Интерфейс мозг-компьютер: Распознавание визуальных электроэцефалографических потенциалов врача при чтении маммограмм

Valentina Sulimova
Tula State University, Tula, Russia
Sergey Bukhonov
Moscow Institute of Physics and Technology, Moscow, Russia
Olga Krasotkina
Markov Processes International, New Jersey, USA
Vadim Mottl
Computing Center of the Russian Academy of Sciences, Moscow, Russia
David Windridge
Middlesex University, London, United Kingdom

Traditional electroencephalography

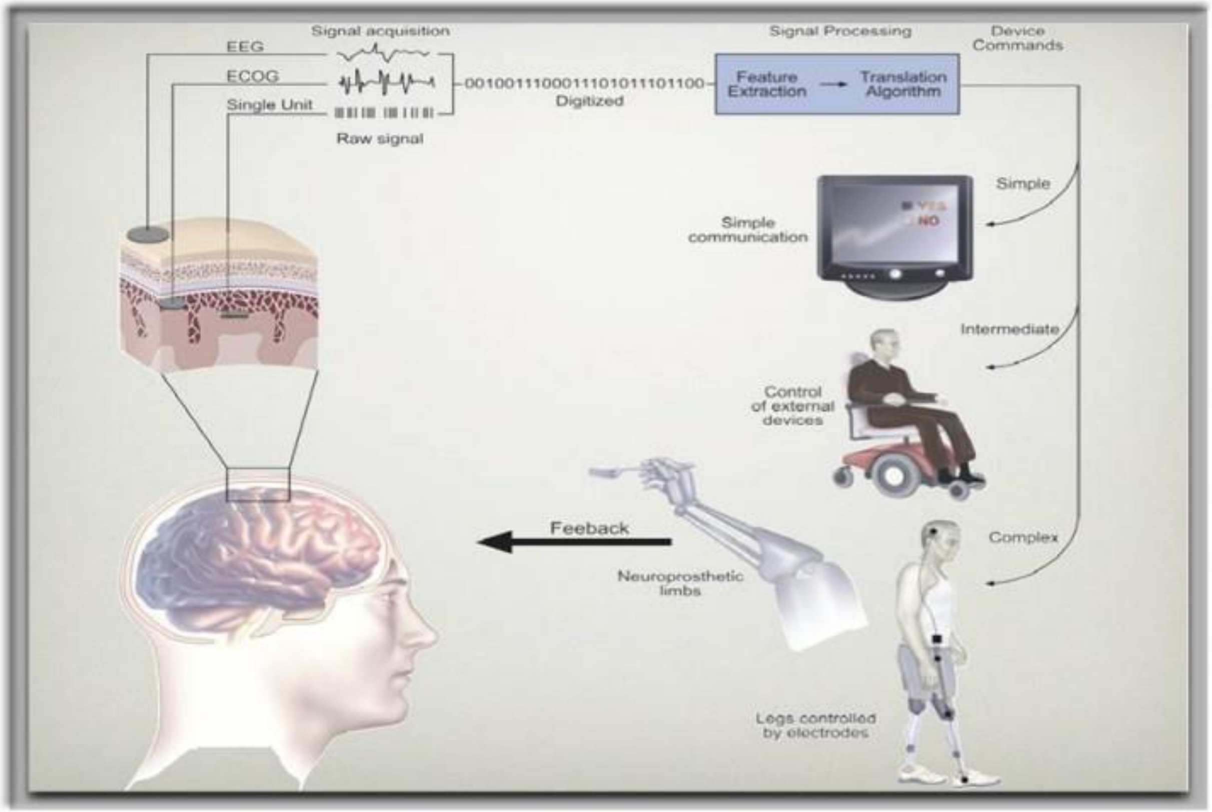
Electroencephalography was originally invented and is broadly used as a means to study mechanisms by which human behavior is generated, in particular, for brain diseases diagnosis.

New role of electroencephalography: the basis of brain-computer interfaces

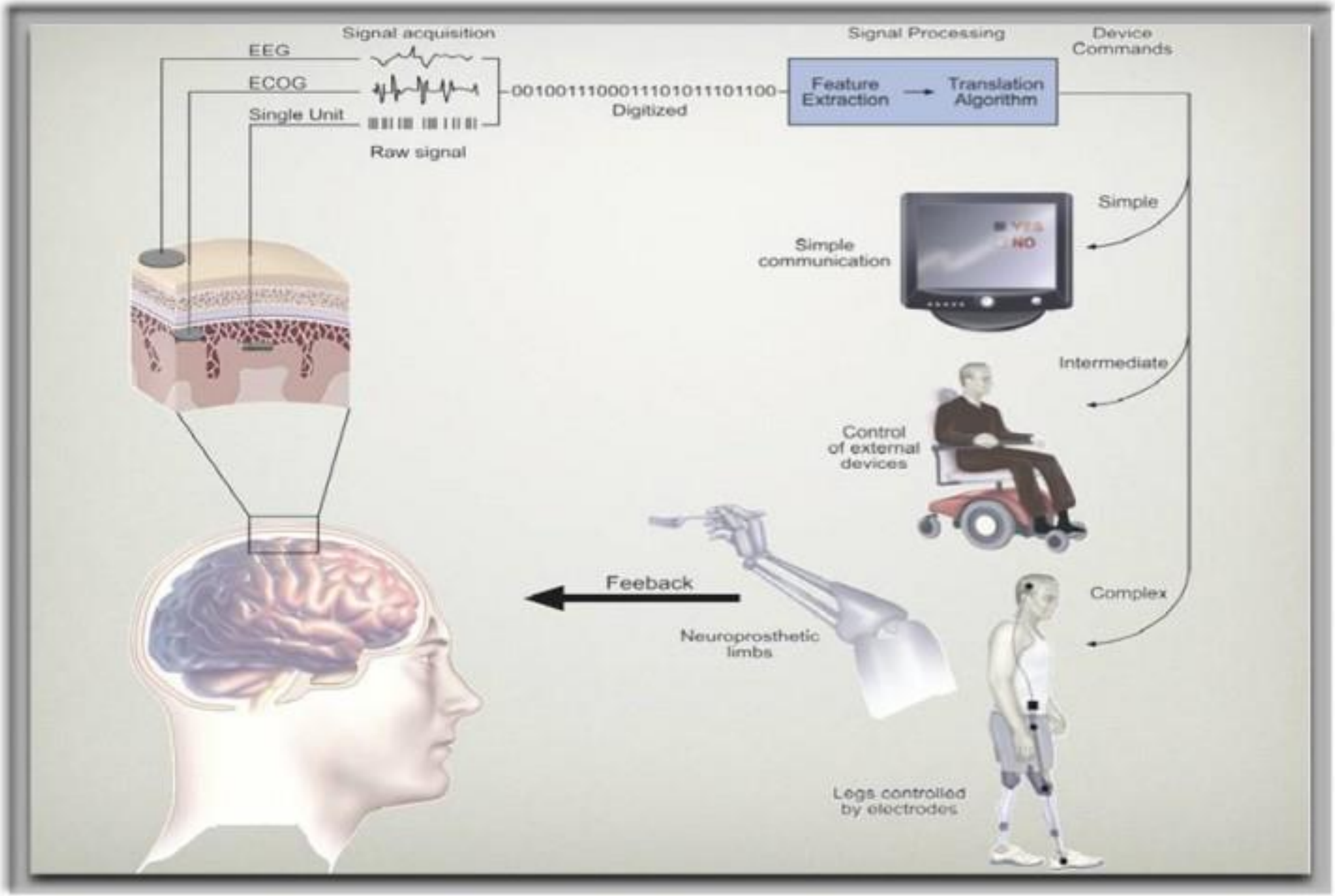
In the past decades, electroencephalography has become the basis of many brain-computer interfaces, which decode neural response to different stimuli into commands that, for instance, operate external devices like brain-driven artificial limbs or invalid chairs.

New role of electroencephalography: the basis of brain-computer interfaces

In the past decades, electroencephalography has become the basis of many brain-computer interfaces, which decode neural response to different stimuli into commands that, for instance, operate external devices like brain-driven artificial limbs or invalid chairs.

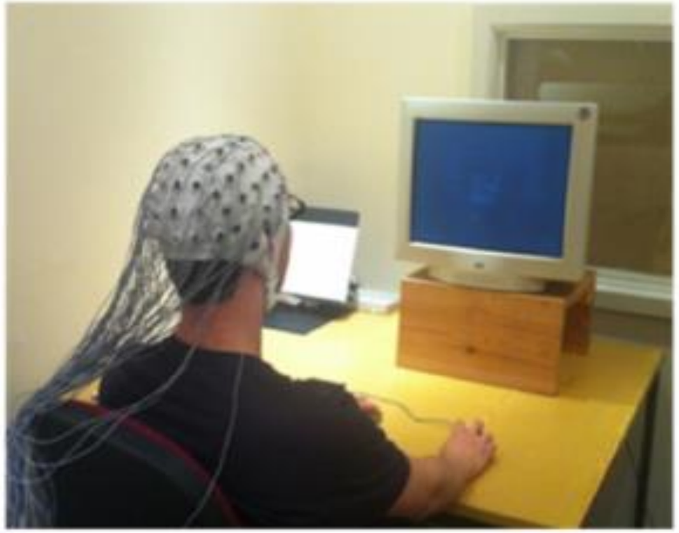


New role of electroencephalography: the basis of brain-computer interfaces

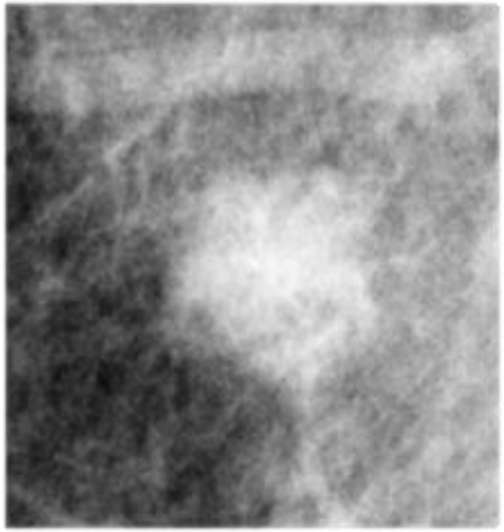


Our novel idea: EEG-based brain computer interface for outstanding X-ray mammologists

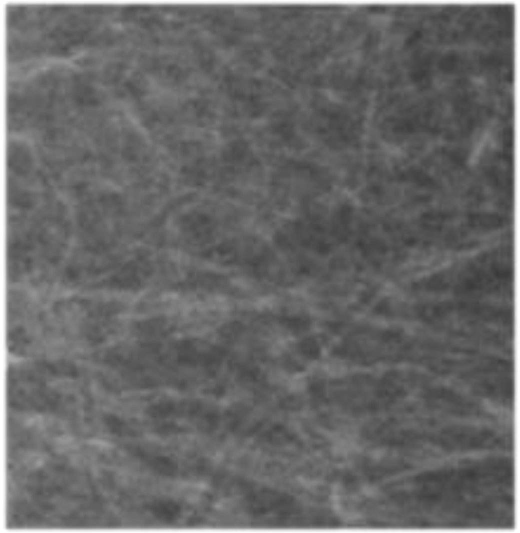
It is assumed that the person whose EEG is processed is an experienced mammologist able to reliably distinguish between X-ray mammograms of women with breast cancer and those of healthy women.



EEG registration in the process of viewing by an expert of rapidly changing mammographic images



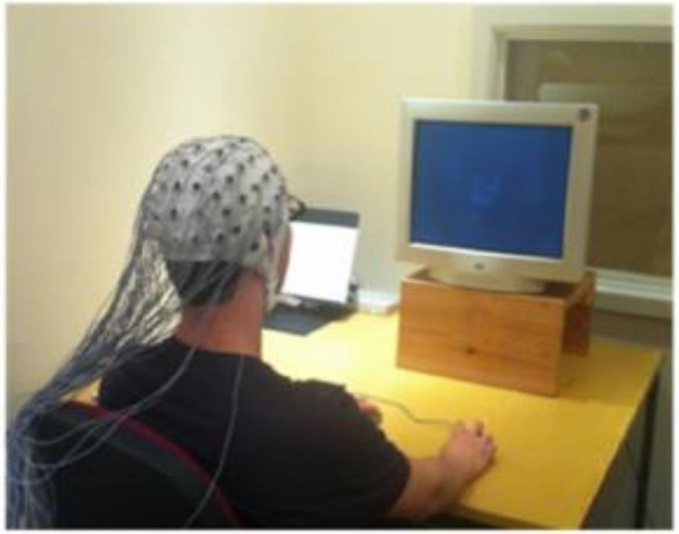
A target image: mammogram with pathology

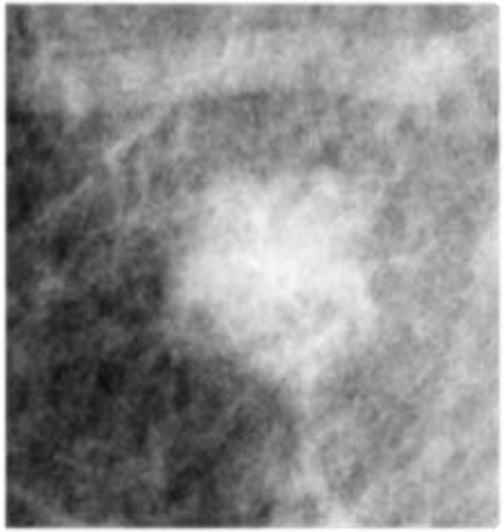


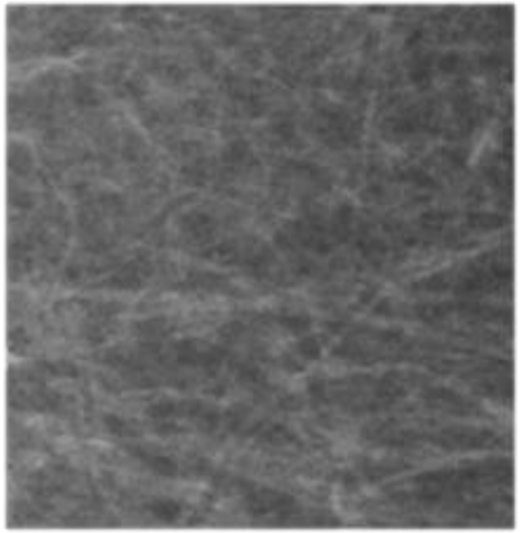
A non-target image: mammogram without pathologies

Our novel idea: EEG-based brain computer interface for outstanding X-ray mammologists

It is assumed that the person whose EEG is processed is an experienced mammologist able to reliably distinguish between X-ray mammograms of women with breast cancer and those of healthy women.



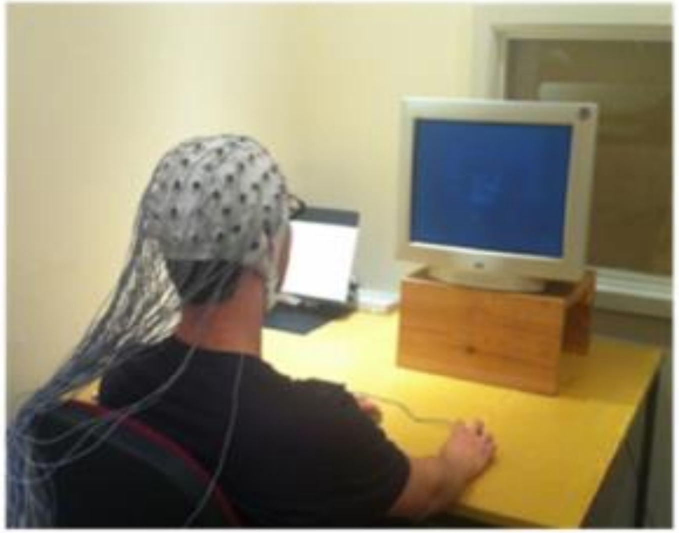


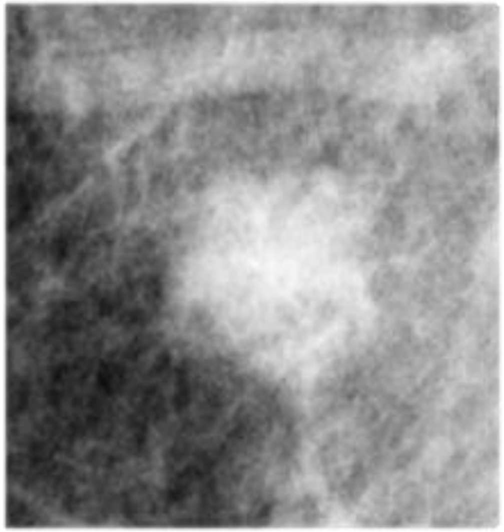


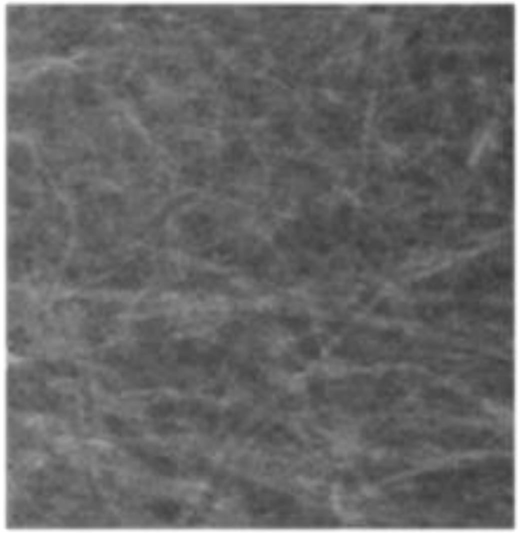
The aim is to essentially improve productivity of the rare pronounced experts by way of, first, accelerating the screening of mammographic images up to ten pictures per second, and, second, immediately detecting the eventual potentials evoked in the expert's EEG by a target (cancer) image among a crowd of non-target ones, before the expert becomes aware of this fact.

Our novel idea: EEG-based brain computer interface for outstanding X-ray mammologists

It is assumed that the person whose EEG is processed is an experienced mammologist able to reliably distinguish between X-ray mammograms of women with breast cancer and those of healthy women.







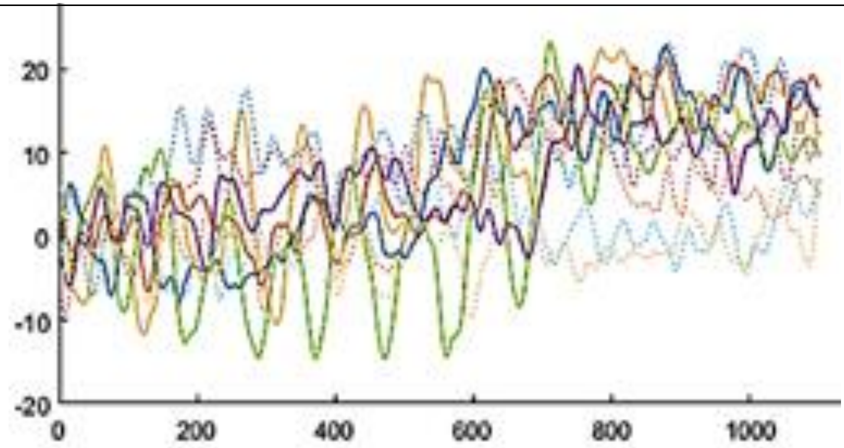
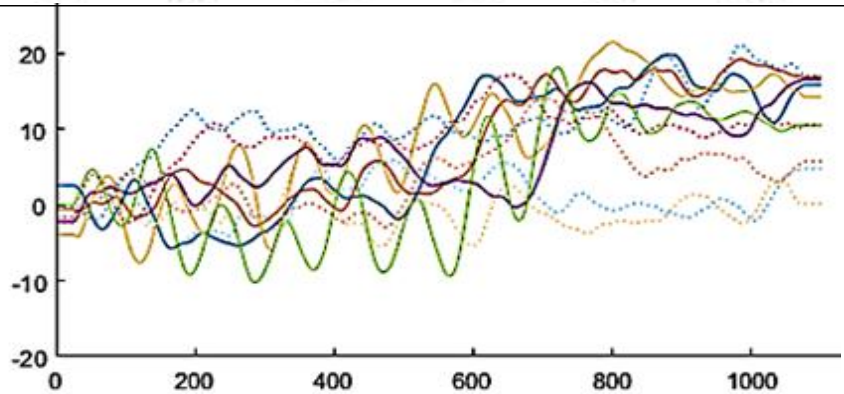
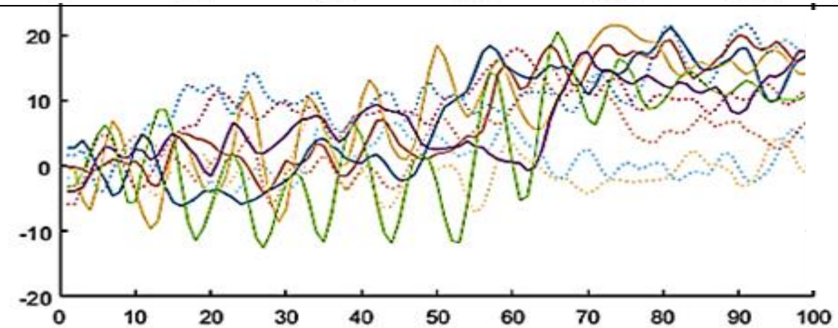
A series of 11 mammograms is shown to the expert during 1100 ms (1.1 seconds)

Two classes of mammogram series	
The target class	The non-target class
The series contains one cancer mammogram	No pathologies in the series

The aim: Finding discriminative features of the 66-channel EEG

Preprocessing of EEG fragments

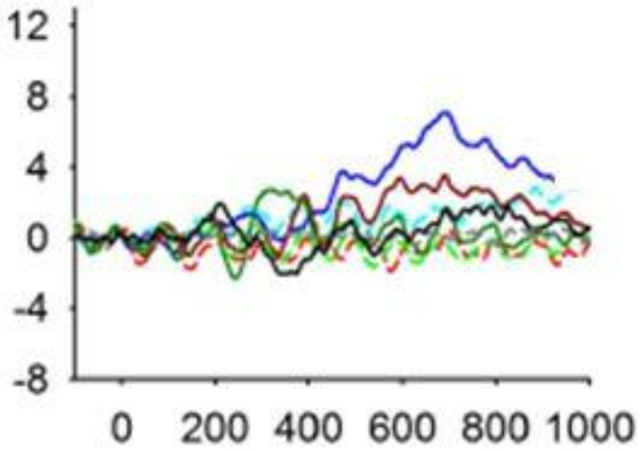
Original fragment 1.1 seconds in length, time sampling 1.1 Khz, 1100 time samples

Before preprocessing $\mathbf{x}=\left(x_i\in\mathbb{R},\, i=1,\ldots,m\right)\in\mathbb{R}^m,\;\; m=1100$	
sliding window smoothing $\mathbf{x}=\left(x_i\in\mathbb{R},\, i=1,\ldots,m\right)\in\mathbb{R}^m,\;\; m=1100$	
11-fold decimation $\mathbf{x}'=\left(x_i\in\mathbb{R},\, i=1,\ldots,n\right)\in\mathbb{R}^m,\;\; n=100$	

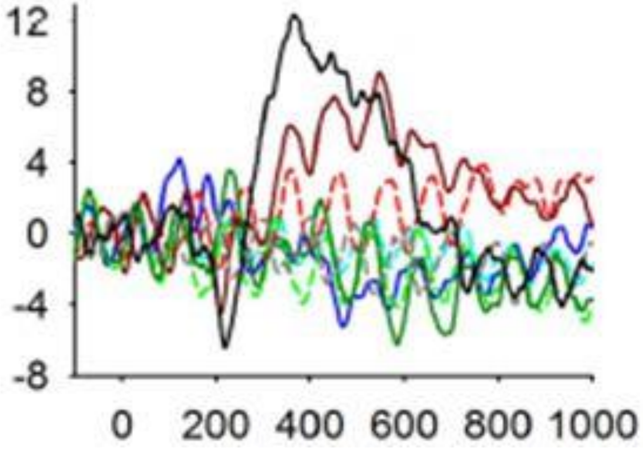
Examples of visual distinction between EEG signals induced by watching of pathological and normal mammograms

Averaged signals registered from several experts

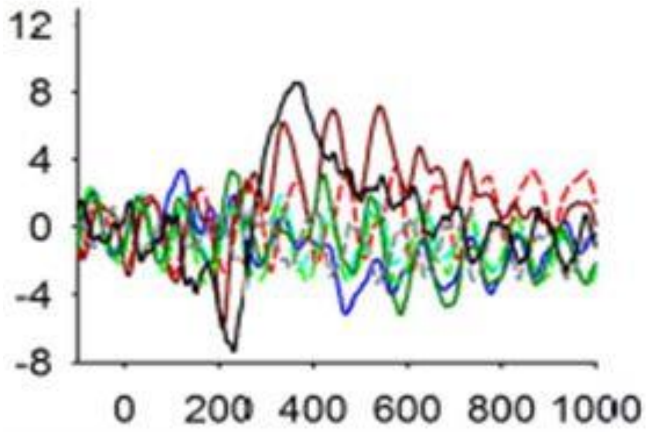
Lead 1



Lead 2



Lead 3



time from start of demonstration, ms

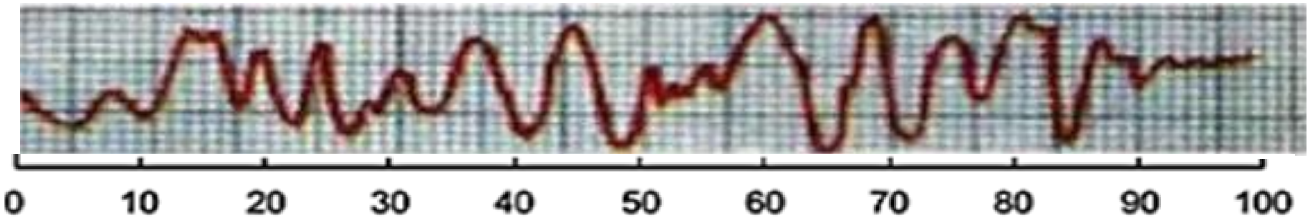
Solid lines – evoked potentials in EEG induced by cancer mammograms

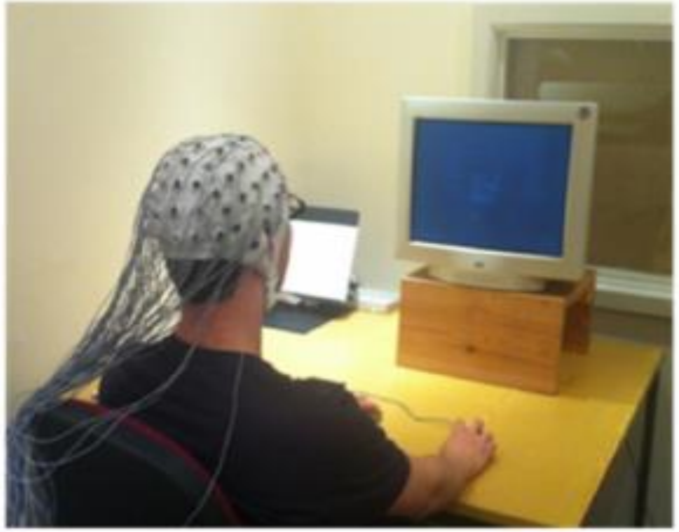
Dashed lines – undisturbed EEG from normal mammograms

The combined feature vector for recognition of image series containing a pathology

Overall number of EEG leads $k = 1,\dots,K,\ \ K = 66$

each represented by a vector of $m = 100$ sequential samples





$\mathbf{x}_k = \left(x_{k,1},\dots,x_{k,m}\right) \in \mathbb{R}^m = \mathbb{R}^{100}$

The entire feature vector $\mathbf{x} = \left(\mathbf{x}_1,\dots,\mathbf{x}_K\right) \in \mathbb{R}^{mK} = \mathbb{R}^{6600}$

A huge dimension of the feature space!

The empirical data set

The training set of EEG fragments	The test set
$\left\{(\mathbf{x}_j \in \mathbb{R}^{6600}, y_j),\ j = 1,\dots,N\right\},\ N = 196$ $y_j = 1$ $N_1 = 98$ target class (one cancer image in the series) $y_j = -1$ $N_{-1} = 98$ non-target class (no cancer image	$N = 558$ $N_1 = 279,\ \ N_{-1} = 279$

The size of the training set is quite moderate!

Specificity of pattern recognition learning in high-dimensional feature space from relatively small training set

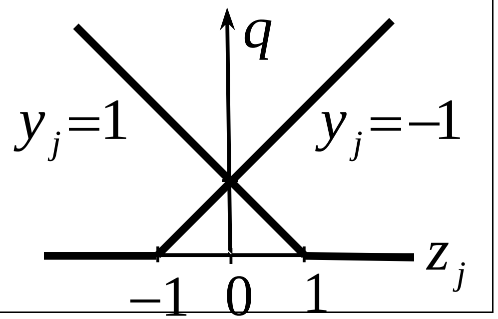
Empirical risk minimization

$$\begin{cases} \frac{1}{N}\sum_{j=1}^{N} q(y_j, z_j) \to \min(\mathbf{a}, b),\ y_j = \pm 1 \\ z_j = \mathbf{a}^T \mathbf{x}_j + b,\ \mathbf{a} \in \mathbb{R}^n, b \in \mathbb{R} \end{cases}$$

$q(y_j, z_j)$ – link (loss) function

SVM
pattern recognition

$$q(y_j, z_j) = \max(0, 1 - y_j z_j)$$



In our case, $n = 6600$, $N = 196$, $n \gg N$:
Overfitting and low generalization performance are inevitable!

Regularization – a way of enhancing the generalization performance
Our proposal: Combination of two novel kinds of regularization functions

$$\underbrace{\alpha \sum_{i=2}^{n} (a_i - a_{i-1})^2}_{\text{Smoothness regularization}} + \underbrace{\gamma \sum_{i=1}^{n} \begin{pmatrix} 2\mu\,|a_i|, & |a_i| \le \mu \\ \mu^2 + a_i^2, & |a_i| > \mu \end{pmatrix}}_{\text{Selective ridge regularization}} + \sum_{j=1}^{N} \underbrace{q(y_j, \mathbf{a}^T \mathbf{x}_j + b)}_{\text{Link function, in our case, SVM}} \to \min\left(\mathbf{a} \in \mathbb{R}^n, b \in \mathbb{R}\right)$$

Two novel kinds of regularization functions

$$\underbrace{\alpha\sum_{i=2}^{n}(a_i-a_{i-1})^2}_{\textit{Smoothness regularization}} + \underbrace{\gamma\sum_{i=1}^{n}\begin{pmatrix}2\mu\,|a_i|, & |a_i|\le\mu \\ \mu^2+a_i^2, & |a_i|>\mu\end{pmatrix}}_{\textit{Selective ridge regularization}} + \sum_{j=1}^{N}\underbrace{q(y_j,\mathbf{a}^T\mathbf{x}_j+b)}_{\textit{Link function, in our case, SVM}} \to \min\left(\mathbf{a}\in\mathbb{R}^n,\, b\in\mathbb{R}\right)$$

Smoothness regularization

$\alpha\sum_{i=2}^{n}(a_i-a_{i-1})^2 \to \min\limits_{(a_1\cdots a_n)}$ Smoothness parameter $0\le\alpha<\infty$	The greater α, the closer to each other become coefficients $(a_1\cdots a_n)$, the more similar will be influence of features $(x_1\cdots x_n)$

Selective (sparse) ridge regularization

$\sum_{i=1}^{n}\begin{pmatrix}2\mu\,|a_i|, & |a_i|\le\mu \\ \mu^2+a_i^2, & |a_i|>\mu\end{pmatrix} \to \min\limits_{(a_1\cdots a_n)}$ Selectivity parameter $0\le\mu<\infty$	The greater μ, the greater number of coefficients become zero $a_i=0$, the greater number of features x_i are suppressed

To learn how to solve such problems jointly, please attend our talk tomorrow at 4 p.m.

“*Linear complexity algorithms for high dimensional SVM and regression problems with smart sparse regularization*”

Two novel kinds of regularization functions

$$\alpha \underbrace{\sum_{i=2}^{n}(a_i - a_{i-1})^2}_{\substack{\textit{Smoothness}\\ \textit{regularization}}} + \gamma \underbrace{\sum_{i=1}^{n}\begin{pmatrix} 2\mu\,|a_i|, & |a_i| \le \mu \\ \mu^2 + a_i^2, & |a_i| > \mu \end{pmatrix}}_{\substack{\textit{Selective ridge}\\ \textit{regularization}}} + \sum_{j=1}^{N} \underbrace{q(y_j, \mathbf{a}^T\mathbf{x}_j + b)}_{\substack{\textit{Link function,}\\ \textit{in our case, SVM}}} \to \min\left(\mathbf{a} \in \mathbb{R}^n, b \in \mathbb{R}\right)$$

Smoothness regularization

$\alpha\sum_{i=2}^{n}(a_i - a_{i-1})^2 \to \min_{(a_1 \cdots a_n)}$ Smoothness parameter $0 \le \alpha < \infty$	The greater α, the closer to each other become coefficients $(a_1 \cdots a_n)$, the more similar will be influence of features $(x_1 \cdots x_n)$

Selective (sparse) ridge regularization

$\sum_{i=1}^{n}\begin{pmatrix} 2\mu\,|a_i|, & |a_i| \le \mu \\ \mu^2 + a_i^2, & |a_i| > \mu \end{pmatrix} \to \min_{(a_1 \cdots a_n)}$ Selectivity parameter $0 \le \mu < \infty$	The greater μ, the greater number of coefficients become zero $a_i = 0$, the greater number of features x_i are suppressed

In this talk, we consider a simplified technique. The idea is to apply these regularizations in turn, first smoothness, then selectivity.

The regularized Support Vector Machine (SVM) for two-class recognition of evoked potentials in EEG

The classical SVM

$$\gamma\sum_{i=2}^{n} a_i^2+\sum_{j=1}^{N}\max\Big(0,\ 1-y_j\,(\mathbf{a}^T\mathbf{x}_j+b)\Big)\to\min\big(\mathbf{a},b\big)$$

The regularized Support Vector Machine (SVM) for two-class recognition of evoked potentials in EEG

The classical SVM

$$\begin{cases} \gamma \sum_{i=2}^{n} a_i^2 + \sum_{j=1}^{N} \delta_j \to \min\left(\mathbf{a}, b, \delta_1, ..., \delta_N\right) \\ y_j (\mathbf{a}^T \mathbf{x}_j + b) \geq 1 - \delta_j,\ \delta_j \geq 0,\ j = 1, ..., N \end{cases}$$

Equivalent formulation
(V. Vapnik, C. Cortes, 1995)

The regularized Support Vector Machine (SVM) for two-class recognition of evoked potentials in EEG

The classical SVM

$$\begin{cases} \gamma \sum_{i=2}^{n} a_i^2 + \sum_{j=1}^{N} \delta_j \to \min\left(\mathbf{a}, b, \delta_1, ..., \delta_N\right) \\ y_j(\mathbf{a}^T \mathbf{x}_j + b) \geq 1 - \delta_j,\ \delta_j \geq 0,\ j = 1, ..., N \end{cases}$$

Equivalent formulation
(V. Vapnik, C. Cortes, 1995)

is applied separately twice to the EEG signals in each of 66 leads, $\mathbf{a}, \mathbf{x}_{k,j} \in \mathbb{R}^{100}$,
$k = 1, ..., 66$, within the bounds of the training set $j = 1, ..., 196$,
obtained by 11-fold decimation, first, from the original EEG signals, and then from the smoothed ones.

The regularized Support Vector Machine (SVM) for two-class recognition of evoked potentials in EEG

The classical SVM

$$\begin{cases} \gamma\sum_{i=2}^{n} a_i^2 + \sum_{j=1}^{N}\delta_j \to \min\left(\mathbf{a}, b, \delta_1, ..., \delta_N\right) \\ y_j(\mathbf{a}^T\mathbf{x}_j + b) \geq 1 - \delta_j,\ \delta_j \geq 0,\ j = 1, ..., N \end{cases}$$

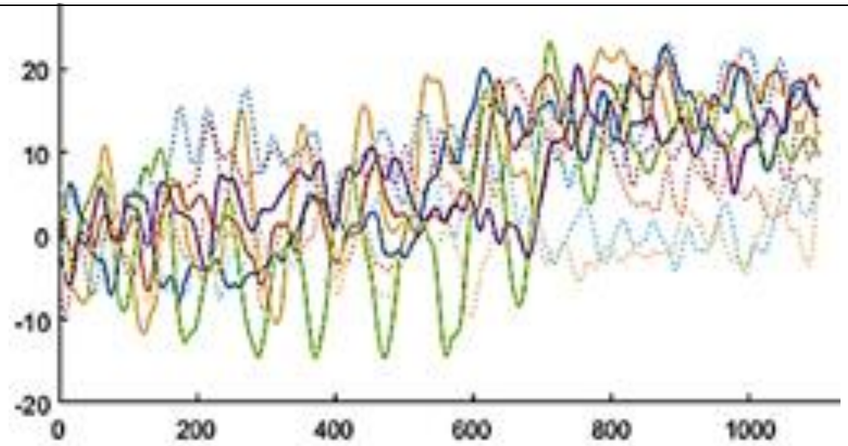
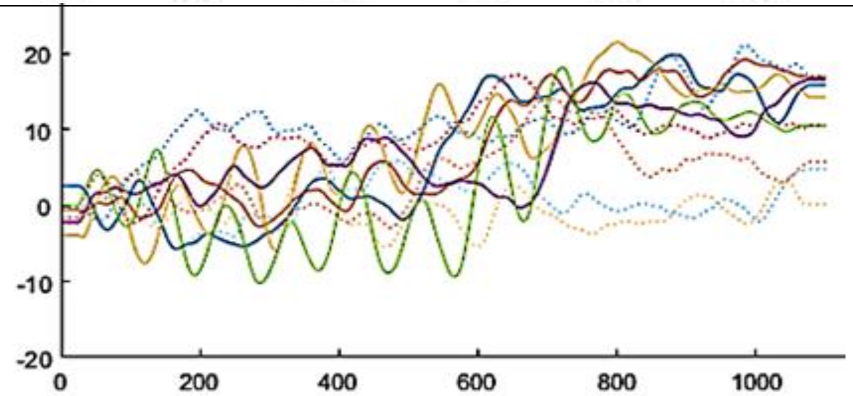
Equivalent formulation (V. Vapnik, C. Cortes, 1995)

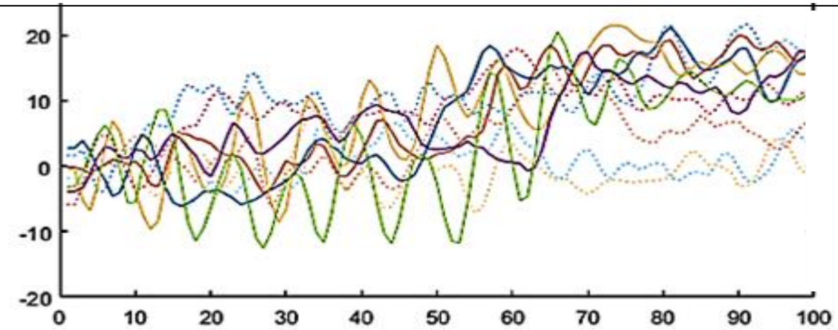
is applied separately twice to the EEG signals in each of 66 leads, $\mathbf{a}, \mathbf{x}_{k,j} \in \mathbb{R}^{100}$, $k = 1, ..., 66$, within the bounds of the training set $j = 1, ..., 196$, obtained by 11-fold decimation, first, from the original EEG signals, and then from the smoothed ones.

Remember:

The training set of EEG fragments	The test set
$\left\{(\mathbf{x}_j \in \mathbb{R}^{6600}, y_j),\ j = 1, ..., N\right\},\ N = 196$ $y_j = 1$ $N_1 = 98$ target class (one cancer image in the series) $y_j = -1$ $N_{-1} = 98$ non-target class (no cancer image	$N = 558$ $N_1 = 279$, $N_{-1} = 279$

Remember: Preprocessing of EEG fragments

Before preprocessing $\mathbf{x} = \left(x_i \in \mathbb{R},\, i = 1,\ldots,m\right) \in \mathbb{R}^m,\;\; m = 1100$	
After sliding window smoothing $\mathbf{x} = \left(x_i \in \mathbb{R},\, i = 1,\ldots,m\right) \in \mathbb{R}^m,\;\; m = 1100$	

11-fold decimation $\mathbf{x}' = \left(x_i \in \mathbb{R},\, i = 1,\ldots,n\right) \in \mathbb{R}^m,\;\; n = 100$	

The regularized Support Vector Machine (SVM) for two-class recognition of evoked potentials in EEG

The classical SVM

$$\begin{cases} \gamma \sum_{i=2}^{n} a_i^2 + \sum_{j=1}^{N} \delta_j \to \min\left(\mathbf{a} \in \mathbb{R}^n, b, \delta_1, \ldots, \delta_N\right) \\ y_j(\mathbf{a}^T \mathbf{x}_j + b) \geq 1 - \delta_j,\ \delta_j \geq 0,\ j = 1, \ldots, N \end{cases}$$

Equivalent formulation (V. Vapnik, C. Cortes, 1995)

is applied separately twice to the EEG signals in each of 66 leads, $\mathbf{a}, \mathbf{x}_{k,j} \in \mathbb{R}^{100}$, $k = 1, \ldots, 66$, within the bounds of the training set $j = 1, \ldots, 196$, obtained by 11-fold decimation, first, from the original EEG signals, and then from the smoothed ones.

The regularized Support Vector Machine (SVM) for two-class recognition of evoked potentials in EEG

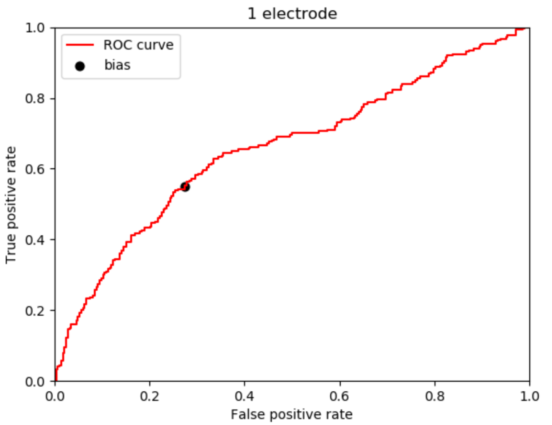
The classical SVM

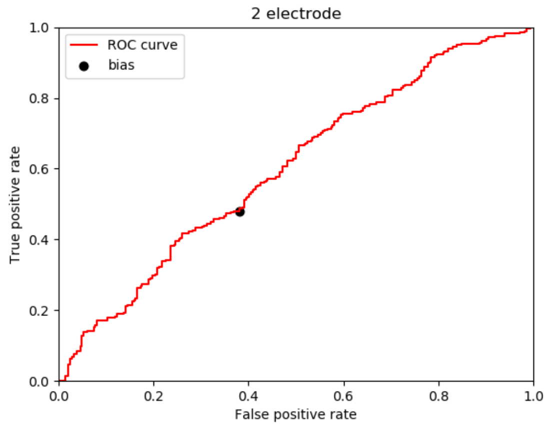
$$\begin{cases} \gamma\sum_{i=2}^{n} a_i^2+\sum_{j=1}^{N}\delta_j \to \min\left(\mathbf{a}\in\mathbb{R}^n, b, \delta_1,\ldots,\delta_N\right) \\ y_j(\mathbf{a}^T\mathbf{x}_j+b)\geq 1-\delta_j,\ \delta_j\geq 0,\ j=1,\ldots,N \end{cases}$$

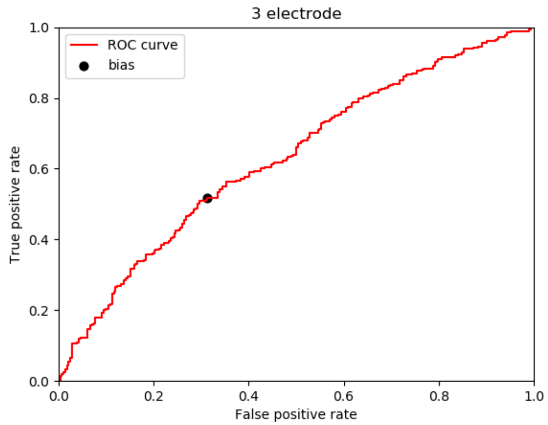
Equivalent formulation (V. Vapnik, C. Cortes, 1995)

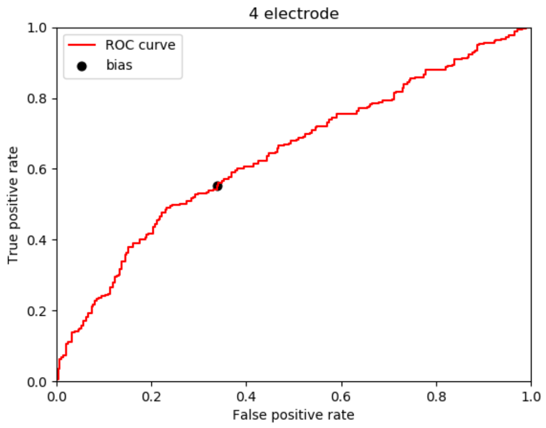
is applied separately twice to the EEG signals in each of 66 leads, $\mathbf{a},\mathbf{x}_{k,j}\in\mathbb{R}^{100}$, $k=1,\ldots,66$, within the bounds of the training set $j=1,\ldots,196$, obtained by 11-fold decimation, first, from the original EEG signals, and then from the smoothed ones. In our experiments, we put $\gamma\ll 1$, i.e., $\gamma>0$ but $\gamma\to 0$.

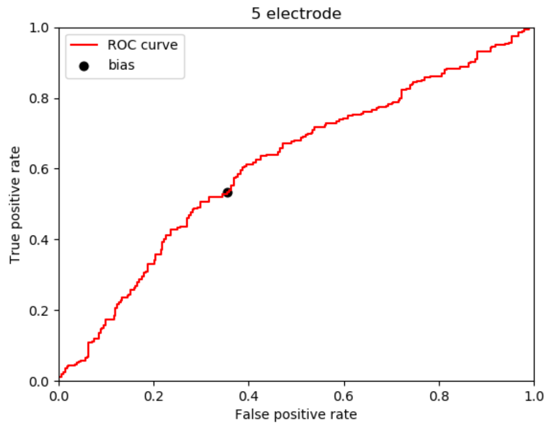
We computed ROC curves (*Receiver Operating Characteristic*) and the respective values of the AUC criterion (*Area Under Curve*) for each of 2×66 results of training (non-smoothed, smoothed), and, in addition, for 2 results (non-smoothed, smoothed), obtained from the concatenation $\mathbf{x}\in\mathbb{R}^{6600}$ of all the 66 EEG fragments as a joint signal of length $n=6600=100\times 66$.

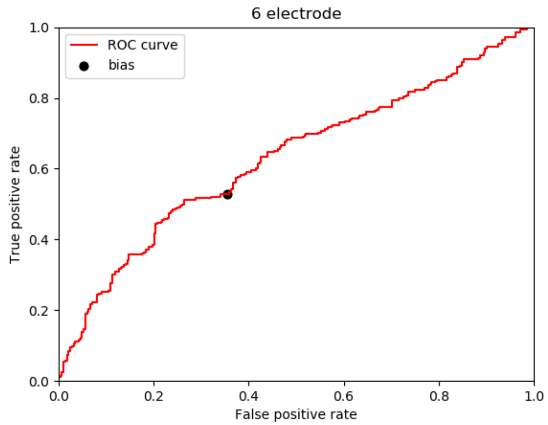
<table>
<tr><td>ROC curves
of the training results</td><td>
</td></tr>
<tr><td></td><td>Electrode 1 AUC=0,6985</td></tr>
</table>

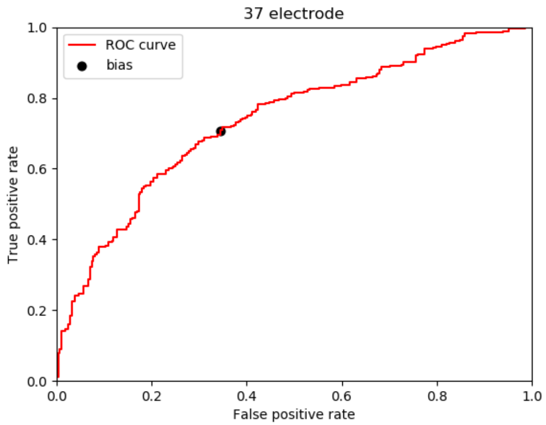
ROC curves of the training results	
	Electrode 2 AUC=0,6611

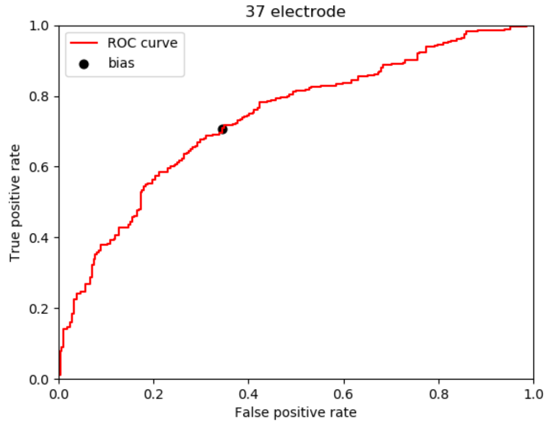
ROC curves of the training results	
	Electrode 3 AUC=0,6901

<table>
<tr><td>ROC curves
of the training results</td><td>
</td></tr>
<tr><td></td><td>Electrode 4 AUC=0,689</td></tr>
</table>

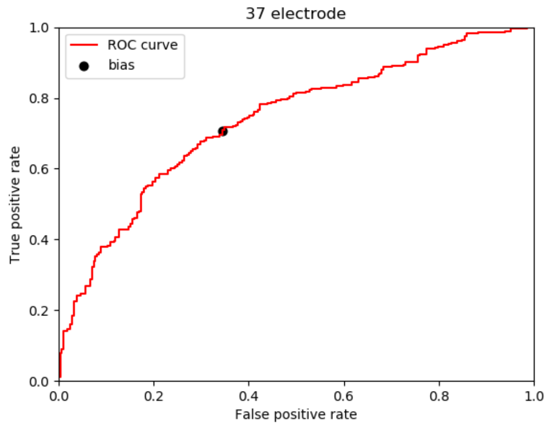
<table>
<tr><td>ROC curves
of the training results</td><td>
</td></tr>
<tr><td></td><td>Electrode 5 AUC=0,6167</td></tr>
</table>

ROC curves of the training results	
	Electrode 6 AUC=0,6806

ROC curves of the training results	
	Electrode 37 AUC=0,8451

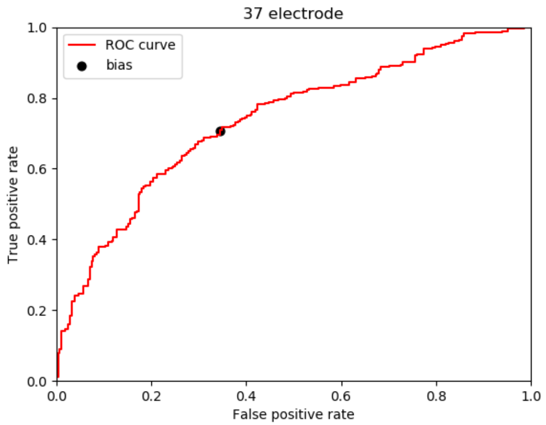
ROC curves of the training results	
	Electrode 37 AUC=0,8451

As we see, the AUC values of the training results are quite low. This means that no single electrode provides a sufficiently reliable recognition of pathologies in mammograms.

ROC curves of the training results	
	Electrode 37 AUC=0,8451

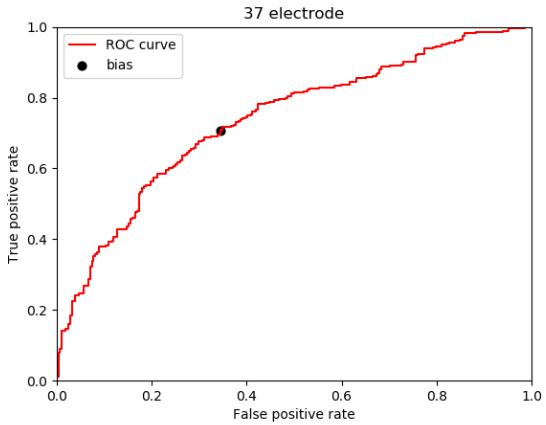
As we see, the AUC values of the training results are quite low. This means that no single electrode provides a sufficiently reliable recognition of pathologies in mammograms.

The joint EEG signal $\mathbf{x} \in \mathbb{R}^{6600}$ as concatenation of all the 66 leads yields AUC=0,815, smaller than the AUC values at the best of the single electrodes.

ROC curves of the training results	
	Electrode 37 $\boxed{\text{AUC=0,8451}}$

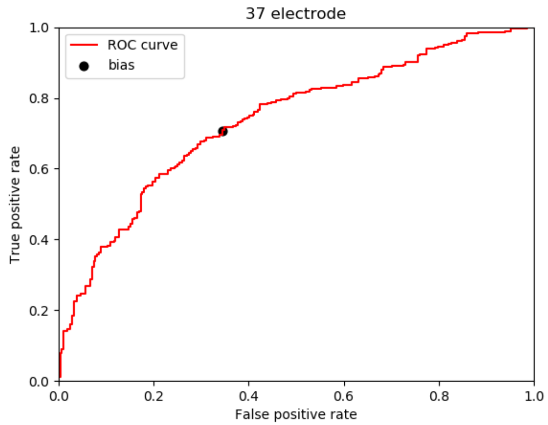
As we see, the AUC values of the training results are quite low. This means that no single electrode provides a sufficiently reliable recognition of pathologies in mammograms.

The joint EEG signal $\mathbf{x}\in\mathbb{R}^{6600}$ as concatenation of all the 66 leads yields $\boxed{\text{AUC=0,815, smaller than the AUC values at the best of the single electrodes}}$.

ROC curves **of the training results**	
	Electrode 37 AUC=0,8451

As we see, the AUC values of the training results are quite low. This means that no single electrode provides a sufficiently reliable recognition of pathologies in mammograms.

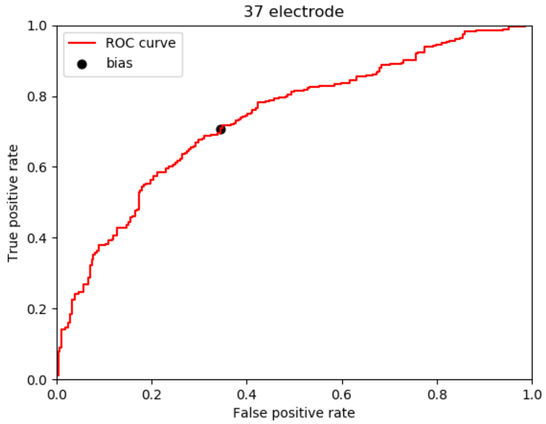
The joint EEG signal $\mathbf{x} \in \mathbb{R}^{6600}$ as concatenation of all the 66 leads yields AUC=0,815, smaller than the AUC values at the best of the single electrodes.

ROC curves of the training results	
	Electrode 37 AUC=0,8451

As we see, the AUC values of the training results are quite low. This means that no single electrode provides a sufficiently reliable recognition of pathologies in mammograms.

The joint EEG signal $\mathbf{x} \in \mathbb{R}^{6600}$ as concatenation of all the 66 leads yields AUC=0,815, smaller than the AUC values at the best of the single electrodes.

This fact witnesses of a deep overfitting: $n = 6600 \gg N = 196$.

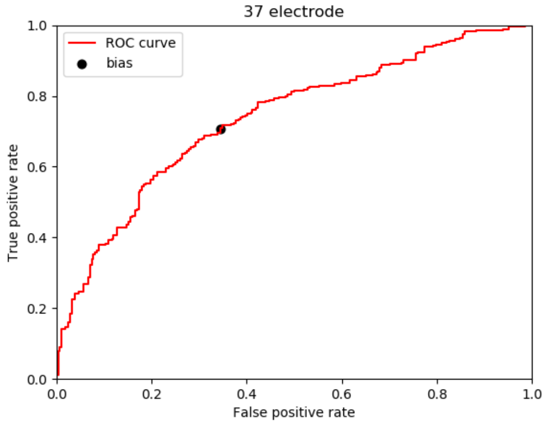
ROC curves of the training results	
	Electrode 37 AUC=0,8451

As we see, the AUC values of the training results are quite low. This means that no single electrode provides a sufficiently reliable recognition of pathologies in mammograms.

The joint EEG signal $\mathbf{x} \in \mathbb{R}^{6600}$ as concatenation of all the 66 leads yields AUC=0,815, smaller than the AUC values at the best of the single electrodes.

This fact witnesses of a deep overfitting: $n = 6600 \gg N = 196$.

Then, we chose 7 electrodes, for which the individual training had showed the best recognition quality in the test set of 558 EEG fragments.

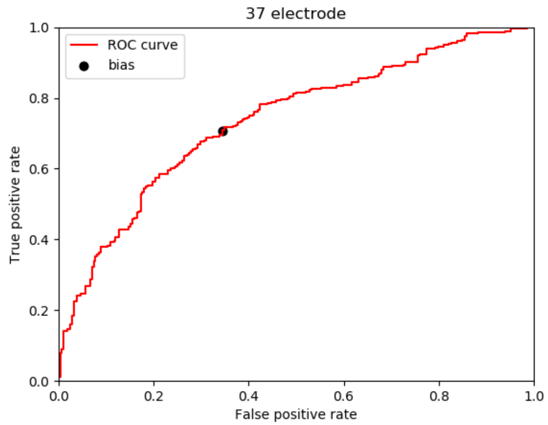
ROC curves of the training results	
	Electrode 37 AUC=0,8451

As we see, the AUC values of the training results are quite low. This means that no single electrode provides a sufficiently reliable recognition of pathologies in mammograms.

The joint EEG signal $\mathbf{x}\in\mathbb{R}^{6600}$ as concatenation of all the 66 leads yields AUC=0,815, smaller than the AUC values at the best of the single electrodes.

This fact witnesses of a deep overfitting: $n=6600 \gg N=196$.

Then, we chose 7 electrodes, for which the individual training had showed the best recognition quality in the test set of 558 EEG fragments. The joint EEG signal $\mathbf{x}\in\mathbb{R}^{7\times 100}=\mathbb{R}^{700}$ as concatenation of the signals at electrodes 27, 28, 30, 33, 37, 42, 53 yields AUC=0,9026, much better than at the single electrodes.

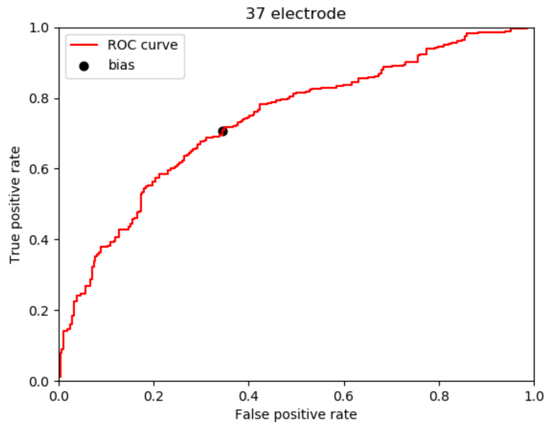
ROC curves of the training results	
	Electrode 37 AUC=0,8451

As we see, the AUC values of the training results are quite low. This means that no single electrode provides a sufficiently reliable recognition of pathologies in mammograms.

The joint EEG signal $\mathbf{x}\in\mathbb{R}^{6600}$ as concatenation of all the 66 leads yields AUC=0,815, smaller than the AUC values at the best of the single electrodes.

This fact witnesses of a deep overfitting: $n=6600 \gg N=196$.

Then, we chose 7 electrodes, for which the individual training had showed the best recognition quality in the test set of 558 EEG fragments. The joint EEG signal $\mathbf{x}\in\mathbb{R}^{7\times 100}=\mathbb{R}^{700}$ as concatenation of the signals at electrodes 27, 28, 30, 33, 37, 42, 53 yields $\boxed{\text{AUC=0,9026}}$, much better than at the single electrodes.

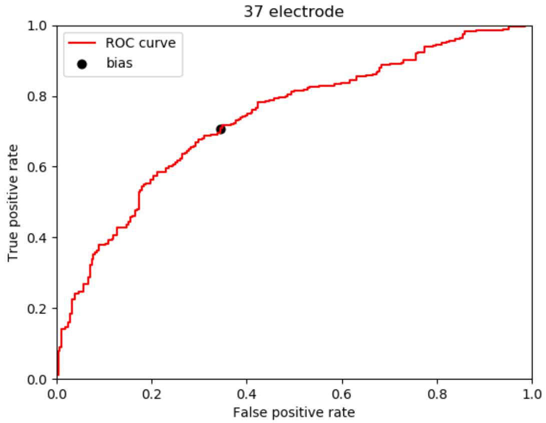
<table>
<tr><td>ROC curves
of the training results</td><td>
</td></tr>
<tr><td></td><td>Electrode 37 AUC=0,8451</td></tr>
</table>

As we see, the AUC values of the training results are quite low. This means that no single electrode provides a sufficiently reliable recognition of pathologies in mammograms.

The joint EEG signal $\mathbf{x} \in \mathbb{R}^{6600}$ as concatenation of all the 66 leads yields AUC=0,815, smaller than the AUC values at the best of the single electrodes.

This fact witnesses of a deep overfitting: $n = 6600 \gg N = 196$.

Then, we chose 7 electrodes, for which the individual training had showed the best recognition quality in the test set of 558 EEG fragments. The joint EEG signal $\mathbf{x} \in \mathbb{R}^{7 \times 100} = \mathbb{R}^{700}$ as concatenation of the signals at electrodes 27, 28, 30, 33, 37, 42, 53 yields AUC=0,9026, much better than at the single electrodes.

ROC curves of the training results	
	Electrode 37 AUC=0,8451

As we see, the AUC values of the training results are quite low. This means that no single electrode provides a sufficiently reliable recognition of pathologies in mammograms.

The joint EEG signal $\mathbf{x}\in\mathbb{R}^{6600}$ as concatenation of all the 66 leads yields AUC=0,815, smaller than the AUC values at the best of the single electrodes.

This fact witnesses of a deep overfitting: $n=6600 \gg N=196$.

Then, we chose 7 electrodes, for which the individual training had showed the best recognition quality in the test set of 558 EEG fragments. The joint EEG signal $\mathbf{x}\in\mathbb{R}^{7\times100}=\mathbb{R}^{700}$ as concatenation of the signals at electrodes 27, 28, 30, 33, 37, 42, 53 yields AUC=0,9026, much better than at the single electrodes.

Thus, overfitting can be essentially relieved by eliminating low-informative electrodes and, thereby, reducing the dimension of the feature space.

Feature-selective Support Vector Machine (SVM) for two-class recognition of evoked potentials in EEG

Remember: Up to now we applied the classical SVM

$$\begin{cases} \gamma \sum_{i=2}^{n} a_i^2 + \sum_{j=1}^{N} \delta_j \to \min \left(\mathbf{a} \in \mathbb{R}^n, b, \delta_1, \ldots, \delta_N \right) \\ y_j (\mathbf{a}^T \mathbf{x}_j + b) \geq 1 - \delta_j,\ \delta_j \geq 0,\ j = 1, \ldots, N,\ \gamma \ll 1 \end{cases}$$

(V. Vapnik, C. Cortes, 1995)

$\mathbf{x}_j \in \mathbb{R}^n$, $n = 700$ – concatenation of 7 EEGs from 7 individually best electrodes 27, 28, 30, 33, 37, 42, 53.

Feature-selective Support Vector Machine (SVM) for two-class recognition of evoked potentials in EEG

Remember: Up to now we applied the classical SVM

$$\begin{cases} \gamma \sum_{i=2}^{n} a_i^2 + \sum_{j=1}^{N} \delta_j \to \min\left(\mathbf{a} \in \mathbb{R}^n, b, \delta_1, \ldots, \delta_N\right) \\ y_j(\mathbf{a}^T \mathbf{x}_j + b) \geq 1 - \delta_j,\ \delta_j \geq 0,\ j = 1, \ldots, N,\ \gamma \ll 1 \end{cases}$$

(V. Vapnik, C. Cortes, 1995)

$\mathbf{x}_j \in \mathbb{R}^n$, $n = 700$ – concatenation of 7 EEGs from 7 individually best electrodes 27, 28, 30, 33, 37, 42, 53.

Now we apply the feature selective-SVM

$$\begin{cases} \gamma \sum_{i=1}^{n} \begin{pmatrix} 2\mu |a_i|, |a_i| \leq \mu \\ \mu^2 + a_i^2, |a_i| > \mu \end{pmatrix} + \sum_{j=1}^{N} \delta_j \to \min\left(\mathbf{a} \in \mathbb{R}^n, b, \delta_1, \ldots, \delta_N\right) \\ y_j(\mathbf{a}^T \mathbf{x}_j + b) \geq 1 - \delta_j,\ \delta_j \geq 0,\ j = 1, \ldots, N,\ \gamma \ll 1 \end{cases}$$

(A. Tatarchuk, 2008)

Tatarchuk, et al. Selectivity supervision in combining pattern-recognition modalities by feature- and kernel-selective Support Vector Machines. Proc. ICPR, 2008.

$\mathbf{x}_j \in \mathbb{R}^n$, $n = 1300$ – concatenation of 13 EEGs from 7 individually best and 5 individually worst electrodes 16, 26, 27, 28, 30, 33, 34, 37, 39, 42, 46, 53, 60.

Feature-selective Support Vector Machine (SVM) for two-class recognition of evoked potentials in EEG

$$\begin{cases} \gamma \sum_{i=1}^{n} \begin{pmatrix} 2\mu \, | a_i |, | a_i | \le \mu \\ \mu^2 + a_i^2, | a_i | > \mu \end{pmatrix} + \sum_{j=1}^{N} \delta_j \to \min \left(\mathbf{a} \in \mathbb{R}^n, b, \delta_1, \ldots, \delta_N \right) \\ y_j (\mathbf{a}^T \mathbf{x}_j + b) \ge 1 - \delta_j, \, \delta_j \ge 0, \; j = 1, \ldots, N, \; \gamma \ll 1 \end{cases}$$

(B. Tatarchuk, 2008)

Tatarchuk, et al. Selectivity supervision in combining pattern-recognition modalities by feature- and kernel-selective Support Vector Machines. Proc. ICPR, 2008.

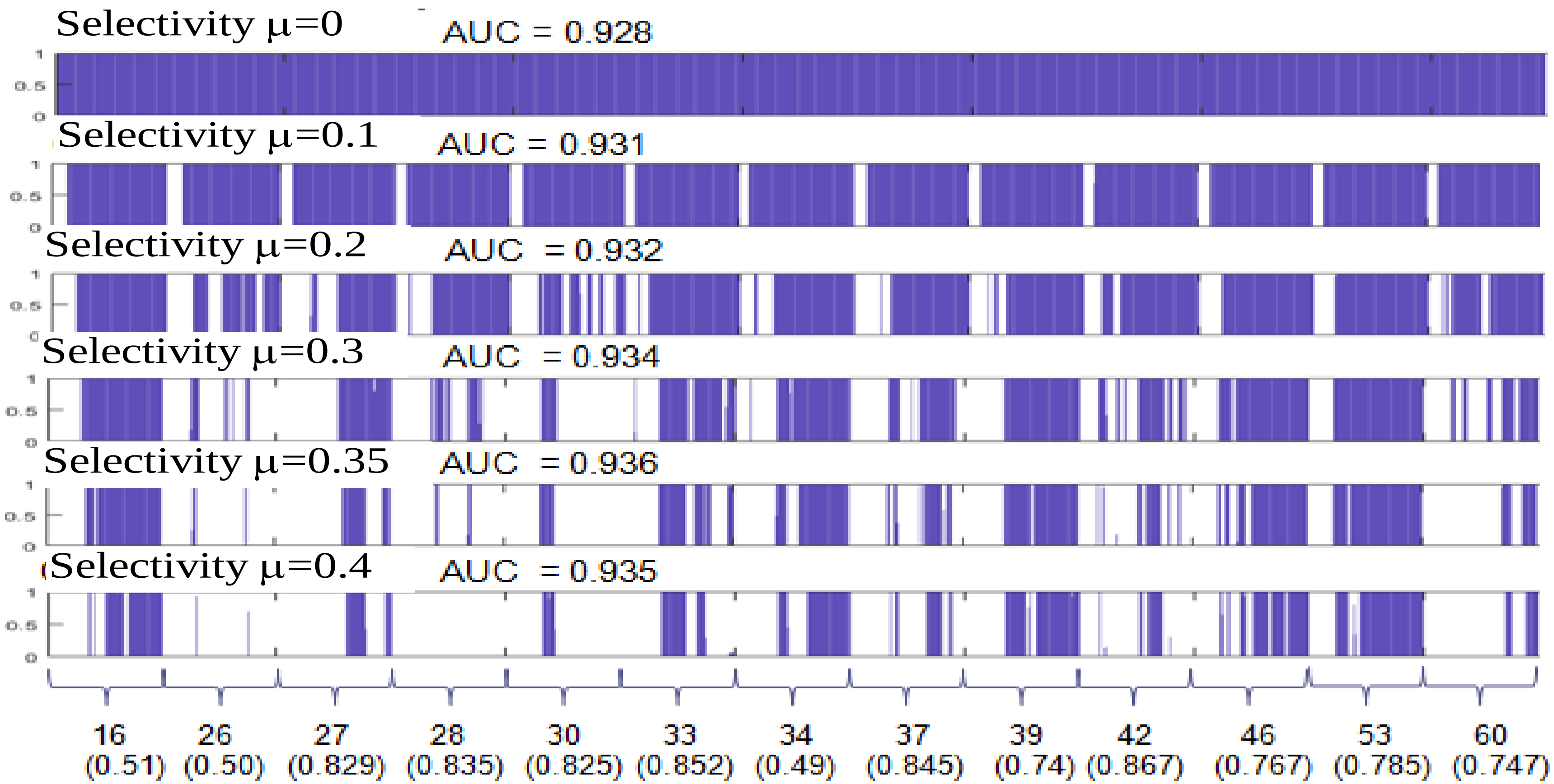
$\mathbf{x}_j \in \mathbb{R}^n$, $n = 1300$ – concatenation of 13 EEGs from 7 individually best and 5 individually worst electrodes 16, 26, 27, 28, 30, 33, 34, 37, 39, 42, 46, 53, 60.

Feature-selective Support Vector Machine (SVM) for two-class recognition of evoked potentials in EEG

$$\begin{cases} \gamma \sum_{i=1}^{n} \begin{pmatrix} 2\mu\,|a_i|, |a_i| \le \mu \\ \mu^2 + a_i^2, |a_i| > \mu \end{pmatrix} + \sum_{j=1}^{N} \delta_j \to \min\left(\mathbf{a} \in \mathbb{R}^n, b, \delta_1, \ldots, \delta_N\right) \\ y_j(\mathbf{a}^T \mathbf{x}_j + b) \ge 1 - \delta_j,\ \delta_j \ge 0,\ j = 1, \ldots, N,\ \gamma \ll 1 \end{cases}$$

(C. Tatarchuk, 2008)

Tatarchuk, et al. Selectivity supervision in combining pattern-recognition modalities by feature- and kernel-selective Support Vector Machines. Proc. ICPR, 2008.



Conclusions

The aim of this study is to essentially improve the productivity of rare pronounced experts by way of,

first, accelerating the screening of mammographic images up to ten pictures per second, and,

second, immediately detecting the eventual potentials evoked in the expert's EEG by a target (cancer) image among a crowd of non-target ones before the expert becomes aware of this fact.

Acknowledgement

We would like to acknowledge support from the Russian Foundation for Basic Research, projects 18-07-01087, 17-07-00993, 17-07-00436, and from Project Dreams4Cars, grant No 731593, United Kingdom.

Thank you!

Questions?