



CZECH TECHNICAL UNIVERSITY IN PRAGUE
Faculty of Nuclear Sciences and Physical Engineering



English Title

Czech Title

Diploma thesis

Author: **Miroslav Kovář**

Supervisor: **M.Sc. M.A. Sebastián Basterrech, Ph.D.**

Academic year: 2018/2019

- Zadání práce -

- Zadání práce (zadní strana) -

Acknowledgment:

Some acknowledgment here.

Author's declaration:

I declare that this research project is entirely my own work and I have listed all the used sources in the bibliography.

Prague, August 8, 2018

Miroslav Kovář

Název práce:

Czech Title

Autor: Miroslav Kovář

Obor: Aplikace přírodních věd

Zaměření: Matematická informatika

Druh práce: Diplomová práce

Vedoucí práce: M.Sc. M.A. Sebastián Basterrech, Ph.D., Artificial Intelligence Center, FEE, CTU Prague

Abstrakt:

Klíčová slova:

Title:

English Title

Author: Miroslav Kovář

Abstract:

Key words:

Contents

1	Classical EEG signal analysis methods	13
1.1	EEG signal	13
1.2	Non-linear time series analysis	15
1.2.1	Attractor	15
1.2.2	Recurrence plot	18
1.3	Frequency domain	18
1.4	Deep learning	18
2	CNNs and CapsNets	19
2.1	CNNs	19
2.1.1	History	19
2.1.2	Description	20
2.1.3	Applications	21
2.2	CapsNets	21
3	Experiments	23
3.1	Dataset	23
3.2	Results	23

Introduction

Chapter 1

Classical EEG signal analysis methods

1.1 EEG signal

Nature, processes in the brain, way of measuring, limitations, complications

Electroencephalography (EEG) is a noninvasive method of measuring fluctuations of electric potentials near the skull caused by synchronized firing of neurons in the upper cortical layers. Electroencephalogram is a record of these fluctuations measured over a period of time. [14]

Although EEG has significantly lower spatial resolution in comparison with other diagnostic techniques such as functional magnetic resonance sampling (fMRI) and magnetoencephalography (MEG) [18] and enables measuring only neural activity near the cortical surface, as a depression diagnostic tool, it has numerous benefits. Importantly, its significantly lower costs [21] [5], high portability, and ease of operation imply increased availability to the patients [16]. Moreover, it is perfectly noninvasive, which means less complications such as claustrophobia or anxiety [13].

Although the science of EEG signal analysis as a diagnostic tool brings compelling clinical promise as a result of the aforementioned benefits, it also presents multiple technical and conceptual challenges.

Definition 1 ([15]). A series $\{X_t\}_{t \in \mathbb{Z}}$ is called **stationary**, if $\{X_t\}_{t \in \mathbb{Z}}$ for any set of times t_1, t_2, \dots, t_n and any $k \in \mathbb{N}$, $P[X_{t_1}, X_{t_2}, \dots, X_{t_n}] = P[X_{t_1+k}, X_{t_2+k}, \dots, X_{t_n+k}]$, i.e. the joint probability distribution of $\{X_t\}_{t \in \mathbb{Z}}$ is not a function of time. It is called **non-stationary**, if it is not stationary.

Definition 2 ([1]). A series $\{X_t\}_{t \in \mathbb{Z}}$ is called (noisy chaotic) **non-linear**, if it satisfies the relation

$$X_t = f(X_{t-1}) + \epsilon_t \quad (1.1)$$

for a general $f : \mathbb{R} \rightarrow \mathbb{R}$.

EEG signals are prone to be infected with *noise* due to imperfect isolation from surrounding environment. They are known to be *transient*, *non-Gaussian*, *non-stationary* and *nonlinear* [8] [19]. Since some patterns do not activate relative to a stimulus, a successful classifier must be able to detect a pattern *regardless of its starting time*, or find one. And finally, EEG records are relatively high dimensional - 16 electrodes sampling at 256 Hz result 4096 data points per second.

Moreover, due to the phenomenon of neural oscillations, patterns may appear in multiple frequency bands, from slow cortical potentials of δ -waves at 0.5-4 Hz, to high γ frequency band at 70-150 Hz.

Patterns of oscillatory activity in various frequency band have been linked to various mental states [3] [2] and diseases such as epilepsy [17], tremor [12], Parkinson's disease and depression [11]. Many of the diseases, including depression, share common oscillatory patterns known as thalamocortical dysrhythmia, characterized by decrease in normal resting-state α (8-12 Hz) activity slowing down to θ (4-8 Hz) frequencies, accompanied by increase in β and γ (25-50 Hz) activity. [20]



Figure 1.1: A comparison of stationary and non-stationary time series. (Courtesy: Protonk)

1.2 Non-linear time series analysis

Modern neuroscientific work suggests that the most plausible research target for explaining the brain dynamics are the assemblies of coupled and synchronously active neurons, and since majority of those assemblies are describable by non-linear differential equations, it is common to apply principles derived from nonlinear dynamics to characterize these neuronal systems.[8] In this section, we will describe some of those principles.

Attractors, Poincare plots, recurrence plots, Lyapunov coefficients, fractal dimension, Hurst exponent, etc.

Definition 3 ([7]). Let $\mathbf{x} \in \mathbb{R}^m$ be an $m \in \mathbb{N}$ dimensional state space vector dependent on time. A deterministic **dynamical system** is described by a set of m first-order differential equations

$$\frac{d}{dt}\mathbf{x}(t) = \mathbf{f}(\mathbf{x}(t)), \quad t \in \mathbb{R} \quad (1.2)$$

In other words, a dynamical system is a model that determines the evolution of a system only given by the initial state, and current state is a function of the previous state. Hence, a total description of a dynamical system is given by the initial state and a set of equations. A non-linear dynamical system is a system where the differential equations describing its dynamics are non-linear. Unlike in a linear system, changes in the initial state of a non-dynamical system are allowed to have a non-linear relationship to the state space trajectory of the system. [8]

1.2.1 Attractor

Definition 4 ([7]). A dynamical system is called *dissipative*, when it is the case that

$$E[\text{div}\mathbf{f}] < 1 \quad (1.3)$$

In other words, average space space volume (containing the initial state) is contracted as the system evolves.

Dissipative dynamical system have the property that a set of initial states (of positive measure) will converge to a subspace of the overall state space. This subspace is a geometrical object called an **attractor**. Example of four basic attractors can be seen on Figure 1.2.

Since most physiogenerated signals are chaotic, their analysis is concerned primarily with *chaotic* (strange) *attractors*. These attractors are complex, and exhibit what is known as *fractal geometry*. For our purposes, this means they can be characterized as having quantifiable self-similarity.¹ An example of a self-similar attractor is shown on Figure 1.3.

One possible approach to non-linear time series analysis consists of the following steps: [19]

1. reconstruction of the dynamics of given system from recorded data,
2. characterization of the reconstructed attractor,
3. checking validity of the results with surrogate data testing.

In the following text, we will describe multiple techniques of characterizing the properties of attractors (and hence ways of describing the dynamics of the dynamical system).

¹Cantor set being a canonical example of self-similarity.

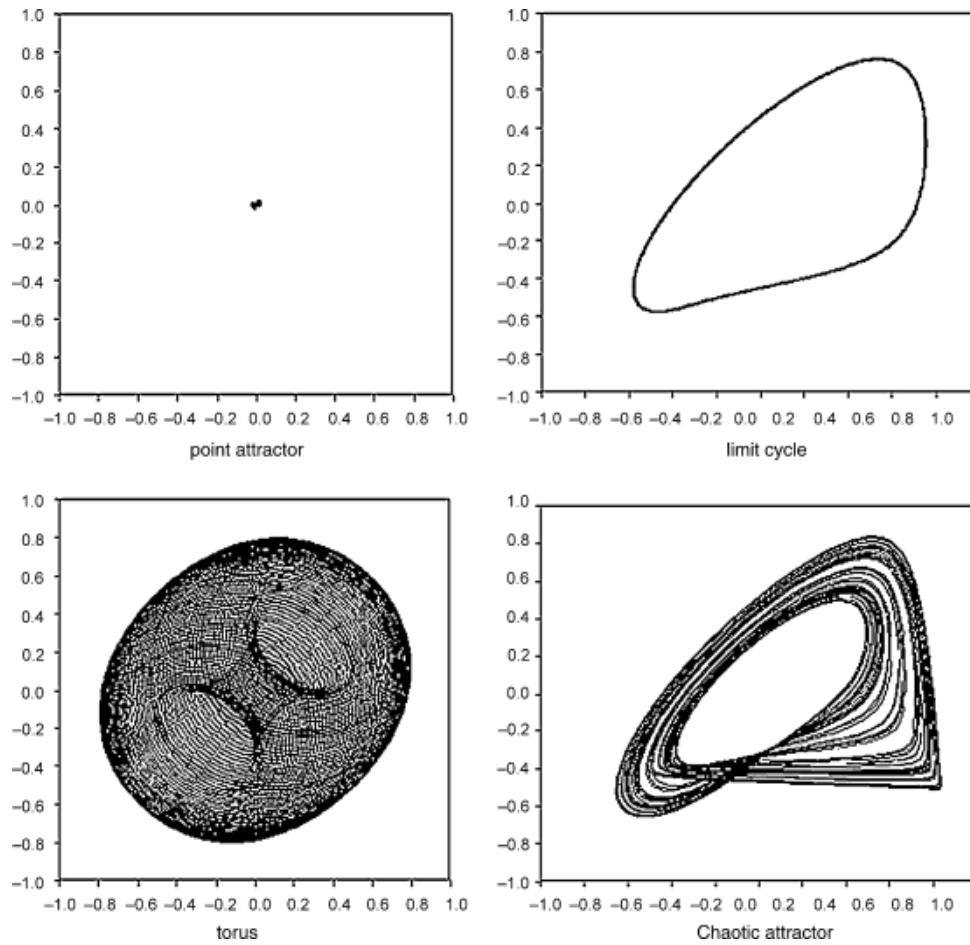


Figure 1.2: Visualization of four common attractor types (units are arbitrary). Left to right, top to bottom: *Point attractor* is the only type of attractor of linear deterministic dissipative systems. *Limit cycle* corresponds to a periodic system. *Torus attractor* corresponds to a quasi-periodic system, resulting (in this example) from a superposition of two periodic oscillations. *Chaotic (strange) attractor*, characterized by strong dependence on the original conditions. ([19])

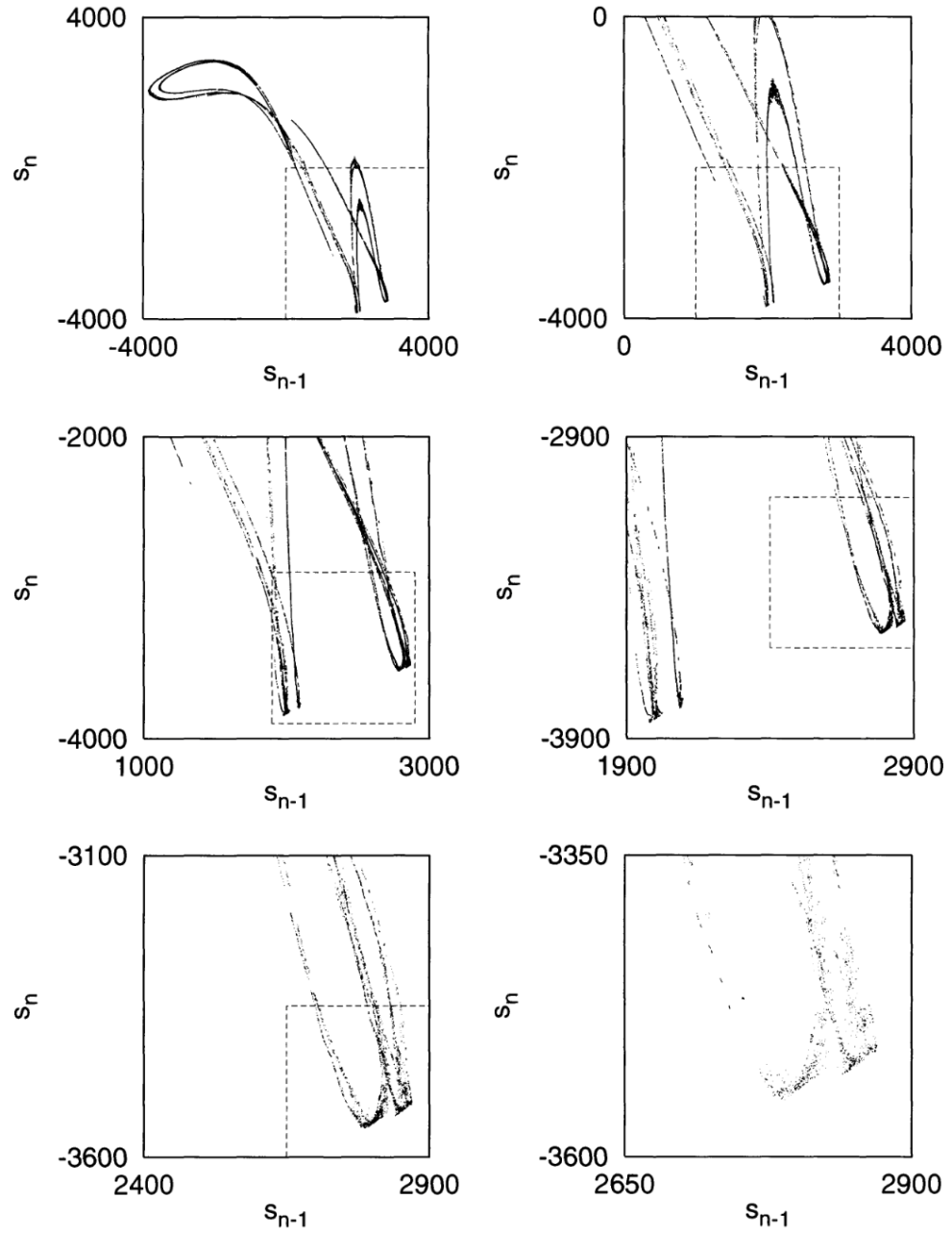


Figure 1.3: Noise-reduced visualization of successive enlargements of highly self-similar attractor. ([7])

1.2.2 Recurrence plot

Recurrence plot is a method of characterizing the attractor underlying a chaotic process.

1.3 Frequency domain

Wavelets, sliding window

1.4 Deep learning

Results in applying deep learning to EEG signal analysis

Chapter 2

CNNs and CapsNets

2.1 CNNs

2.1.1 History

The classical approach to image pattern recognition consists of the following stages:

preprocessing: suppressing unwanted distortions and noise, enhancement beneficial for further processing,

object segmentation: separating disparate objects from the background,

feature extraction: gathering relevant information about the properties of the objects, removing irrelevant variations,

classification: categorizing segmented objects based on obtained features into classes.

The preprocessing step may require additional assumptions about the data or further processing, which are potentially too restrictive or too broad. Getting around this limitation requires dealing with complications such as high dimensionality of the input (number of pixels) and desirability of invariance towards a number of allowable distortions and geometrical transformations.

Artificial neural networks in combination with gradient-based learning are one possible solution to the problem. By gradually optimizing a set of weights based on a training data set using a differentiable error function, they provide a framework for learning a suitable set of assumptions automatically from the data.

One of the oldest neural network architectures, fully connected multi-layer perceptron (FC-MLP), can be used for image pattern recognition. However, it has the following drawbacks:

parameter explosion: the number of parameters of such network is exponential in the number of layers, increasing the capacity of the network and therefore need for more data,

no invariance: no invariance even with respect to common geometrical transformation such as translation, rotation and scaling,

ignoring input topology: natural images exhibit strong local structure and high correlation between intensities of neighboring pixels, but FC-MLPs are unstructured - inputs can be presented in any order.

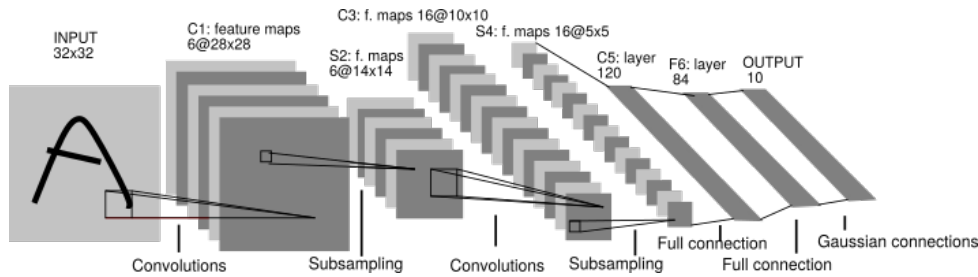


Figure 2.1: LeNet-5 architecture [10].

Although the main idea dates back 1980 with K. Fukushima's neocognitron [4], the back-propagation algorithm was not known at the time. The first convolutional architecture successfully applied on an image pattern recognition problem by attempting to solve the aforementioned problems, dubbed LeNet-5, was proposed in 1998 by Y. LeCun, L. Bottou, Y. Bengio and P. Haffner [9].

2.1.2 Description

Bearing resemblance to visual processing in biological organisms¹, LeNet-5 proposed the following design principles to enforce *shift, scale and distortion invariance*: [10]

local receptive fields: each neuron in a layer receives input from a small neighborhood in the previous layer,

shared weights: each layer is composed of neurons organized in planes within which each neuron have the same weight vector (feature map),

spatial subsampling: adding a subsampling layers, which reduce the resolution of the previous layer by averaging or taking the maximal value of neighboring pixels in the previous layer.

Local receptive fields enable the network to synthesize filters that produce strong response to elementary salient features in the early layers (such as lines, edges and corners in a visual input, and equivalence in other modalities), and then learn to combine them in the subsequent layers to produce higher-order feature detectors.

Shared weights principle exploits the fact that translation-invariant features have the property that one feature detector can be used accross the entire image. Since neural units in a layer with differing receptive fields possess the same feature map, the same feature detecting operation (convolution with feature map kernel followed by additive bias and a application of a non-linear function) is performed on differing parts of the image. A single convolutional layer is composed of multiple feature detecting planes.

Spatial subsampling is purposed to ensure scale and distortion invariance² by reducing the precision at which a feature is encoded in a feature map by reducing its resolution - when scale and distortion invariance is assumed, the exact location of a feature becomes less important and is allowed to exhibit slight positional variance. Although this simple solution performs well in many practical situations, it

¹As early as in 1968, D. H. Hubel and T.N. Wiesel discovered that some cells (called simple cells) in cat's primary visual cortex (V1) with small receptive fields (shared by neighboring neurons) are sensitive to straight lines and edges of light of particular orientation, and other cells (called complex cells) with larger receptive fields further in the visual cortex also respond to straight lines and edges, but with invariance to translation [6].

²Whether it achieves this goal has been famously doubted by Geoffrey Hinton: "The pooling operation used in convolutional neural networks is a big mistake and the fact that it works so well is a disaster." []

has multiple drawbacks, as mentioned in section 2.2 (relative representation of features is not encoded in the feature map representation)

2.1.3 Applications

Maybe mention an example of how LeNet-5 was improved on subsequently (AlexNet, ResNet, etc.)

2.2 CapsNets

Does it make sense trying them? I found a only a few successful implementations. Maybe it would be better to try those after we have some results already, because it seems risky - we might end up with nothing.

Chapter 3

Experiments

3.1 Dataset

Size of our dataset, conditions during trials, labels, etc.

3.2 Results

Conclusion

Bibliography

- [1] Peter J Bickel and Peter Bühlmann. What is a linear process? *Proceedings of the National Academy of Sciences*, 93(22):12128–12131, 1996.
- [2] György Buzsáki and Andreas Draguhn. Neuronal oscillations in cortical networks. *science*, 304(5679):1926–1929, 2004.
- [3] Ryan T Canolty, Erik Edwards, Sarang S Dalal, Maryam Soltani, Srikantan S Nagarajan, Heidi E Kirsch, Mitchel S Berger, Nicholas M Barbaro, and Robert T Knight. High gamma power is phase-locked to theta oscillations in human neocortex. *science*, 313(5793):1626–1628, 2006.
- [4] Kunihiro Fukushima and Sei Miyake. Neocognitron: A self-organizing neural network model for a mechanism of visual pattern recognition. In *Competition and cooperation in neural nets*, pages 267–285. Springer, 1982.
- [5] Matti Hämäläinen, Riitta Hari, Risto J Ilmoniemi, Jukka Knuutila, and Olli V Lounasmaa. Magnetoencephalography—theory, instrumentation, and applications to noninvasive studies of the working human brain. *Reviews of modern Physics*, 65(2):413, 1993.
- [6] D. H. Hubel and T. N. Wiesel. Receptive fields and functional architecture of monkey striate cortex. *The Journal of Physiology*, 1968.
- [7] Holger Kantz and Thomas Schreiber. *Nonlinear time series analysis*, volume 7. Cambridge university press, 2004.
- [8] Alexander Ya Kaplan, Andrew A Fingelkurts, Alexander A Fingelkurts, Sergei V Borisov, and Boris S Darkhovsky. Nonstationary nature of the brain activity as revealed by eeg/meg: methodological, practical and conceptual challenges. *Signal processing*, 85(11):2190–2212, 2005.
- [9] Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2323, 1998.
- [10] Yann LeCun, Patrick Haffner, Léon Bottou, and Yoshua Bengio. Object recognition with gradient-based learning. In *Shape, contour and grouping in computer vision*, pages 319–345. Springer, 1999.
- [11] Rodolfo R Llinás, Urs Ribary, Daniel Jeanmonod, Eugene Kronberg, and Partha P Mitra. Thalamocortical dysrhythmia: a neurological and neuropsychiatric syndrome characterized by magnetoencephalography. *Proceedings of the National Academy of Sciences*, 96(26):15222–15227, 1999.
- [12] JH McAuley and CD Marsden. Physiological and pathological tremors and rhythmic central motor control. *Brain*, 123(8):1545–1567, 2000.

- [13] Kieran J Murphy and James A Brunberg. Adult claustrophobia, anxiety and sedation in mri. *Magnetic resonance imaging*, 15(1):51–54, 1997.
- [14] Paul L Nunez, Ramesh Srinivasan, et al. *Electric fields of the brain: the neurophysics of EEG*. Oxford University Press, USA, 2006.
- [15] Maurice Bertram Priestley. Non-linear and non-stationary time series analysis. 1988.
- [16] Teal L Schultz. Technical tips: Mri compatible eeg electrodes: advantages, disadvantages, and financial feasibility in a clinical setting. *The Neurodiagnostic Journal*, 52(1):69–81, 2012.
- [17] Vladimir Shusterman and William C Troy. From baseline to epileptiform activity: a path to synchronized rhythmicity in large-scale neural networks. *Physical Review E*, 77(6):061911, 2008.
- [18] Ramesh Srinivasan. Methods to improve the spatial resolution of eeg. *International Journal of Bioelectromagnetism*, 1(1):102–111, 1999.
- [19] C. J. Stam. Nonlinear dynamical analysis of EEG and MEG: Review of an emerging field. *Clinical Neurophysiology*, 116(10):2266–2301, 2005.
- [20] Sven Vanneste, Jae-Jin Song, and Dirk De Ridder. Thalamocortical dysrhythmia detected by machine learning. *Nature communications*, 9(1):1103, 2018.
- [21] Paul M Vespa, Val Nenov, and Marc R Nuwer. Continuous eeg monitoring in the intensive care unit: early findings and clinical efficacy. *Journal of Clinical Neurophysiology*, 16(1):1–13, 1999.