Stellenbosch
UNIVERSITY
IYUNIVESITHI
UNIVERSITEIT

# Quadcopter for Autonomous Underground Mine Mapping And Exploration

by

Nico Epler

Thesis presented in fulfilment of the requirements for the degree of Master of Engineering (Electronic) in the Faculty of Engineering at Stellenbosch University

Supervisor: Dr Callen Fisher

September 2025

# Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: September 2025 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .

# Abstract

The English abstract.

# Opsomming

Die Afrikaanse uittreksel.

# Acknowledgements

I would like to thank my dog, Muffin. I also would like to thank the inventor of the incubator; without him/her, I would not be here. Finally, I would like to thank Dr Herman Kamper for this amazing report template.

# Contents

# Nomenclature

### Variables and functions

| | |
|---|---|
| $p(x)$ | Probability density function with respect to variable $x$. |
| $P(A)$ | Probability of event $A$ occurring. |
| $\varepsilon$ | The Bayes error. |
| $\varepsilon_u$ | The Bhattacharyya bound. |
| $B$ | The Bhattacharyya distance. |
| $s$ | An HMM state. A subscript is used to refer to a particular state, e.g. $s_i$ refers to the $i^{\text{th}}$ state of an HMM. |
| $\mathbf{S}$ | A set of HMM states. |
| $\mathbf{F}$ | A set of frames. |
| $\mathbf{o}_f$ | Observation (feature) vector associated with frame $f$. |
| $\gamma_s(\mathbf{o}_f)$ | A posteriori probability of the observation vector $\mathbf{o}_f$ being generated by HMM state $s$. |
| $\mu$ | Statistical mean vector. |
| $\Sigma$ | Statistical covariance matrix. |
| $L(\mathbf{S})$ | Log likelihood of the set of HMM states $\mathbf{S}$ generating the training set observation vectors assigned to the states in that set. |
| $\mathcal{N}(\mathbf{x}\|\mu, \Sigma)$ | Multivariate Gaussian PDF with mean $\mu$ and covariance matrix $\Sigma$. |
| $a_{ij}$ | The probability of a transition from HMM state $s_i$ to state $s_j$. |
| $N$ | Total number of frames or number of tokens, depending on the context. |
| $D$ | Number of deletion errors. |
| $I$ | Number of insertion errors. |
| $S$ | Number of substitution errors. |

## Acronyms and abbreviations

| | |
|---|---|
| AE | Afrikaans English |
| AID | accent identification |
| ASR | automatic speech recognition |
| AST | African Speech Technology |
| CE | Cape Flats English |
| DCD | dialect-context-dependent |
| DNN | deep neural network |
| G2P | grapheme-to-phoneme |
| GMM | Gaussian mixture model |
| HMM | hidden Markov model |
| HTK | Hidden Markov Model Toolkit |
| IE | Indian South African English |
| IPA | International Phonetic Alphabet |
| LM | language model |
| LMS | language model scaling factor |
| MFCC | Mel-frequency cepstral coefficient |
| MLLR | maximum likelihood linear regression |
| OOV | out-of-vocabulary |
| PD | pronunciation dictionary |
| PDF | probability density function |
| SAE | South African English |
| SAMPA | Speech Assessment Methods Phonetic Alphabet |

# Chapter 1

# Introduction

The last few years have seen great advances in speech recognition. Much of this progress is due to the resurgence of neural networks; most speech systems now rely on deep neural networks (DNNs) with millions of parameters [1, 2]. However, as the complexity of these models has grown, so has their reliance on labelled training data. Currently, system development requires large corpora of transcribed speech audio data, texts for language modelling, and pronunciation dictionaries. Despite speech applications becoming available in more languages, it is hard to imagine that resource collection at the required scale would be possible for all 7000 languages spoken in the world today.

I really like apples.

## 1.1. Section heading

This is some section with two table in it: Table 6.1 and Table 6.2.

**Table 1.1:** Performance of the unconstrained segmental Bayesian model on TIDigits1 over iterations in which the reference set is refined.

| Metric | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| WER (%) | 35.4 | 23.5 | 21.5 | 21.2 | 22.9 |
| Average cluster purity (%) | 86.5 | 89.7 | 89.2 | 88.5 | 86.6 |
| Word boundary $F$-score (%) | 70.6 | 72.2 | 71.8 | 70.9 | 69.4 |
| Clusters covering 90% of data | 20 | 13 | 13 | 13 | 13 |

**Table 1.2:** A table with an example of using multiple columns.

| | Accuracy (%) | | |
|---|---|---|---|
| Model | Intermediate | Output | Bitrate |
| Baseline | 27.5 | 26.4 | 116 |
| VQ-VAE | 26.0 | 22.1 | 190 |
| CatVAE | 28.7 | 24.3 | 215 |

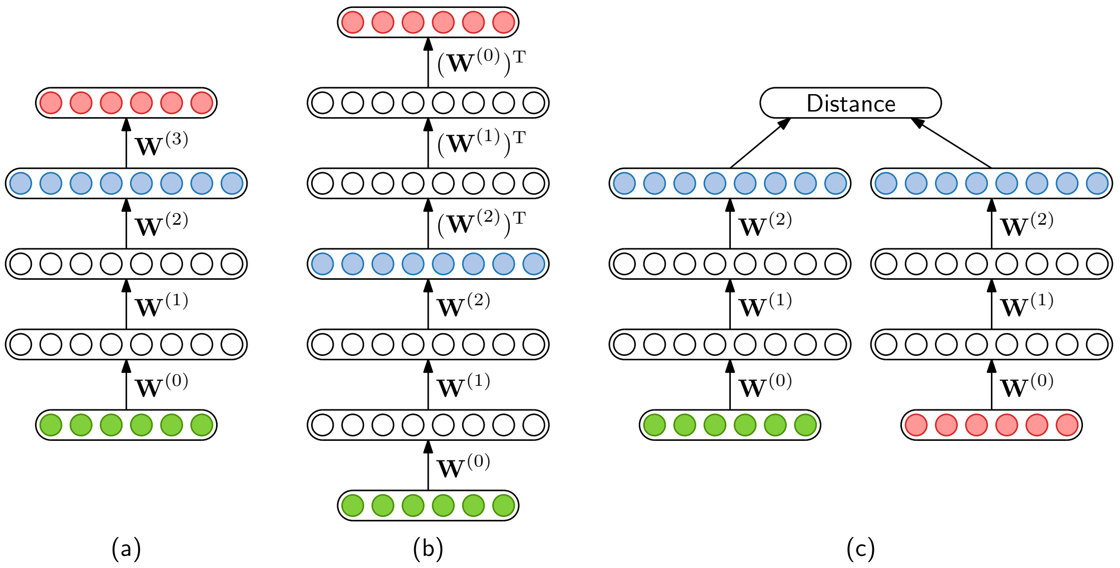
$\mathbf{W}^{(3)}$
$\mathbf{W}^{(2)}$
$\mathbf{W}^{(1)}$
$\mathbf{W}^{(0)}$
$(\mathbf{W}^{(0)})^\mathrm{T}$
$(\mathbf{W}^{(1)})^\mathrm{T}$
$(\mathbf{W}^{(2)})^\mathrm{T}$
Distance
(a)
(b)
(c)

**Figure 1.1:** (a) The cAE as used in this chapter. The encoding layer (blue) is chosen based on performance on a development set. (b) The cAE with symmetrical tied weights. The encoding from the middle layer (blue) is always used. (c) The siamese DNN. The cosine distance between aligned frames (green and red) is either minimized or maximized depending on whether the frames belong to the same (discovered) word or not. A cAE can be seen as a type of DNN [1].

This is a new page, showing what the page headings looks like, and showing how to refer to a figure like Figure 6.1.

The following is an example of an equation:

$$P(\mathbf{z}|\boldsymbol{\alpha}) = \int_{\boldsymbol{\pi}} P(\mathbf{z}|\boldsymbol{\pi})\, p(\boldsymbol{\pi}|\boldsymbol{\alpha})\, \mathrm{d}\boldsymbol{\pi} = \int_{\boldsymbol{\pi}} \prod_{k=1}^{K} \pi_k^{N_k} \frac{1}{B(\boldsymbol{\alpha})} \prod_{k=1}^{K} \pi_k^{\alpha_k - 1}\, \mathrm{d}\boldsymbol{\pi} \tag{1.1}$$

which you can subsequently refer to as (6.1) or Equation 6.1. But make sure to consistently use the one or the other (and not mix the two ways of referring to equations).

## 1.2. Contributions

The following papers resulted from the work presented here:

*Conference paper 1:*
L. Skywalker, D. Vadar, and O. W. Kenobi, "A comparison between father-son and master-apprentice relationships in space conflict situations," in *Proceedings of the International Conference on Action, Space and Star Politics (ICASSP)*, 2020.

*Journal paper 1:*
L. Skywalker, and L. Organa, "Identifying weaknesses in large evil corporations," *IEEE Transactions on the Exploration of the Outer Rim*, vol. 21, pp. 154–174, 2021.

# Chapter 2

# Introduction

The last few years have seen great advances in speech recognition. Much of this progress is due to the resurgence of neural networks; most speech systems now rely on deep neural networks (DNNs) with millions of parameters [1, 2]. However, as the complexity of these models has grown, so has their reliance on labelled training data. Currently, system development requires large corpora of transcribed speech audio data, texts for language modelling, and pronunciation dictionaries. Despite speech applications becoming available in more languages, it is hard to imagine that resource collection at the required scale would be possible for all 7000 languages spoken in the world today.

I really like apples.

## 2.1. Section heading

This is some section with two table in it: Table 6.1 and Table 6.2.

**Table 2.1:** Performance of the unconstrained segmental Bayesian model on TIDigits1 over iterations in which the reference set is refined.

| Metric | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| WER (%) | 35.4 | 23.5 | 21.5 | 21.2 | 22.9 |
| Average cluster purity (%) | 86.5 | 89.7 | 89.2 | 88.5 | 86.6 |
| Word boundary $F$-score (%) | 70.6 | 72.2 | 71.8 | 70.9 | 69.4 |
| Clusters covering 90% of data | 20 | 13 | 13 | 13 | 13 |

**Table 2.2:** A table with an example of using multiple columns.

| | Accuracy (%) | | |
|---|---|---|---|
| Model | Intermediate | Output | Bitrate |
| Baseline | 27.5 | 26.4 | 116 |
| VQ-VAE | 26.0 | 22.1 | 190 |
| CatVAE | 28.7 | 24.3 | 215 |

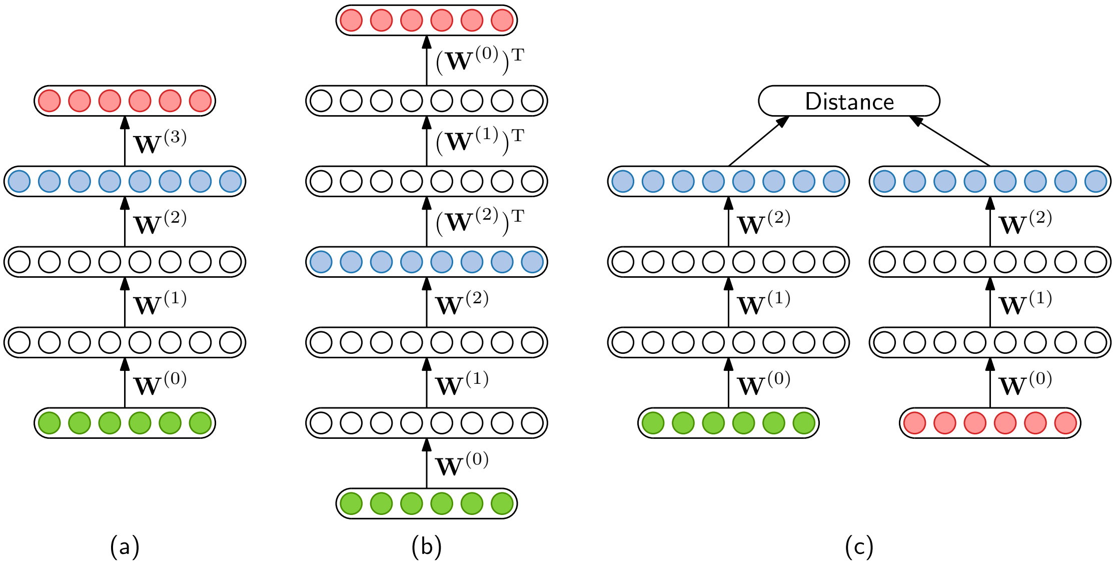
$\mathbf{W}^{(3)}$
$\mathbf{W}^{(2)}$
$\mathbf{W}^{(1)}$
$\mathbf{W}^{(0)}$
$(\mathbf{W}^{(0)})^{\mathrm{T}}$
$(\mathbf{W}^{(1)})^{\mathrm{T}}$
$(\mathbf{W}^{(2)})^{\mathrm{T}}$
Distance
(a)
(b)
(c)

**Figure 2.1:** (a) The cAE as used in this chapter. The encoding layer (blue) is chosen based on performance on a development set. (b) The cAE with symmetrical tied weights. The encoding from the middle layer (blue) is always used. (c) The siamese DNN. The cosine distance between aligned frames (green and red) is either minimized or maximized depending on whether the frames belong to the same (discovered) word or not. A cAE can be seen as a type of DNN [1].

This is a new page, showing what the page headings looks like, and showing how to refer to a figure like Figure 6.1.

The following is an example of an equation:

$$P(\mathbf{z}|\boldsymbol{\alpha}) = \int_{\boldsymbol{\pi}} P(\mathbf{z}|\boldsymbol{\pi})\, p(\boldsymbol{\pi}|\boldsymbol{\alpha})\, \mathrm{d}\boldsymbol{\pi} = \int_{\boldsymbol{\pi}} \prod_{k=1}^{K} \pi_k^{N_k} \frac{1}{B(\boldsymbol{\alpha})} \prod_{k=1}^{K} \pi_k^{\alpha_k - 1}\, \mathrm{d}\boldsymbol{\pi} \tag{2.1}$$

which you can subsequently refer to as (6.1) or Equation 6.1. But make sure to consistently use the one or the other (and not mix the two ways of referring to equations).

## 2.2. Contributions

The following papers resulted from the work presented here:

*Conference paper 1:*
L. Skywalker, D. Vadar, and O. W. Kenobi, “A comparison between father-son and master-apprentice relationships in space conflict situations,” in *Proceedings of the International Conference on Action, Space and Star Politics (ICASSP)*, 2020.

*Journal paper 1:*
L. Skywalker, and L. Organa, “Identifying weaknesses in large evil corporations,” *IEEE Transactions on the Exploration of the Outer Rim*, vol. 21, pp. 154–174, 2021.

# Chapter 3

# Introduction

The last few years have seen great advances in speech recognition. Much of this progress is due to the resurgence of neural networks; most speech systems now rely on deep neural networks (DNNs) with millions of parameters [1, 2]. However, as the complexity of these models has grown, so has their reliance on labelled training data. Currently, system development requires large corpora of transcribed speech audio data, texts for language modelling, and pronunciation dictionaries. Despite speech applications becoming available in more languages, it is hard to imagine that resource collection at the required scale would be possible for all 7000 languages spoken in the world today.

I really like apples.

## 3.1. Section heading

This is some section with two table in it: Table 6.1 and Table 6.2.

**Table 3.1:** Performance of the unconstrained segmental Bayesian model on TIDigits1 over iterations in which the reference set is refined.

| Metric | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| WER (%) | 35.4 | 23.5 | 21.5 | 21.2 | 22.9 |
| Average cluster purity (%) | 86.5 | 89.7 | 89.2 | 88.5 | 86.6 |
| Word boundary $F$-score (%) | 70.6 | 72.2 | 71.8 | 70.9 | 69.4 |
| Clusters covering 90% of data | 20 | 13 | 13 | 13 | 13 |

**Table 3.2:** A table with an example of using multiple columns.

| | Accuracy (%) | | |
|---|---|---|---|
| Model | Intermediate | Output | Bitrate |
| Baseline | 27.5 | 26.4 | 116 |
| VQ-VAE | 26.0 | 22.1 | 190 |
| CatVAE | 28.7 | 24.3 | 215 |

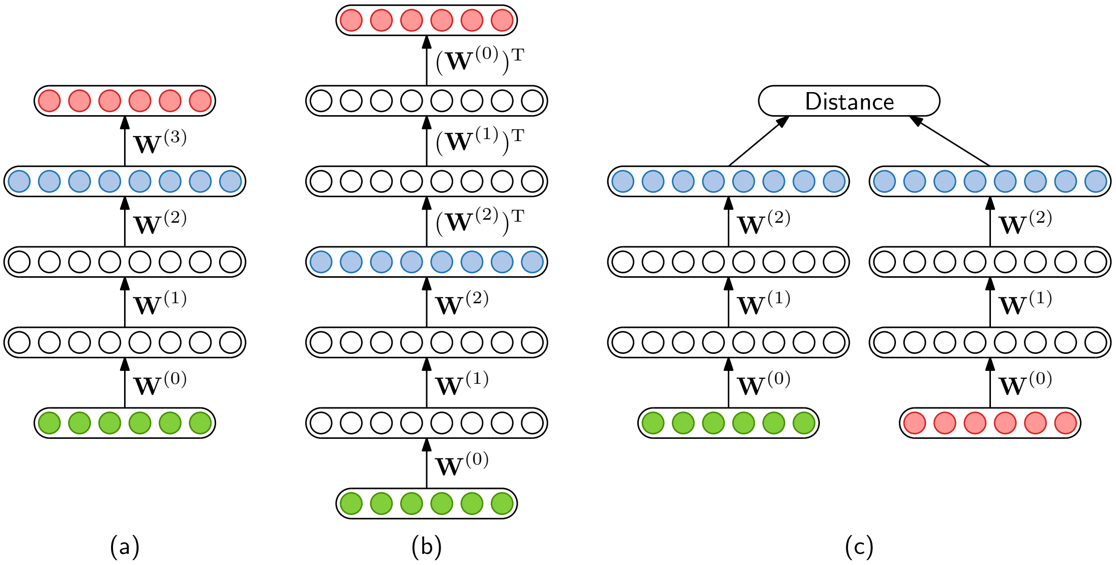

**Figure 3.1:** (a) The cAE as used in this chapter. The encoding layer (blue) is chosen based on performance on a development set. (b) The cAE with symmetrical tied weights. The encoding from the middle layer (blue) is always used. (c) The siamese DNN. The cosine distance between aligned frames (green and red) is either minimized or maximized depending on whether the frames belong to the same (discovered) word or not. A cAE can be seen as a type of DNN [1].

This is a new page, showing what the page headings looks like, and showing how to refer to a figure like Figure 6.1.

The following is an example of an equation:

$$P(\mathbf{z}|\boldsymbol{\alpha}) = \int_{\boldsymbol{\pi}} P(\mathbf{z}|\boldsymbol{\pi})\, p(\boldsymbol{\pi}|\boldsymbol{\alpha})\, \mathrm{d}\boldsymbol{\pi} = \int_{\boldsymbol{\pi}} \prod_{k=1}^{K} \pi_k^{N_k} \frac{1}{B(\boldsymbol{\alpha})} \prod_{k=1}^{K} \pi_k^{\alpha_k - 1}\, \mathrm{d}\boldsymbol{\pi} \tag{3.1}$$

which you can subsequently refer to as (6.1) or Equation 6.1. But make sure to consistently use the one or the other (and not mix the two ways of referring to equations).

## 3.2. Contributions

The following papers resulted from the work presented here:

*Conference paper 1:*
L. Skywalker, D. Vadar, and O. W. Kenobi, "A comparison between father-son and master-apprentice relationships in space conflict situations," in *Proceedings of the International Conference on Action, Space and Star Politics (ICASSP)*, 2020.

*Journal paper 1:*
L. Skywalker, and L. Organa, "Identifying weaknesses in large evil corporations," *IEEE Transactions on the Exploration of the Outer Rim*, vol. 21, pp. 154–174, 2021.

# Chapter 4

# Introduction

The last few years have seen great advances in speech recognition. Much of this progress is due to the resurgence of neural networks; most speech systems now rely on deep neural networks (DNNs) with millions of parameters [1, 2]. However, as the complexity of these models has grown, so has their reliance on labelled training data. Currently, system development requires large corpora of transcribed speech audio data, texts for language modelling, and pronunciation dictionaries. Despite speech applications becoming available in more languages, it is hard to imagine that resource collection at the required scale would be possible for all 7000 languages spoken in the world today.

I really like apples.

## 4.1. Section heading

This is some section with two table in it: Table 6.1 and Table 6.2.

**Table 4.1:** Performance of the unconstrained segmental Bayesian model on TIDigits1 over iterations in which the reference set is refined.

| Metric | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| WER (%) | 35.4 | 23.5 | 21.5 | 21.2 | 22.9 |
| Average cluster purity (%) | 86.5 | 89.7 | 89.2 | 88.5 | 86.6 |
| Word boundary $F$-score (%) | 70.6 | 72.2 | 71.8 | 70.9 | 69.4 |
| Clusters covering 90% of data | 20 | 13 | 13 | 13 | 13 |

**Table 4.2:** A table with an example of using multiple columns.

| | Accuracy (%) | | |
|---|---|---|---|
| Model | Intermediate | Output | Bitrate |
| Baseline | 27.5 | 26.4 | 116 |
| VQ-VAE | 26.0 | 22.1 | 190 |
| CatVAE | 28.7 | 24.3 | 215 |

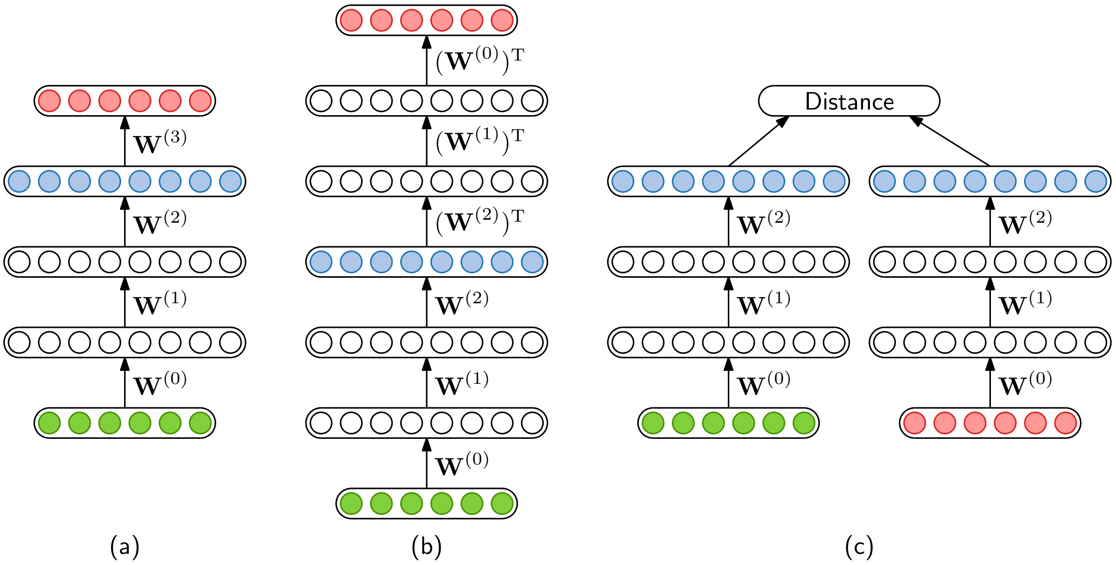

**Figure 4.1:** (a) The cAE as used in this chapter. The encoding layer (blue) is chosen based on performance on a development set. (b) The cAE with symmetrical tied weights. The encoding from the middle layer (blue) is always used. (c) The siamese DNN. The cosine distance between aligned frames (green and red) is either minimized or maximized depending on whether the frames belong to the same (discovered) word or not. A cAE can be seen as a type of DNN [1].

This is a new page, showing what the page headings looks like, and showing how to refer to a figure like Figure 6.1.

The following is an example of an equation:

$$P(\mathbf{z}|\boldsymbol{\alpha}) = \int_{\boldsymbol{\pi}} P(\mathbf{z}|\boldsymbol{\pi})\, p(\boldsymbol{\pi}|\boldsymbol{\alpha})\, \mathrm{d}\boldsymbol{\pi} = \int_{\boldsymbol{\pi}} \prod_{k=1}^{K} \pi_k^{N_k} \frac{1}{B(\boldsymbol{\alpha})} \prod_{k=1}^{K} \pi_k^{\alpha_k - 1}\, \mathrm{d}\boldsymbol{\pi} \tag{4.1}$$

which you can subsequently refer to as (6.1) or Equation 6.1. But make sure to consistently use the one or the other (and not mix the two ways of referring to equations).

## 4.2. Contributions

The following papers resulted from the work presented here:

*Conference paper 1:*
L. Skywalker, D. Vadar, and O. W. Kenobi, "A comparison between father-son and master-apprentice relationships in space conflict situations," in *Proceedings of the International Conference on Action, Space and Star Politics (ICASSP)*, 2020.

*Journal paper 1:*
L. Skywalker, and L. Organa, "Identifying weaknesses in large evil corporations," *IEEE Transactions on the Exploration of the Outer Rim*, vol. 21, pp. 154–174, 2021.

# Chapter 5

# Introduction

The last few years have seen great advances in speech recognition. Much of this progress is due to the resurgence of neural networks; most speech systems now rely on deep neural networks (DNNs) with millions of parameters [1, 2]. However, as the complexity of these models has grown, so has their reliance on labelled training data. Currently, system development requires large corpora of transcribed speech audio data, texts for language modelling, and pronunciation dictionaries. Despite speech applications becoming available in more languages, it is hard to imagine that resource collection at the required scale would be possible for all 7000 languages spoken in the world today.

I really like apples.

## 5.1. Section heading

This is some section with two table in it: Table 6.1 and Table 6.2.

**Table 5.1:** Performance of the unconstrained segmental Bayesian model on TIDigits1 over iterations in which the reference set is refined.

| Metric | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| WER (%) | $35.4$ | $23.5$ | $21.5$ | $21.2$ | $22.9$ |
| Average cluster purity (%) | $86.5$ | $89.7$ | $89.2$ | $88.5$ | $86.6$ |
| Word boundary $F$-score (%) | $70.6$ | $72.2$ | $71.8$ | $70.9$ | $69.4$ |
| Clusters covering 90% of data | 20 | 13 | 13 | 13 | 13 |

**Table 5.2:** A table with an example of using multiple columns.

| | Accuracy (%) | | |
|---|---|---|---|
| Model | Intermediate | Output | Bitrate |
| Baseline | 27.5 | 26.4 | 116 |
| VQ-VAE | 26.0 | 22.1 | 190 |
| CatVAE | 28.7 | 24.3 | 215 |

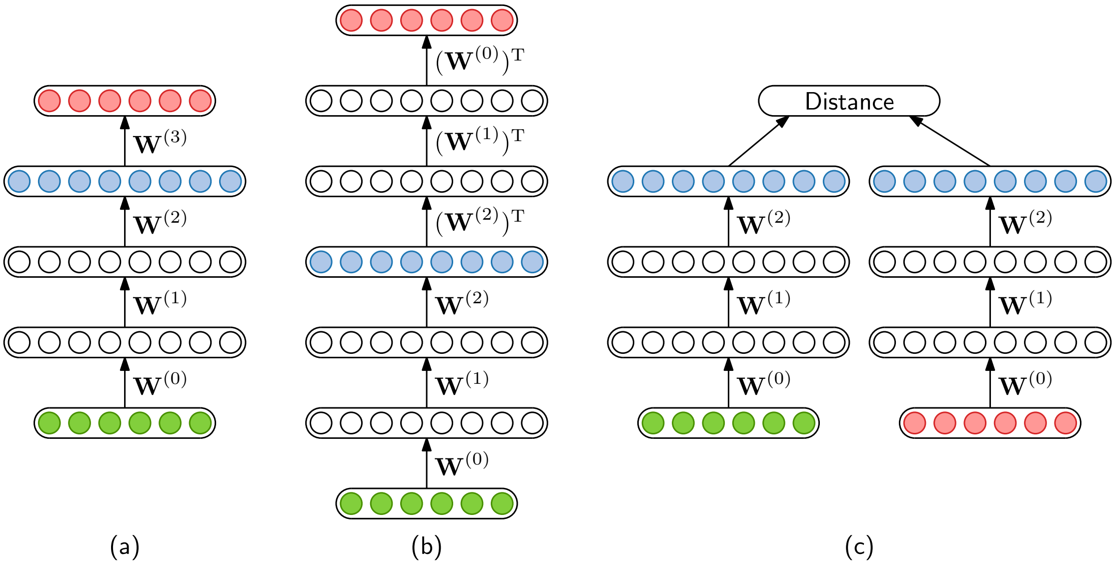

**Figure 5.1:** (a) The cAE as used in this chapter. The encoding layer (blue) is chosen based on performance on a development set. (b) The cAE with symmetrical tied weights. The encoding from the middle layer (blue) is always used. (c) The siamese DNN. The cosine distance between aligned frames (green and red) is either minimized or maximized depending on whether the frames belong to the same (discovered) word or not. A cAE can be seen as a type of DNN [1].

This is a new page, showing what the page headings looks like, and showing how to refer to a figure like Figure 6.1.

The following is an example of an equation:

$$P(\mathbf{z}|\boldsymbol{\alpha}) = \int_{\boldsymbol{\pi}} P(\mathbf{z}|\boldsymbol{\pi})\, p(\boldsymbol{\pi}|\boldsymbol{\alpha})\, \mathrm{d}\boldsymbol{\pi} = \int_{\boldsymbol{\pi}} \prod_{k=1}^{K} \pi_k^{N_k} \frac{1}{B(\boldsymbol{\alpha})} \prod_{k=1}^{K} \pi_k^{\alpha_k - 1}\, \mathrm{d}\boldsymbol{\pi} \tag{5.1}$$

which you can subsequently refer to as (6.1) or Equation 6.1. But make sure to consistently use the one or the other (and not mix the two ways of referring to equations).

## 5.2. Contributions

The following papers resulted from the work presented here:

*Conference paper 1:*
L. Skywalker, D. Vadar, and O. W. Kenobi, “A comparison between father-son and master-apprentice relationships in space conflict situations,” in *Proceedings of the International Conference on Action, Space and Star Politics (ICASSP)*, 2020.

*Journal paper 1:*
L. Skywalker, and L. Organa, “Identifying weaknesses in large evil corporations,” *IEEE Transactions on the Exploration of the Outer Rim*, vol. 21, pp. 154–174, 2021.

# Chapter 6

# Introduction

The last few years have seen great advances in speech recognition. Much of this progress is due to the resurgence of neural networks; most speech systems now rely on deep neural networks (DNNs) with millions of parameters [1, 2]. However, as the complexity of these models has grown, so has their reliance on labelled training data. Currently, system development requires large corpora of transcribed speech audio data, texts for language modelling, and pronunciation dictionaries. Despite speech applications becoming available in more languages, it is hard to imagine that resource collection at the required scale would be possible for all 7000 languages spoken in the world today.

I really like apples.

## 6.1. Section heading

This is some section with two table in it: Table 6.1 and Table 6.2.

**Table 6.1:** Performance of the unconstrained segmental Bayesian model on TIDigits1 over iterations in which the reference set is refined.

| Metric | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| WER (%) | 35.4 | 23.5 | 21.5 | 21.2 | 22.9 |
| Average cluster purity (%) | 86.5 | 89.7 | 89.2 | 88.5 | 86.6 |
| Word boundary $F$-score (%) | 70.6 | 72.2 | 71.8 | 70.9 | 69.4 |
| Clusters covering 90% of data | 20 | 13 | 13 | 13 | 13 |

**Table 6.2:** A table with an example of using multiple columns.

| | Accuracy (%) | | |
|---|---|---|---|
| Model | Intermediate | Output | Bitrate |
| Baseline | 27.5 | 26.4 | 116 |
| VQ-VAE | 26.0 | 22.1 | 190 |
| CatVAE | 28.7 | 24.3 | 215 |

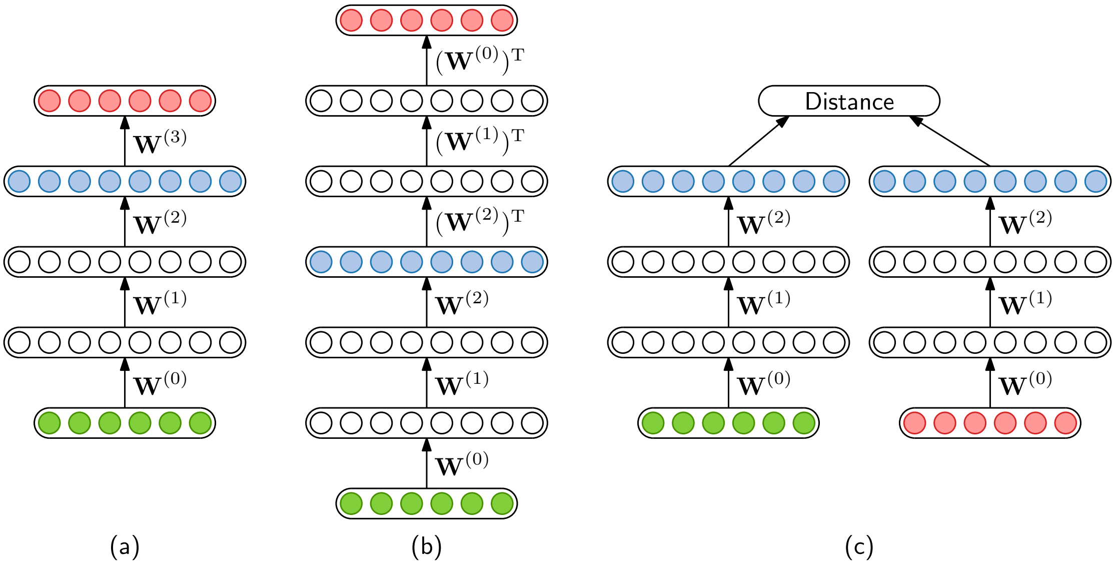
$\mathbf{W}^{(3)}$
$\mathbf{W}^{(2)}$
$\mathbf{W}^{(1)}$
$\mathbf{W}^{(0)}$
$(\mathbf{W}^{(0)})^\mathrm{T}$
$(\mathbf{W}^{(1)})^\mathrm{T}$
$(\mathbf{W}^{(2)})^\mathrm{T}$
Distance
(a)
(b)
(c)

**Figure 6.1:** (a) The cAE as used in this chapter. The encoding layer (blue) is chosen based on performance on a development set. (b) The cAE with symmetrical tied weights. The encoding from the middle layer (blue) is always used. (c) The siamese DNN. The cosine distance between aligned frames (green and red) is either minimized or maximized depending on whether the frames belong to the same (discovered) word or not. A cAE can be seen as a type of DNN [1].

This is a new page, showing what the page headings looks like, and showing how to refer to a figure like Figure 6.1.

The following is an example of an equation:

$$P(\mathbf{z}|\boldsymbol{\alpha}) = \int_{\boldsymbol{\pi}} P(\mathbf{z}|\boldsymbol{\pi})\, p(\boldsymbol{\pi}|\boldsymbol{\alpha})\, \mathrm{d}\boldsymbol{\pi} = \int_{\boldsymbol{\pi}} \prod_{k=1}^{K} \pi_k^{N_k} \frac{1}{B(\boldsymbol{\alpha})} \prod_{k=1}^{K} \pi_k^{\alpha_k - 1}\, \mathrm{d}\boldsymbol{\pi} \tag{6.1}$$

which you can subsequently refer to as (6.1) or Equation 6.1. But make sure to consistently use the one or the other (and not mix the two ways of referring to equations).

## 6.2. Contributions

The following papers resulted from the work presented here:

*Conference paper 1:*
L. Skywalker, D. Vadar, and O. W. Kenobi, “A comparison between father-son and master-apprentice relationships in space conflict situations,” in *Proceedings of the International Conference on Action, Space and Star Politics (ICASSP)*, 2020.

*Journal paper 1:*
L. Skywalker, and L. Organa, “Identifying weaknesses in large evil corporations,” *IEEE Transactions on the Exploration of the Outer Rim*, vol. 21, pp. 154–174, 2021.

# Chapter 7

# Summary and Conclusion

# Bibliography

[1] G. E. Dahl, D. Yu, L. Deng, and A. Acero, “Context-dependent pre-trained deep neural networks for large-vocabulary speech recognition,” *IEEE Trans. Audio, Speech, Language Process.*, vol. 20, no. 1, pp. 30–42, 2012.

[2] G. Hinton, L. Deng, D. Yu, G. E. Dahl, A.-R. Mohamed, N. Jaitly, A. Senior, V. Vanhoucke, P. Nguyen, T. N. Sainath, and B. Kingsbury, “Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups,” *IEEE Signal Process. Mag.*, vol. 29, no. 6, pp. 82–97, 2012.

# Appendix A

# Project Planning Schedule

This is an appendix.

# Appendix B

# Outcomes Compliance

This is another appendix.