



UNIVERSIDADE DO MINHO

DEPARTAMENTO DE INFORMÁTICA

Tecnologias e Aplicações  
Chinese MNIST Autoencoder

Jorge Mota (A85272)

June 29, 2021

# Contents

<b>1</b>	<b>Introdução</b>	<b>3</b>
<b>2</b>	<b>Preparação do dataset</b>	<b>4</b>
<b>3</b>	<b>Rede</b>	<b>5</b>
3.1	Treinamento . . . . .	6
<b>4</b>	<b>Conclusão</b>	<b>8</b>

# Chapter 1

## Introdução

Este projeto consiste no desenvolvimneto de redes generativas num contexto à escolha do aluno, e foi escolhido o desenvolvimento de autoencoder para caracteres chineses.

Autoencoders são tipos de redes neurais que aprendem a codificar de forma eficiente conjuntos de dados, usamos autoencoders concretos e temos extrações de features discretas sem recorrer a VAE, sendo que o espaço latente são vetores fixos ao invés de misturas de distribuições

O dataset escolhido para este projeto foi o MNIST chinês, que inclui conjunto de figuras com números chineses escrito à mão.

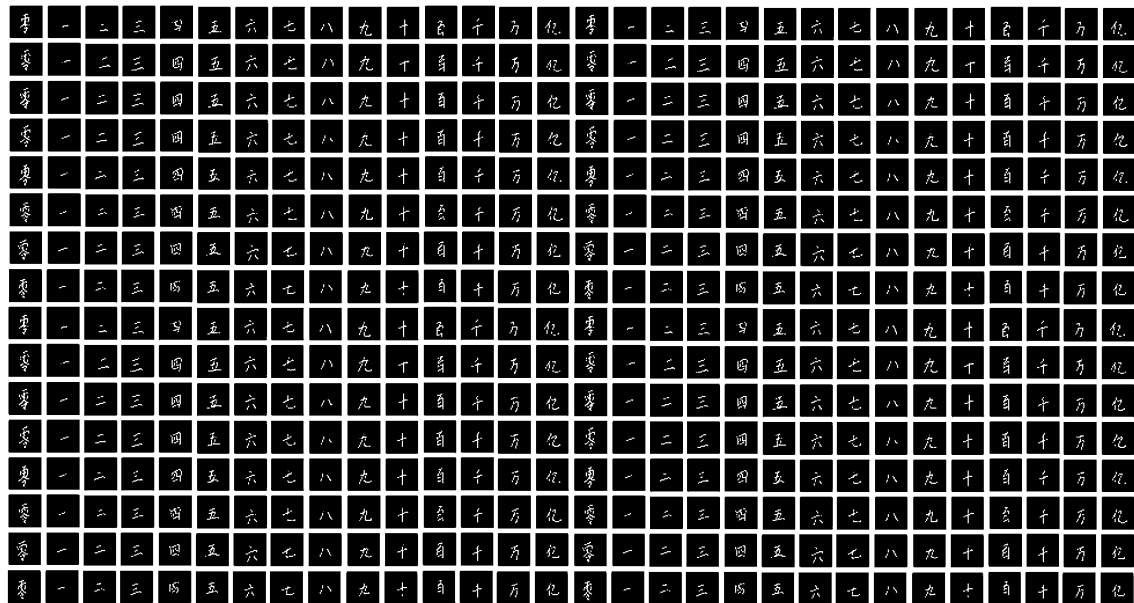


Figure 1.1: Caption

## Chapter 2

# Preparação do dataset

O principal caso de estudo é o dataset utilizado para treinar a rede neural, e para isto utilizei o *chinese MNIST dataset*, que contém cerca de 15'000 imagens de números escritos à mão. Para preparar os dados para treino, compus um script para dividir o dataset de forma justa para ter um *train set* e um *validation set* usado no treinamento com uma proporção de 70% e 30% respectivamente.

O produto do script de preparação separou o dataset na diretoria `chinese_mnist/` por `train_images` e `val_images`, em que cada uma destas diretorias tem pastas que identificam a label do conjunto de imagens que elas contêm (embora as labels não sejam utilizadas no treino da rede).

```
.
├── train_images
│   ├── 1
│   ├── 2
│   ├── ...
│   ├── 14
│   └── 15
└── val_images
    ├── 1
    ├── 2
    ├── ...
    ├── 14
    └── 15
```

# Chapter 3

## Rede

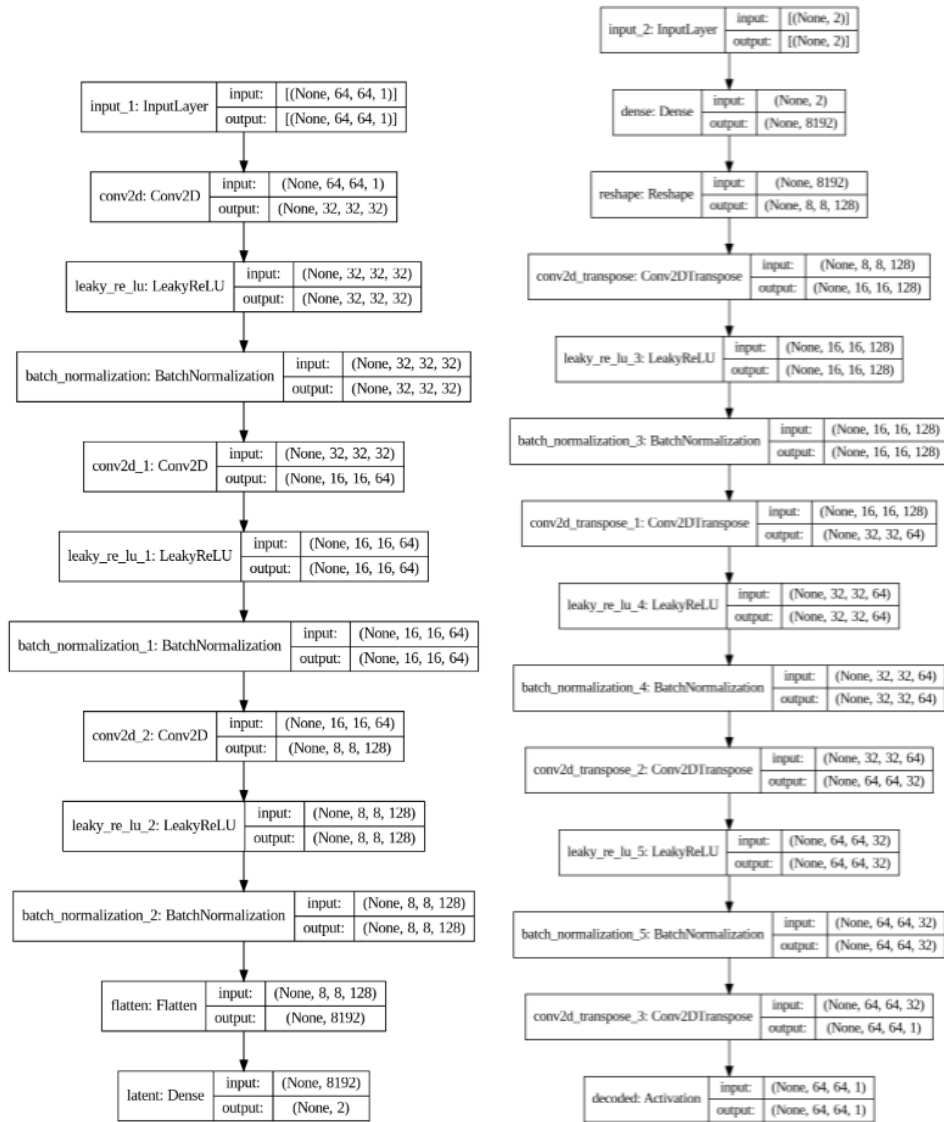


Figure 3.1: Autoencoder (Encoder e Decoder)

Esta rede é composta por três blocos convolucionais (`Conv2D` com função de ativação e `BatchNormalization`) de 128, 64 e 32 filtros. A simplicidade da rede deve-se ao facto de o dataset ser simples o suficiente para ser treinado com esta rede pequena, o que também se pode notar em exemplos para o dataset MNIST, que de certa forma é semelhante ao escolhido para este projeto.

### 3.1 Treinamento

Para treinar esta rede, como é um tipo de aprendizagem não supervisionada, passa-se como label a própria imagem ao mapear o dataset (`img, img`).

Foram realizados treinos com cerca de 25 e 50 *epochs* e os resultados não se mostraram de grande relevância, principalmente devido ao espaço latente ser de dimensão pequeno para o conjunto de dados usado para treinar.

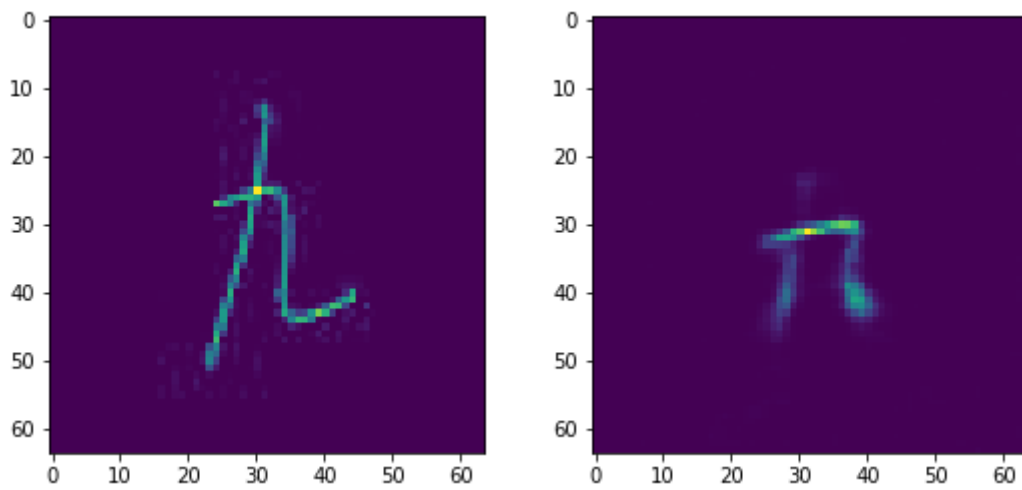


Figure 3.2: Original vs Reconstruída

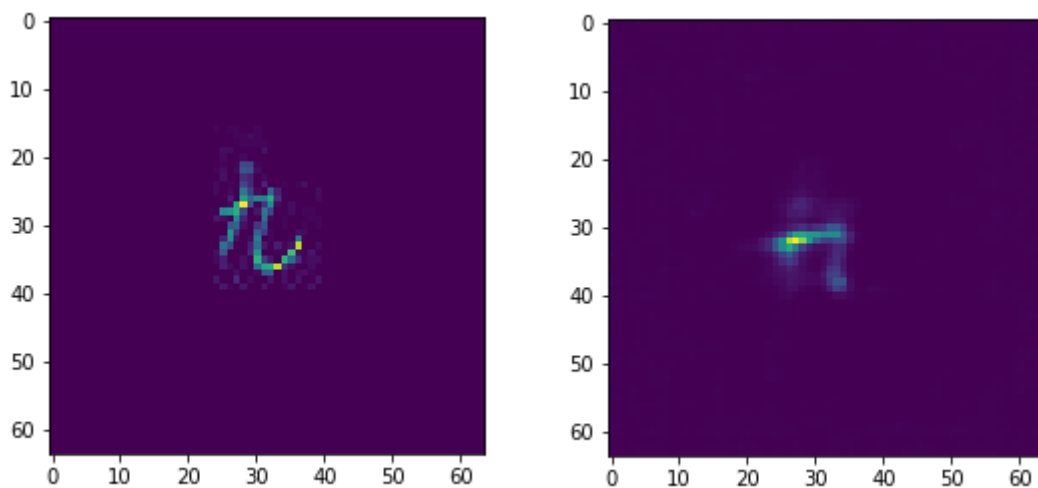


Figure 3.3: Original vs Reconstruída

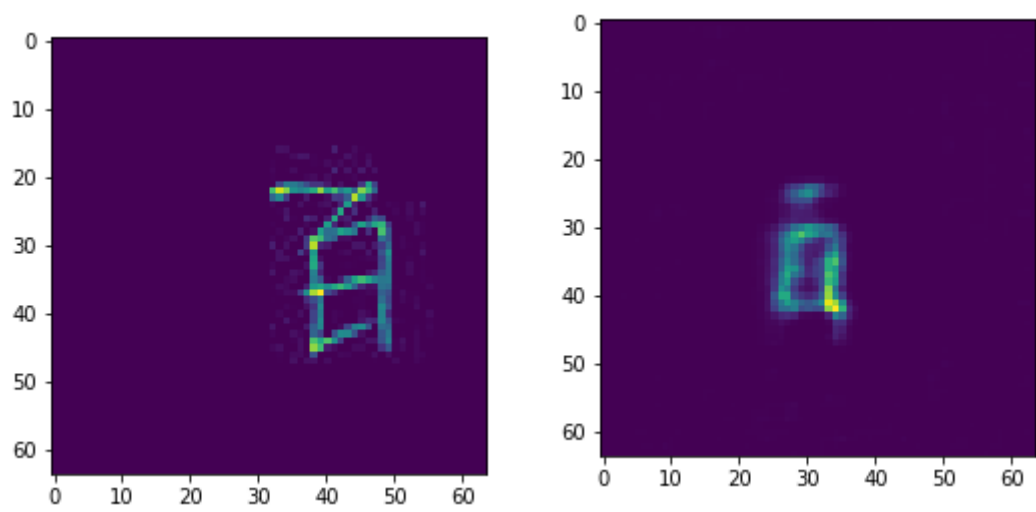


Figure 3.4: Original vs Reconstruída

## Chapter 4

# Conclusão

Devido à insuficiência da fiabilidade dos resultados, pode-se notar que há muitos aspetos que se podem mudar, especialmente a nível de treino. Deu para perceber os conceitos chave no desenvolvimento de redes neuronais não supervisionadas para autoencoders