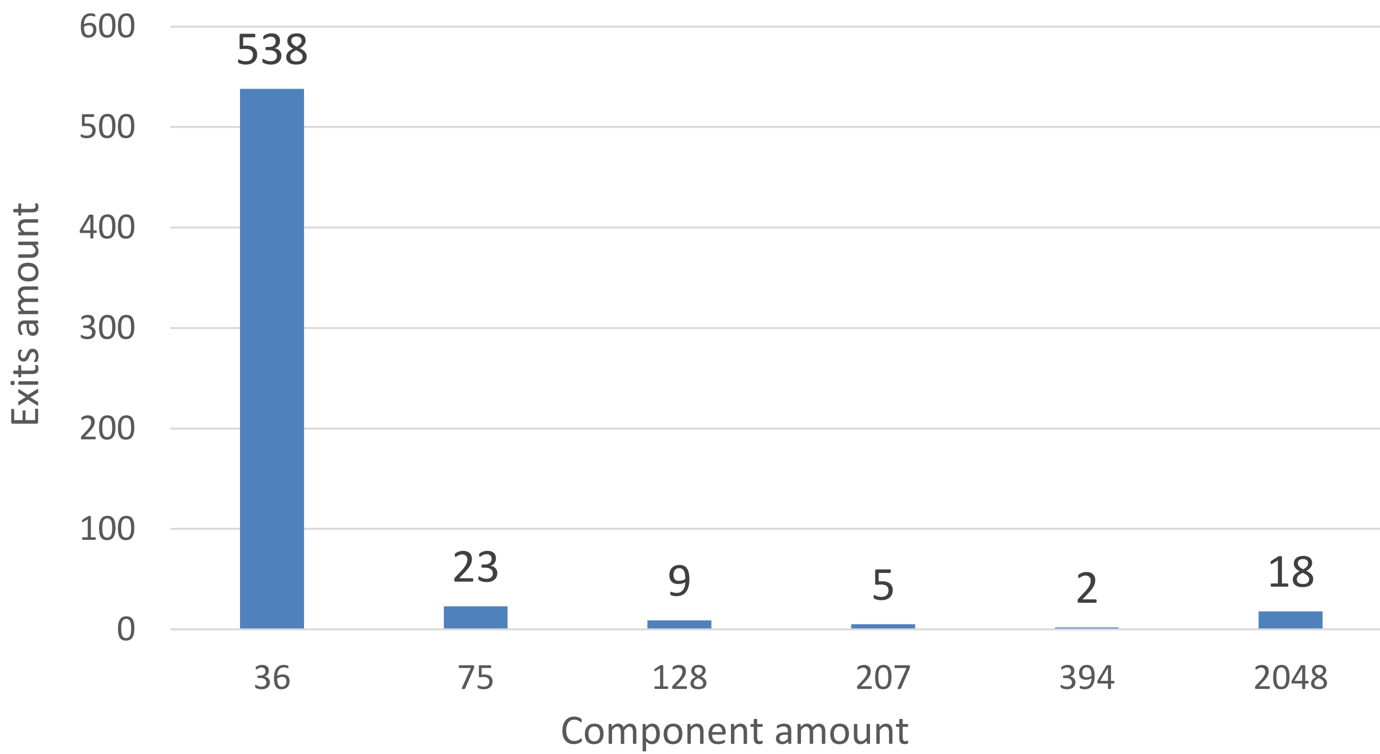


TensorFlow

Face recognition results for the k-NN classifier (LFW)

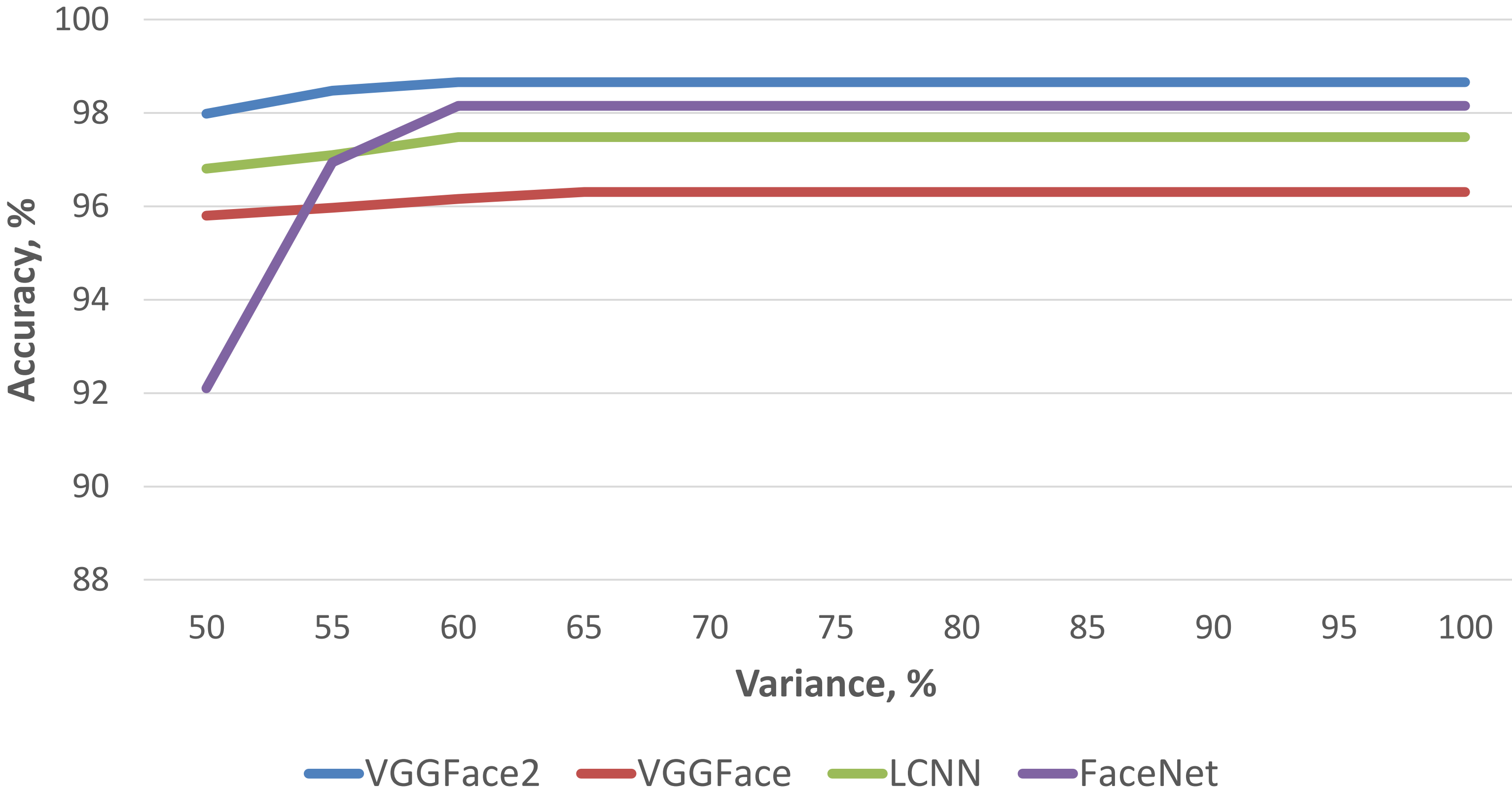
Metric	Classifier	VGGFace	LCNN	VGGFace2	FaceNet
Accuracy (%)	k-NN, all components	96.31	97.48	98.66	98.15
	k-NN, fixed no. of components	94.10	96.21	96.95	97.36
	sequential k-NN, fixed no. of components	95.97	96.97	98.32	98.15
	sequential k-NN, variable no. of components	95.80	96.81	97.98	97.94
Time (ms)	k-NN, all components	50.39	4.88	34.78	9.37
	k-NN, fixed no. of components	2.53	2.50	2.53	2.52
	sequential k-NN, fixed no. of components	2.92	2.54	2.84	2.69
	sequential k-NN, variable no. of components	3.20	2.23	2.81	1.87

Face recognition results for the k-NN classifier (LFW)



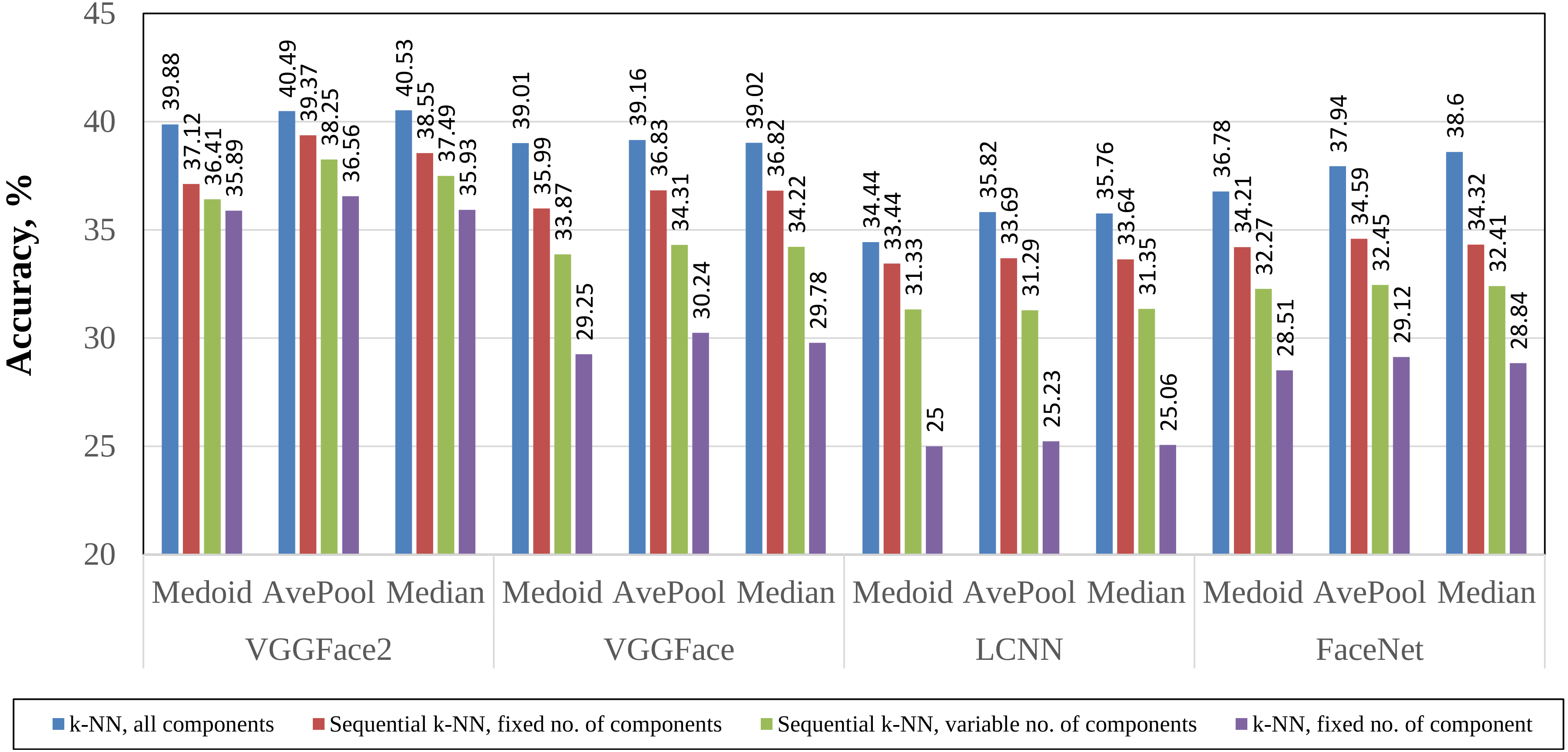
AIST

Face recognition results for the k-NN classifier (LFW)

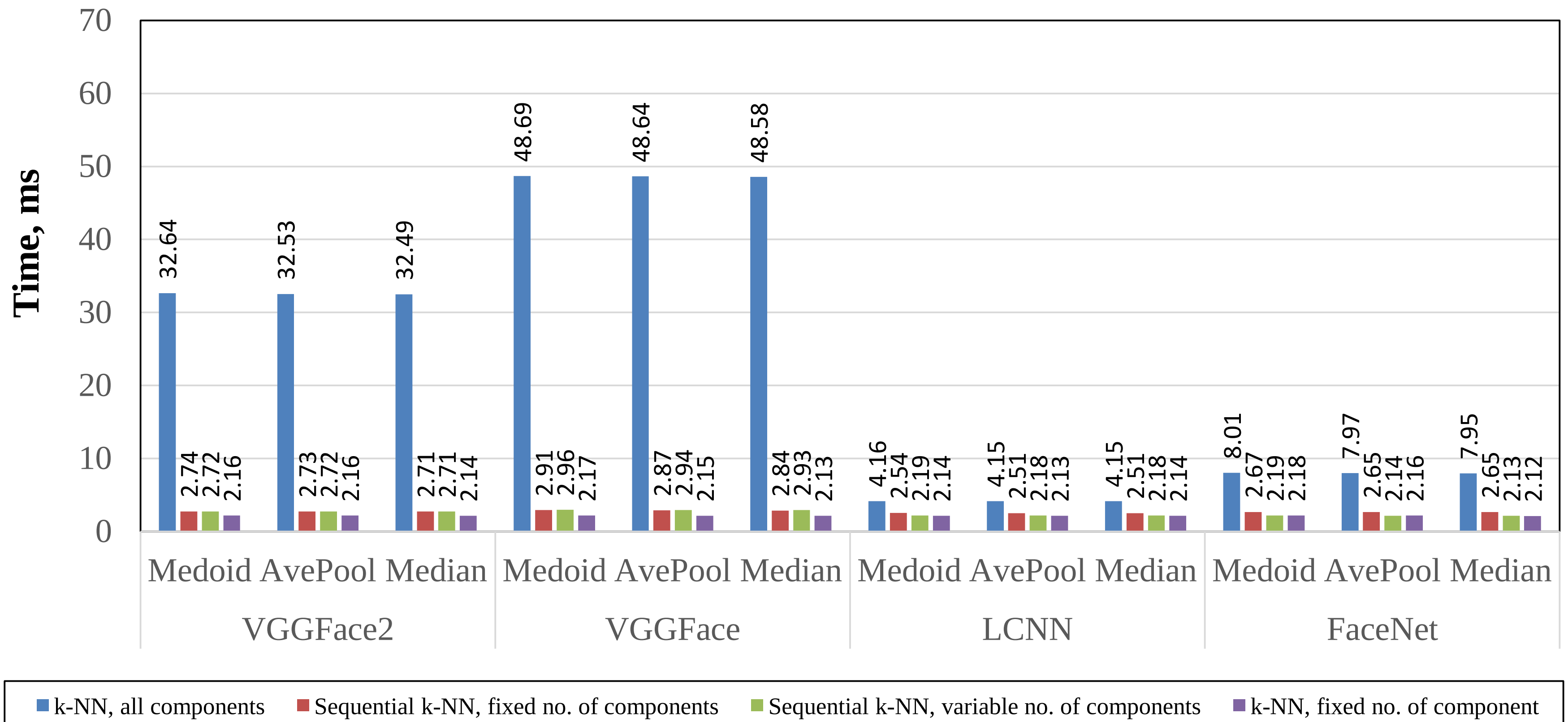


AIST

Face recognition results for the k-NN classifier (YTF)



Face recognition results for the k-NN classifier (YTF)



Conclusion

- We proposed the modification of the k-NN classification method based on sequential analysis of high-dimensional features
- The proposed approach is up to 8-10 times faster when compared to the original k-NN method, while the accuracy decreases only by 0.1-0.9%

Future work

- To examine sophisticated classifiers in order to increase the recognition accuracy
- To use more difficult dissimilarity measures
- To study contemporary aggregation techniques instead of simple averaging of individual features

Thank you for attention!