

Undergraduate Thesis

---

# HiC matrix correctign using ICE

---

Felix Karg

Examiner: Prof. Dr. Backofen

Advisers: Dr. Mehmet Tekman, Joachim Wolff

Albert-Ludwigs-University Freiburg

Faculty of Engineering

Department of Computer Science

Chair for Bioinformatics

July 10<sup>th</sup>, 2019

**Writing Period**

10. 04. 2019 – 10. 07. 2019

**Examiner**

Prof. Dr. Backofen

**Advisers**

Dr. Mehmet Tekman, Joachim Wolff

Bachelor Thesis

---

# HiC matrix correctign using ICE

---

Felix Karg

Gutachter: Prof. Dr. Backofen

Betreuer: Dr. Mehmet Tekman, Joachim Wolff

Albert-Ludwigs-Universität Freiburg

Technische Fakultät

Institut für Informatik

Lehrstuhl für Bioinformatik

10. Juli 2019

**Bearbeitungszeit**

10. 04. 2016 – 10. 07. 2016

**Gutachter**

Prof. Dr. Backofen

**Betreuer**

Dr. Mehmet Tekman, Joachim Wolff

# Declaration

I hereby declare, that I am the sole author and composer of my thesis and that no other sources or learning aids, other than those listed, have been used. Furthermore, I declare that I have acknowledged the work of others by providing detailed references of said work.

I hereby also declare, that my Thesis has not been prepared for another examination or assignment, either wholly or excerpts thereof.

---

Place, Date

---

Signature



# Abstract

foo bar





# Zusammenfassung

German version is only needed for an undergraduate thesis.



# Inhaltsverzeichnis

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Template Structure . . . . .	1
1.2	setup.tex . . . . .	2
1.3	Advice . . . . .	4
<b>2</b>	<b>Related Work</b>	<b>7</b>
<b>3</b>	<b>Background</b>	<b>9</b>
<b>4</b>	<b>Approach</b>	<b>11</b>
4.1	Problem Definition . . . . .	11
4.2	First Part of the Approach . . . . .	11
4.3	N-th Part of the Approach . . . . .	11
<b>5</b>	<b>Experiments</b>	<b>13</b>
<b>6</b>	<b>Conclusion</b>	<b>15</b>
<b>7</b>	<b>Acknowledgments</b>	<b>17</b>
	<b>Bibliography</b>	<b>21</b>



# Abbildungsverzeichnis

1	Tikz Example . . . . .	10
2	Caption that appears in the figlist . . . . .	14



# Tabellenverzeichnis

1	Table caption . . . . .	13
---	-------------------------	----





# List of Algorithms

1	Stochastic Gradient Descent: Neural Network . . . . .	10
---	---	----



# 1 Introduction

This is a template for an undergraduate or master's thesis. The first sections are concerned with the template itself. If this is your first thesis, consider reading Section 1.3.

Of course, the structure of this thesis is only an example. Discuss with your adviser what structure fits best for your thesis.

## 1.1 Template Structure

- To compile the document either run the makefile or run your compiler on the file ‘thesis\_main.tex’. The included makefile requires latexmk which automatically runs bibtex and recompiles your thesis as often as needed. Also it automatically places all output files (aux, bbl, ...) in the folder ‘out’. As the pdf also goes in there, the makefile copies the pdf file to the parent folder. There is also a makefile in the chapters folder, to ensure you can also compile from this directory.
- The file ‘setup.tex’ includes the packages and defines commands. For more details see Section 1.2.
- Each chapter goes into a separate document, the files can be found in the folder chapters.

- The bib folder contains the .bib files, I'd suggest to create multiple bib files for different topics. If you add some or rename the existing ones, don't forget to also change this in thesis\_main.tex. You can then cite as usual [1, 2, 3].
- The template is written in a way that eases the switch from scrbook to book class. So if you're not a fan of KOMA you can just replace the documentclass in the main file. The only thing that needs to be changed in setup.tex is the caption styling, see the comments there.

## 1.2 setup.tex

Edit setup.tex according to your needs. The file contains two sections, one for package includes, and one for defining commands. At the end of the includes and commands there is a section that can safely be removed if you don't need algorithms or tikz. Also don't forget to adapt the pdf hypersetup!!

setup.tex defines:

- some new commands for remembering to do stuff:
  - `\todo{Do this!}`: **(TODO: Do this!)**
  - `\extend{Write more when new results are out!}`:  
**(EXTEND: Write more when new results are out!)**
  - `\draft{Hacky text!}`: **(DRAFT: Hacky text!)**
- some commands for referencing, 'in `\chapref{chap:introduction}`' produces 'in Chapter 1'
  - `\chapref{}`

- `\secref{sec:XY}`
  - `\eqref{}`
  - `\figref{}`
  - `\tabref{}`
  - the colors of the Uni’s corporate design, accessible with `{\color{UniX} Colored Text}`
    - UniBlue
    - UniRed
    - UniGrey
  - a command for naming matrices `\mat{G}`, **G**, and naming vectors `\vec{a}`, **a**. This overwrites the default behavior of having an arrow over vectors, sticking to the naming conventions normal font for scalars, bold-lowercase for vectors, and bold-uppercase for matrices.
  - named equations:
- ```

\begin{align}
d(a,b) &= d(b,a) \\
\end{align}
\eqname{symmetry}

```

$$d(a, b) = d(b, a) \tag{1}$$

symmetry

## 1.3 Advice

This section gives some advice how to write a thesis ranging from writing style to formatting. To be sure, ask your advisor about his/her preferences.

For a more complete list we recommend to read Donald Knuth's paper on mathematical writing. (At least the first paragraph). [http://jmlr.csail.mit.edu/reviewing-papers/knuth\\_mathematical\\_writing.pdf](http://jmlr.csail.mit.edu/reviewing-papers/knuth_mathematical_writing.pdf)

- If you use formulae pay close attention to be consistent throughout the thesis!
- Usually in a thesis you don't write 'In [24] the data is..'. You have more space than a paper has, so write 'AuthorXY et al. prepare the data... [24]'. Also pay attention to the placement: The citation is at the end of the sentence before the full stop with a no-break space. ... last word~\cite{XY}.
- Pay attention to comma usage, there is a big difference between English and German. '...the fact that bla...' etc.
- Do not write 'don't ', 'can't' etc. Write 'do not', 'can not'.
- If an equation is at the end of a sentence, add a full stop. If it's not the end, add a comma:  $a = b + c$  (1),
- Avoid footnotes if possible.
- Use ‘ ’ for citing, not " ".
- It's important to look for spelling mistakes in your thesis. There are also tools like aspell that can help you find such mistakes. This is never an excuse not to properly read your thesis again, but it can help. You can find an introduction under <https://git.fachschaft.tf/fachschaft/aspell>.

- If have things like a graph or any other drawings consider using tikz, if you need function graphs or diagrams consider using pgfplots. This has the advantage that the style will be more consistent (same font, formatting options etc.) than when you use some external program.
- Discuss with your advisor whether to use passive voice or not. In most computer science papers passive voice is avoided. It's harder to read, more likely to produce errors, and most of the times less precise. Of course there are situations where the passive voice fits but in scientific papers they are rare. Compare the sentence: 'We created the wheel to solve this.' to 'The wheel was created to solve this', you don't know who did it, making it harder to understand what is your contribution and what is not.
- In tables avoid vertical lines, keep them clean and neat. See 1 for an example. More details can be found in the 'Small Guide to Making Nice Tables' <https://www.inf.ethz.ch/personal/markusp/teaching/guides/guide-tables.pdf>





## 2 Related Work

Give a brief overview of the work relevant for your thesis.



## 3 Background

Explain the math and notation.



**Abbildung 1:** Use tikz to draw nice graphs!

---

**Algorithm 1** Stochastic Gradient Descent: Neural Network

---

Create a mini batch of  $m$  samples  $\mathbf{x}_0 \dots \mathbf{x}_{m-1}$

**foreach** sample  $\mathbf{x}$  **do**

$\mathbf{a}^{\mathbf{x},0} \leftarrow \mathbf{x}$

▷ Set input activation

**foreach** Layer  $l \in \{1 \dots L-1\}$  **do**

▷ Forward pass

$\mathbf{z}^{\mathbf{x},l} \leftarrow \mathbf{W}^l \mathbf{a}^{\mathbf{x},l-1} + \mathbf{b}^l$

$\mathbf{a}^{\mathbf{x},l} \leftarrow \varphi(\mathbf{z}^{\mathbf{x},l})$

**end for**

$\delta^{\mathbf{x},L} \leftarrow \nabla_{\mathbf{a}} C_{\mathbf{x}} \odot \varphi'(\mathbf{z}^{\mathbf{x},L})$

▷ Compute error

**foreach** Layer  $l \in L-1, L-2 \dots 2$  **do**

▷ Backpropagate error

$\delta^{\mathbf{x},l} \leftarrow ((\mathbf{W}^{l+1})^T \delta^{\mathbf{x},l+1}) \odot \varphi'(\mathbf{z}^{\mathbf{x},l})$

**end for**

**end for**

**foreach**  $l \in L, L-1 \dots 2$  **do**

▷

▷ Gradient descent

$\mathbf{W}^l \leftarrow \mathbf{W}^l - \frac{\eta}{m} \sum_{\mathbf{x}} \delta^{\mathbf{x},l} (\mathbf{a}^{\mathbf{x},l-1})^T$

$\mathbf{b}^l \leftarrow \mathbf{b}^l - \frac{\eta}{m} \sum_{\mathbf{x}} \delta^{\mathbf{x},l}$

**end for**

---

## 4 Approach

The approach usually starts with the problem definition and continues with what you have done. Try to give an intuition first and describe everything with words and then be more formal like ‘Let  $g$  be ...’.

### 4.1 Problem Definition

Start with a very short motivation why this is important. Then, as stated above, describe the problem with words before getting formal.

### 4.2 First Part of the Approach

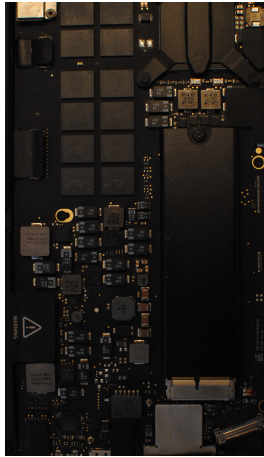
### 4.3 N-th Part of the Approach



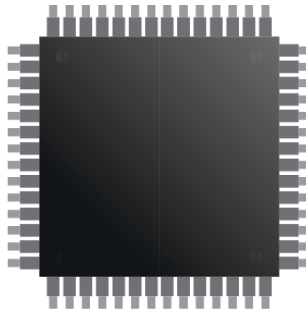
## 5 Experiments

| Type | Accuracy         |
|------|------------------|
| A    | $82.47 \pm 3.21$ |
| B    | $78.47 \pm 2.43$ |
| C    | $84.30 \pm 2.35$ |
| D    | $86.81 \pm 3.01$ |

**Tabelle 1: Table caption.** foo bar...



(a) Some cool graphic



(b) Some cool related graphic

**Abbildung 2: Caption that appears under the fig** Lorem ipsum dolor sit amet, consectetur adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetur id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.



## 6 Conclusion



## 7 Acknowledgments

First and foremost, I would like to thank...

- advisers
- examiner
- person1 for the dataset
- person2 for the great suggestion
- proofreaders



## ToDo Counters

To Dos: 1;   1

Parts to extend: 1;   1

Draft parts: 1;   1



# Literaturverzeichnis

- [1] D. Kingma and J. Ba, “Adam: A method for stochastic optimization,” *arXiv preprint arXiv:1412.6980*, 2014.
- [2] J. Bromley, J. W. Bentz, L. Bottou, I. Guyon, Y. LeCun, C. Moore, E. Säckinger, and R. Shah, “Signature verification using a “siamese” time delay neural network,” *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 7, no. 04, pp. 669–688, 1993.
- [3] M. Muja and D. G. Lowe, “Fast approximate nearest neighbors with automatic algorithm configuration.,” *VISAPP (1)*, vol. 2, no. 331-340, p. 2, 2009.

