

Instituto Nacional de Pesquisas Espaciais
Ministério da Ciência, Tecnologia, Inovações e Comunicações

COMPUTAÇÃO APLICADA
CAP 378 - TÓPICOS EM OBSERVAÇÃO DA TERRA

**Extração de características para compressão
de imagens de satélite**

Gabriel Augusto Lins Leal Pinheiro
Pedro Alexandre dos Santos

12 de setembro de 2018
São José dos Campos, São Paulo

Resumo

Quando você tira uma foto usando um celular, a imagem precisa ser armazenada, e para tal técnicas de compressão são empregadas, e o mesmo ocorre quando você assiste um vídeo, ouve uma música, entre várias outras atividades cotidianas. Hoje, a compressão é uma tecnologia essencial, que permite disponibilizar informação. Satélites de imageamento contém mecanismo responsáveis para o armazenamento e transmissão dos dados de captura. Com o avanço desses satélites e do aumento da resolução espacial, a utilização de técnicas para a compressão das imagens tornou-se uma tarefa importante, principalmente, no contexto de sensoriamento remoto, onde a imagem deve ser transmitida integralmente logo após a aquisição, devido a limitação da taxa de transmissão e de recursos a bordo. Neste contexto, o presente trabalho irá abordar métodos de compressão de imagens de satélite embarcado e solo.

Sumário

1	Introdução	5
1.1	Objetivo	6
2	Revisão teórica	8
2.1	Compressão	8
2.2	Compressão de imagens	8
2.2.1	Medidas de desempenho e qualidade	10
2.2.2	Mapeador	11
2.2.3	Quantização	12
2.2.4	Métodos de Entropia	13
2.2.5	O estado da arte	14
2.3	Aprendizado de máquina	15
2.3.1	Aprendizagem supervisionada	16
2.3.2	Aprendizagem não supervisionada	16
2.4	Rede de auto-codificação	16
2.5	Quanto aos sistemas embarcados	18
3	Resultados e Discussões	19
3.1	Atividades Realizadas	19
3.2	Dados	19
3.3	Tensorflow	20
3.4	Rede	23
3.4.1	Topologia	23
3.4.2	Função de ativação	24
3.4.3	Treinando a rede	25

SUMÁRIO

3.5 Aplicação	27
4 Conclusão	28
5 Bibliografia	28

Lista de Figuras

1	Estrutura básica de uma rede de auto-codificação.	17
2	Exemplos de amostras das classes que compõe o conjunto de dados.	20
3	Grafo da rede de auto-codificação.	21
4	Bloco “saida” do grafo representando a última camada da rede de auto-codificação.	22
5	Comportamento da função de ativação sigmoid.	24
6	Função de ativação Relu e Elu.	25
7	Comparação entre a figura original e o resultado obtido através da recuperação da codificação de 97,1% aproximadamente do dado de entrada.	26
8	Comparação entre a figura original e o resultado obtido através da recuperação da codificação de 98,7% aproximadamente do dado de entrada.	26

Capítulo 1

Introdução

Em 1858, o balonista G. Tournachon fez fotografias de Paris de seu balão marcando o nascimento do sensoriamento remoto moderno. Após este momento, um grande número de imagens foram tiradas por diversos tipos de plataformas, no entanto, a maioria dessas imagens não eram úteis para a elaboração de mapas ou propósitos científicos.

A era da fotografia aérea sistemática começou no início da Primeira Guerra Mundial, atingindo seu clímax durante a Guerra Fria. Outro evento significativo nesta história foi o desenvolvimento de satélites artificiais que aconteceram na segunda metade do século 20, e em 1999, quando o primeiro satélite comercial IKONOS iniciou a coleta de imagens digitais com alta resolução, Graham 1999; Baumann 2009. A modificação da plataforma de aquisição levou também a novos problemas, como enviar as imagens?

Uma imagem digital é composta por uma grade de linhas e colunas, onde o menor elemento é chamado de *pixel*. Uma imagem exibe e captura informações em um tempo e espaço específicos. Ela também pode conter camadas, comumente conhecidas como bandas, que estão por sua vez associados a diferentes faixas de comprimentos de onda. Em sensoriamento remoto, uma imagem pode ter uma única ou centenas de bandas.

Nos últimos anos do século XX, uma grande quantidade de imagens digitais se tornou disponível, e a necessidade de melhorar a qualidade e capacidade de processamento dessas imagens levou ao início de uma nova disciplina:

CAPÍTULO 1. INTRODUÇÃO

o processamento digital de imagens em dados de sensoriamento remoto. Ao mesmo tempo, esse aumento leva à necessidade de novos métodos de armazenamento dos dados, desta forma surge uma outra linha de pesquisa a compressão digital de imagens. Não somente, o armazenamento pode fazer uso destas técnicas mas a transmissão também, o que permite tratar o problema de envio das imagens, anteriormente citado.

Usualmente uma imagem pode apresentar informação (“pixels”) altamente correlacionada o que implica a necessidade de algoritmos que potencializem as componentes da informação. Nesse sentido, é possível aplicar técnicas de redução de dimensionalidade para reduzir a redundância, representando a informação em um espaço de dimensão menor, com a menor perda de informação possível, Sorzano, Vargas e Montano 2014, neste caso a análise de componentes principais, poderia ser vista como uma técnica de compressão.

Em 2006, Hinton e Salakhutdinov 2006, Hinton deu origem a pesquisas em redes neurais profundas, introduzindo uma rede neural de auto-codificação, seguindo de algumas comparações com algoritmos tradicionais, como a análise de componentes principais (PCA), que por sua vez encontra as direções de maior variação no conjunto de dados e representa cada ponto por suas coordenadas ao longo dessas direções. Uma rede neural de auto-codificação é uma generalização não linear do PCA e um algoritmo de aprendizado não supervisionado que é treinado para reproduzir os valores entrada, isto é, o valor de saída deve ser semelhantes ao da entrada.

A rede de auto-codificação é composto de duas partes: uma rede codificadora responsável por compactar os dados em um espaço de baixa-dimensão latente, o que é descrito pela função $h = f(x)$ e uma rede decodificadora para recuperar o dado real, representado pela função $r = g(h)$.

1.1 Objetivo

O objetivo desta monografia é um estudo e uma revisão da rede neural de auto-codificação para a compressão de imagens de satélite. Assim, pretendemos explorar as capacidades de sua não linearidade em comparação com

CAPÍTULO 1. INTRODUÇÃO

outras técnicas, e verificamos se esta é realmente capaz de lidar com imagens de satélite, ou seja, se uma reconstrução dos dados é possível.

Capítulo 2

Revisão teórica

2.1 Compressão

Compressão de dados é o processo de codificar informação usando um número de “bits” menor que a representação original. Surgiu entre 1940 e 1950 com os trabalhos de Claude Shannon em teoria de informação, C. E. Shannon s.d.; Claude E. Shannon 1993, o qual por sua vez fornece o arcabouço matemático para as chamadas técnicas de compressão sem perda. As técnicas de compressão com perda tem sua fundamentação na teoria de razão e distorção, Berger 1975. Assim, a ideia de compressão esta intimamente conectada com estatística de inferência.

2.2 Compressão de imagens

Compressão de imagens é a técnica de reduzir o tamanho da imagem com a menor degradação de qualidade possível. Existem diversos algoritmos de compressão de dados que poderiam ser aplicados para tal tarefa, mas o resultado é muito aquém do desejável, em outras palavras, técnicas usuais de compressão não são adequadas para lidar com imagens, isto ocorre pois elas não são capazes de lidar com a redundância de informação espacial, isto é, uma vez que uma imagem pode ser vista como uma matriz bidimensional de *pixels*, a informação de um *pixel* pode parcialmente ser deduzida a partir de

CAPÍTULO 2. REVISÃO TEÓRICA

seus vizinhos, ou os vizinhos compartilham informação entre si.

Melhorias nessas técnicas levam em consideração redundância e informação desnecessária, esta última lida com componentes da imagens que não perceptíveis para os humanos, isto é, informação irrelevante à visão humana. Um ponto interessante quanto a esse tipo de técnica é que máquinas não apresentam as mesmas limitações que os humanos, então, estas componentes poderiam ser interessantes para elas?

As técnicas tradicionais de compressão podem ser esboçadas com os seguintes elementos principais:

- **mapeador:** o papel deste é reduzir a correlação espacial, para tal utiliza de métodos como a predição de *pixel* ou a transformada da imagem para outro espaço de representação. A predição é usada para prever o valor de um *pixel* a partir dos seus vizinhos, um exemplo, é o método de modulação de código de pulso diferencial, enquanto o papel da transformada é mapear a imagem do domínio do espaço para a frequência, aplicando-se, por exemplo, a transformada discreta do cosseno ou uma transformada de ondaletas;
- **quantizador:** o papel deste é remover componentes dos valores mapeados descartando informação não perceptível para os humanos. Uma técnica usual é o descarte de elementos de alta-frequência, seja por remoção ou sobre quantização. Somente está presente nos algoritmos com perda;
- **codificador de símbolo:** explora a redundância decorrente das repetições de símbolos, por inferência, buscando uma representação mais compacta, isto é, atribui aos símbolos de maior probabilidade códigos de palavras menores. Normalmente são associados a codificação por entropia, e dois métodos tradicionais para lidar com este tipo de redundância são a codificação de Huffman e a codificação aritmética.

Os três elementos acima compõem o sistema de codificação. A decodificação, por sua vez, apresenta apenas dois elementos um decodificador de símbolos, que realiza o papel inverso do codificador e um mapeador inverso,

CAPÍTULO 2. REVISÃO TEÓRICA

o qual toma, por exemplo, uma imagem no domínio da frequência e produz uma correspondente no domínio do espaço.

Basicamente, pode-se ter dois tipos de compressão:

- Compressão sem perda: o sistema de codificação faz uso apenas do mapeador e da codificação por entropia. Seu uso se dá principalmente em aplicações onde perda de informação não é tolerável;
- Compressão com perda: o algoritmo de compressão da imagem tolera uma perda na quantidade de informação com o propósito de aumentar a taxa de compressão, de tal forma que a recuperação do dado seja a mais próxima possível da imagem original. Atualmente, algoritmos de compressão com perda são amplamente utilizados em transmissão de imagens via internet.

2.2.1 Medidas de desempenho e qualidade

A avaliação da qualidade de uma compressão se pode dar de duas maneiras: subjetiva e objetiva, isto é, tem-se critérios de fidelidade objetivos e critérios de fidelidade subjetivos.

No caso dos objetivos, a perda de informação pode ser expressa formalmente por uma expressão matemática. Uma medida de distorção para compressão com perda é o erro quadrático médio, MSE :

$$MSE = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} (\hat{f}(x, y) - f(x, y))^2, \quad (2.1)$$

onde $f(x, y)$ é a imagem original, $\hat{f}(x, y)$ é a imagem reconstruída e $M \times N$ é a dimensão da imagem. O erro associado a cada coordenada (x, y) é

$$e(x, y) = \hat{f}(x, y) - f(x, y) \quad (2.2)$$

e o erro total entre duas imagens é dado por

$$MN \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [\hat{f}(x, y) - f(x, y)]. \quad (2.3)$$

CAPÍTULO 2. REVISÃO TEÓRICA

A relação sinal-ruído ($PSNR$) é uma outra medida muito utilizada, sendo também aplicável a algoritmos de compressão sem perda. É mensurada em dB, e dada por

$$PSNR = 10 \log_{10} \frac{(2^B - 1)^2}{(MSE + 1/12)}, \quad (2.4)$$

onde B é a faixa dinâmica em bits da imagem original, e o papel do fator $1/12$ é o de eliminar divergências no valor de $PSNR$ quando $MSE \rightarrow 0$.

A média quadrática da relação sinal-ruído para uma imagem reconstruída é dada por

$$SNR_{ms} = \frac{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \hat{f}(x, y)^2}{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [\hat{f}(x, y) - f(x, y)]^2}. \quad (2.5)$$

Os métodos de avaliação objetivos somente são capazes de lidar com diferenças exclusivamente de maneira quantitativa, o que é adequado quando a imagem reconstruída será utilizada como um dado de entrada para outros algoritmos de processamento. Note entretanto, que uma imagem com alto valor de MSE pode conter informações visuais mais representativas, do que a mesma imagem com um valor menor de MSE .

Então, as avaliações subjetivas surgem para lidar com a forma que o sistema visual humano lida com as imagens, e portanto, avaliar a qualidade de uma imagem segundo este aspecto. Uma análise deste tipo, poderia ser realizada, apresentando um conjunto de imagens descomprimidas e a sua contra-parte original, para uma pessoa, tal que este indivíduo, qualifique as imagens, isto é, atribua nível de qualidade às imagens.

2.2.2 Mapeador

As técnicas de mapeamento são variadas, todavia não são utilizadas sozinhas o que torna difícil encontrar exemplos. Neste caso, decidiu-se por utilizar a modulação por codificação de pulso diferencial descartando-se a quantização para ilustrar tal ideia. A técnica original foi desenvolvida por Cutler, em 1952, O'Neal s.d., e é uma técnica de compressão com perda (decorrente do processo de quantização).

CAPÍTULO 2. REVISÃO TEÓRICA

Inicialmente, transmite-se a primeira amostra do sinal, e cada amostra subsequente será transmitido a diferença entre a amostra atual e a amostra anterior. A ideia de se utilizar a diferença é que se pressupõe que ela requer menos bits de representação que a informação original. Neste caso, ela pode ser vista como um modelo de previsão, onde o valor de um pixel é estimado dado seu antecessor.

2.2.3 Quantização

Existem diversas técnicas de quantização, entretanto, em geral, estas dependem da saída do mapeador, isto é, a imagem agora está apresentada no espaço das frequências, ou é uma representação de erros. Assim, decidiu-se apresentar aqui uma técnica que permite ilustrar as ideias por trás desta etapa.

Codificação por truncamento de bloco

A codificação de truncamento de bloco foi desenvolvida por Delp e Mitchell, em 1979, Delp e O. Mitchell 1979. Como é uma técnica de quantização apresenta uma perda natural. O seu princípio de funcionamento é quantificar os *pixels* de uma imagem preservando o primeiro e o segundo momento estatístico. A técnica inicia, com a divisão da imagem em blocos, cada qual tendo suas propriedades calculadas individualmente. Para cada bloco, calcula-se o valor médio e a variância dos valores de cinza. Então, constrói-se uma mapa binário, onde se o valor do nível de cinza para um dado pixel for maior que o valor médio, a entrada recebe valor 1, caso contrário 0. Devem, então, ser armazenados o mapa binário, o valor médio, e a variância para cada bloco.

Na descompressão, os valores no mapa binário com entrada 1 são trocados pelo valor médio mais a variância multiplicada por um fator ω , enquanto os valores de 0 são trocados pelo valor médio menos a variância multiplicada pelo inverso do valor ω , onde este é definido, como a raiz quadrada da razão, entre o número de 0 e 1 naquele mapa binário.

2.2.4 Métodos de Entropia

Alguns dos métodos mais conhecidos de compressão por entropia são:

Codificação por comprimento de corrida

A codificação por comprimento de corrida é uma técnica de compressão aplicada a sequências de símbolos repetidos. Ela apresenta um funcionamento bem simples: considere a ocorrência de uma repetição contínua de um determinado símbolo, por exemplo, AAAAAAAAAAAAA, é possível se adotar a representação dada pelo par (12, A), onde o primeiro elemento identifica o número de ocorrências, enquanto o segundo identifica o símbolo, Salomon 2012.

Codificação de Shannon-Fano

A codificação de Shannon-Fano, C. E. Shannon s.d.; Claude E. Shannon 1993 é um método de estatístico de compressão, foi introduzido, em 1948, por Claude Shannon em seu famoso artigo “A Mathematical Theory of Communication” e atribuído à Robert Fano.

Este método introduz a ideia de gerar códigos de tamanho variável para cada símbolo do conjunto de dados original de acordo com sua probabilidade de ocorrência. Desta forma, símbolos com alta probabilidade de ocorrência são representados por símbolos com menor quantidade de bits, e símbolos com baixa probabilidade de ocorrer são codificados por símbolos com um número maior de bits.

O algoritmo é bastante simples, para ilustrar suas diferenças em relações aos outros algoritmos que apresentam as mesmas ideias, apresenta-se o processo de codificação deste:

1. faz-se a levantamento das probabilidades de ocorrência de cada símbolo nos dados originais;
2. ordena-se lista de probabilidades em ordem decrescente;

CAPÍTULO 2. REVISÃO TEÓRICA

3. dividi-se a lista de probabilidades em duas, sem alterar a ordem dos elementos, de forma, que cada uma dessas partes tenha aproximadamente a mesma probabilidade, isto é, a soma de probabilidades de cada nova lista deve ser o mais próximo de 50%, quanto possível;
4. a cada uma dessas parte atribuí-se um número, 0 para a primeira lista e 1 para a segunda;
5. a cada nova lista que tiver mais de um símbolo, aplica-se o mesmo processo, concatenando-se os símbolos 0 ou 1 atribuídos a cada etapa, até todos as listas terem apenas um dígito;
6. a sequencia de símbolos associados a cada lista mapeia a codificação do mesmo.

Codificação de Huffman

A codificação de Huffman, Huffman 1952, assim como a de Shannon-Fano, é um método que utiliza a probabilidade da ocorrência de um símbolo para atribuir um símbolo de tamanho variável. Foi desenvolvido em 1952 por David A. Huffman, e publicado no artigo “A Method for the Construction of Minimum-Redundancy Codes”.

A ideia central da técnica é construir uma árvore binária completa, chamada de árvore de Huffman, a qual é construída recursivamente a partir da junção os dois símbolos de menor probabilidade, que são então somados em símbolos auxiliares e colocados no conjunto de símbolos. O processo é encerrado quando todos os símbolos forem combinados em elementos auxiliares, formando uma árvore binária, o qual é então percorrida, atribuindo-se valores binários para as arestas, com os códigos gerados a partir deste percurso.

2.2.5 O estado da arte

As técnicas de compressão surgiram em decorrência do trabalho de Shannon sobre quantificação de informação, que culminou na entropia de Shannon. Desta forma, as primeiras técnicas de compressão desenvolvidas são as de

CAPÍTULO 2. REVISÃO TEÓRICA

símbolo. Somente posteriormente, com aplicações em novas áreas e novas demandas, surgem o mapeador e o quantificador.

Em 2018, as técnicas mais modernas ainda mantêm a mesma estrutura, entretanto, diversos refinamentos surgiram para cada etapa, como novas transformadas para o mapeador, por exemplo, a LBT (lifting-based reversible hierarchical lapped biorthogonal transform), utilizada no JPEG-X, Dufaux, J. Sullivan e Ebrahimi 2009, ou a utilização de aprendizado de máquina no processo de codificação por entropia, por exemplo, o FLIF (Free Lossless Image Format), Sneyers e Wuille 2016, utiliza de árvores de decisão para adaptar um algoritmo de codificação aritmética.

E, uma próxima geração, já está em desenvolvimento, o qual, por exemplo emprega redes neurais profundas como mecanismo de compressão, sendo que neste caso, a rede deve aprender todos os elementos básicos definidos acima.

2.3 Aprendizado de máquina

Devido ao avanço computacional e da quantidade massiva de dados disponíveis e gerados a cada momento, a aprendizagem de máquina tornou-se uma tecnologia amplamente aplicada em diversas ferramentas e serviços de nosso cotidiano, sendo muitas dessas vezes, de forma imperceptível.

O aprendizado de máquina pode ser definido como um campo de estudo sobre as habilidades dos computadores em aprenderem sem serem explicitamente programados (Munoz 2014). De acordo com T. M. Mitchell et al. 1997, em um problema de aprendizagem um programa de computador aprende a partir da experiência E em relação a alguma tarefa T com uma performance P , se sua performance em T , mensurada por P , melhora com a experiência E .

Uma observação ou exemplo (\mathbf{x}) é formada por um ou mais atributos, no qual podem ser representados na forma vetorial $\mathbf{x} = (x_1, \dots, x_D) \in \mathbb{R}^D$, simbolizando as características de um determinado objeto. Além disso, uma observação pode ou não conter um rótulo, que refere-se a classe ou o resultado que pertence a um determinado exemplo.

Dessa forma, observações que são usadas durante o processo de treina-

CAPÍTULO 2. REVISÃO TEÓRICA

mento de um modelo são denominadas como conjunto de treinamento e podem ser associadas como a experiência pelo qual o algoritmo de aprendizagem irá ter de aprender.

2.3.1 Aprendizagem supervisionada

A aprendizagem de máquina supervisionada (do inglês *Supervised Learning*) consiste em treinar um modelo a partir de um par de entrada, formado pelo exemplo e seu rótulo correspondente. Por conta disso, neste paradigma já se conhece as categorias existentes do conjunto de dados. A classificação e a regressão, são duas abordagens da aprendizagem supervisionada que visam prever as saídas lidando com valores discretos e contínuos, respectivamente.

Considere \mathcal{X} sendo o domínio das observações e \mathcal{Y} o domínio dos rótulos. Seja $P(\mathbf{x}, y)$ a probabilidade conjunta, possivelmente desconhecida, sobre as observações e rótulos $\mathcal{X} \times \mathcal{Y}$, um algoritmo de aprendizagem supervisionada tem por objetivo treinar uma função $f : \mathcal{X} \mapsto \mathcal{Y}$, tal que $f(\mathbf{x})$ consiga prever o rótulo y correto para uma dada observação \mathbf{x} , a partir de um conjunto de treino $\{(x_i, y_i)\}_{i=1}^n \stackrel{i.i.d.}{\sim} P(\mathbf{x}, y)$ (Zhu e Goldberg 2009).

2.3.2 Aprendizagem não supervisionada

Algoritmos de aprendizagem não supervisionada (do inglês *Unsupervised Learning*) são utilizados quando deseja-se encontrar algum padrão na estrutura do conjunto de dados. Comumente métodos como, clusterização, detecção de novidade e redução da dimensionalidade são aplicados neste tipo de paradigma de aprendizagem, onde o conjunto de treino $\{x_i\}_{i=1}^n$ não diz nada a respeito sobre os rótulos de cada observação.

2.4 Rede de auto-codificação

Uma rede de auto-codificação (em inglês *Autoencoder*) é classificado como um algoritmo auto-supervisionado - um tipo de aprendizagem não supervisionada, no qual o programa é treinado para que sua saída seja a mesma que a

CAPÍTULO 2. REVISÃO TEÓRICA

entrada. Uma rede de auto-codificação é uma rede neural artificial, cujo seu objetivo é aprender a fazer uma representação dos dados, para que possa ser mapeada (codificada) em um espaço de dimensão menor, com a finalidade da redução de dimensionalidade.

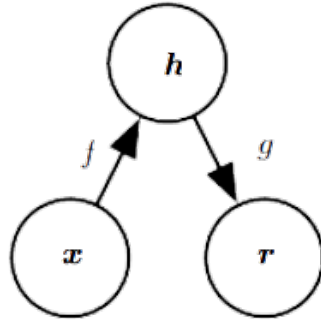


Figura 1: Estrutura básica de uma rede de auto-codificação.

Uma rede de auto-codificação básica é composta de três camadas. A Figura 1 mostra a estrutura geral dessa rede, onde da primeira à segunda camada (codificador), o modelo aprende a representar os dados em uma codificação menor, e da segunda à terceira camada (decodificador), o modelo aprende a descomprimir o código, resultando em uma saída que corresponda ao dado original.

Desta forma, é uma rede neural não-recorrente, *feedforward*, muito semelhante a um perceptron multicamadas (MLP) com uma ou mais camadas ocultas. O processo de aprendizagem é quantificado e alcançado minimizando a função *MSE* entre a reconstrução e a imagem original:

$$loss = \| x - g(f(x)) \|^2 \quad (2.6)$$

onde $f(\cdot)$ representa a função que mapeia a codificação e $g(\cdot)$ a decodificação. Esse modelo pode ser generalizado para incluir mais camadas ocultas, mas a camada de entrada e saída tem sempre o mesmo número de neurônios e, em geral, há alguma simetria, de tal forma que o primeiro conjunto de camadas ocultas (codificador) tem a mesma cardinalidade do segundo conjunto (decodificador), e para a camada i , com n entradas e m saídas, a camada

$N - 1 - i$ terá m entradas e n saídas, onde N é o número de camadas e a primeira camada possui índice 0.

2.5 Quanto aos sistemas embarcados

Em sua grande maioria, os satélites operam em um modo de aquisição e envio dos dados obtidos por sensores. Observe, contudo, que a evolução dos sensores, levou a um aumento significativo no volume dos dados, levando a uma importante questão das missões espaciais atuais, como balancear a largura de banda para transmissão de dados com o volume de dados sendo adquirido.

É importante adicionar ainda que estes sistemas embarcados apresentam várias limitações, por exemplo, na capacidade de armazenamento, no consumo de energia elétrica, na capacidade de dissipação. Assim, como a necessidade de lidar com um ambiente extremo, com baixa temperatura, alto nível de radiação, tempestades geomagnéticas, entre outras.

Neste caso, tem-se que a técnica de compressão deve apresentar baixa complexidade, para que possa ser implementada fisicamente, seja em uma FPGA ou um ASIC. Outra necessidade é a capacidade de lidar com dados brutos recolhidos diretamente do sensor, que podem apresentar diversos tipos de ruídos, reduzindo drasticamente a eficiência da compressão.

Outro requisito é a capacidade de lidar com perdas de blocos de dados. É usual em comunicações na situação desenvolvida acima, a perda de informação. Neste caso, uma técnica de compressão que torne o dado não decodificável após a perda de um único pacote é inadequada, em outras palavras, existe alguma necessidade de resiliência.

Capítulo 3

Resultados e Discussões

3.1 Atividades Realizadas

Este trabalho tem por objetivo investigar técnicas de compressão de imagem de satélite, assim como também, fazer a implementação de redes neurais profundas como uma possível alternativa para a substituição de técnicas convencionais de compressão. Neste capítulo, será abordado o processo de desenvolvimento das rede de auto-codificação, levando em consideração o erro gerado e a capacidade de compressão da rede.

3.2 Dados

O conjunto de dados utilizado neste trabalho é composto por 500.000 imagens do DeepSat (SAT-4) Airborne Dataset (Crawford 2018), originalmente, extraídas do banco de imagens do National Agriculture Imagery Program (NAIP). Cada amostra é uma imagem formada por 28×28 *pixels* e 4 bandas: vermelho, verde, azul e infravermelho.

A princípio, os recortes de imagens foram retiradas da base original, que é formada por 1.500 imagens de satélite (cada uma composta por aproximadamente 6.000×7.000 pixel com resolução espacial de 1 metro) cobrindo diferentes partes do estado de Califórnia. A Figura 2 apresenta amostras das diferentes classes que formam o conjunto de dados final, composta por:

CAPÍTULO 3. RESULTADOS E DISCUSSÕES

terra estéril, árvores, pastagem e outros, sendo esta última classe contendo, geralmente, construções de prédios ou casa e estradas.

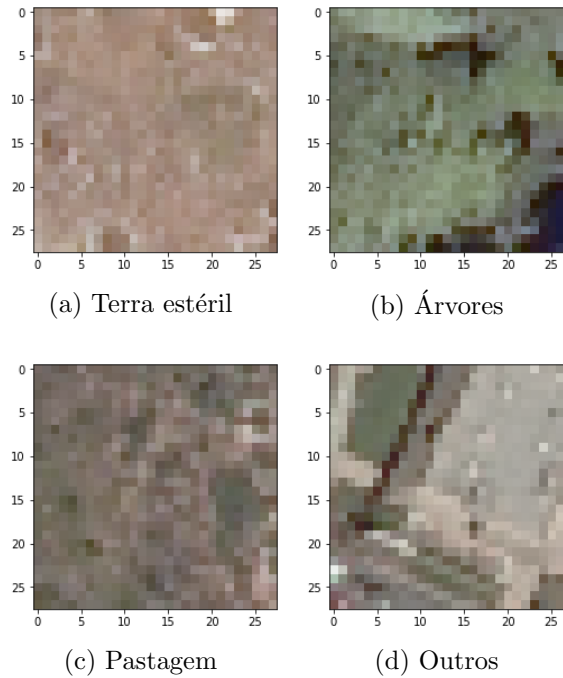


Figura 2: Exemplos de amostras das classes que compõe o conjunto de dados.

3.3 Tensorflow

Tensorflow é uma biblioteca de computação numérica com código aberto, desenvolvida pela Google. Atualmente, é amplamente utilizada em algoritmos de aprendizado de máquina, como redes neurais, por permitir criar grafos de fluxo onde cada nó representa uma operação computacional. Dessa forma, trazendo inúmeras vantagens, tais como processamento paralelo e suporte a várias plataformas (Hope, Resheff e Lieder 2017).

Para a implementação da rede de auto-codificação foi utilizado o Tensorflow, visto o potencial e facilidades que a ferramenta oferece. A Figura 3 apresenta o grafo gerado pela biblioteca, onde um nó acaba representando uma camada ou alguma transformação empregado durante o treinamento da

CAPÍTULO 3. RESULTADOS E DISCUSSÕES

rede.

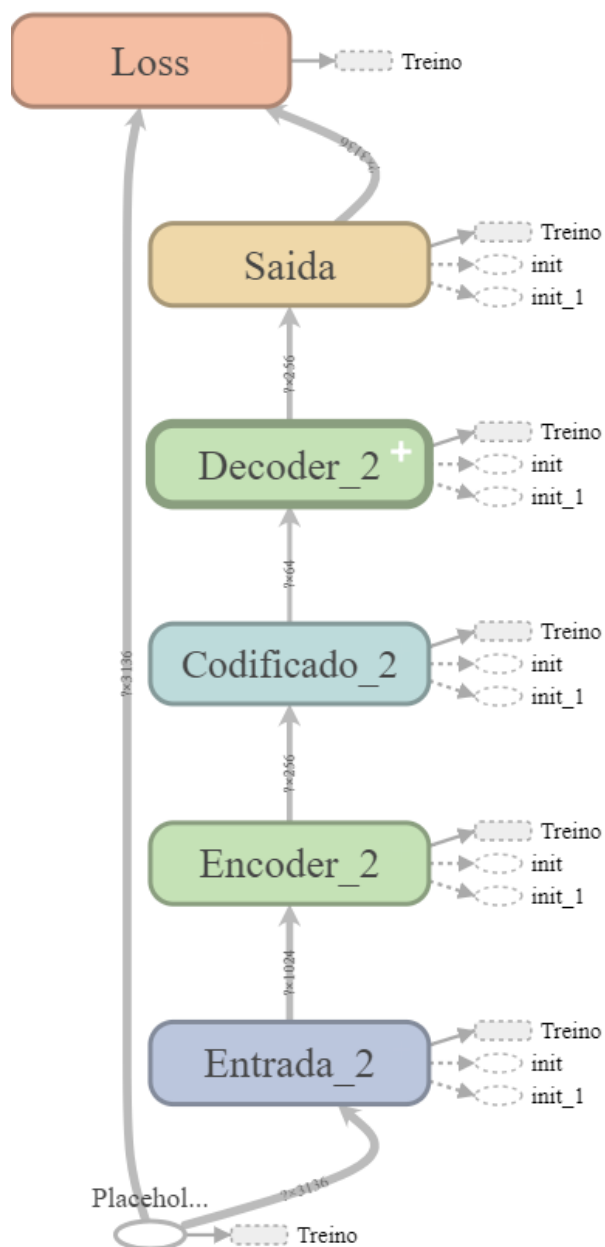


Figura 3: Grafo da rede de auto-codificação.

Em geral, cada nó no grafo corresponde a alguma operação sobre uma

CAPÍTULO 3. RESULTADOS E DISCUSSÕES

entrada recebida, para que assim possa gerar uma saída, que poderá ou não ser utilizada nos demais nós que compõe o grafo, caso haja alguma dependência. Uma dependência surge quando um nó requer como entrada o cálculo realizado a partir da operação de um outro nó. Assim sendo, chama-se de transformação o processo de manipulação do dado que ocorre em uma sequência de nós, por meio de um conjunto de entrada e saída.

A Figura 4 apresenta detalhadamente as operações que ocorrem no bloco “Saída” do grafo. Conforme a Figura 3, é possível notar que sua entrada é dada pela decodificação dos dados, realizada no bloco anterior, que acabam sendo multiplicados pelo pesos, somados pelo viés e aplicados sobre a função de ativação sigmoid da camada atual para a geração da saída. Dessa mesma forma, é feito esse processo de transformação dos dados nos demais bloco do grafo, porém cada qual com o seu objetivo.

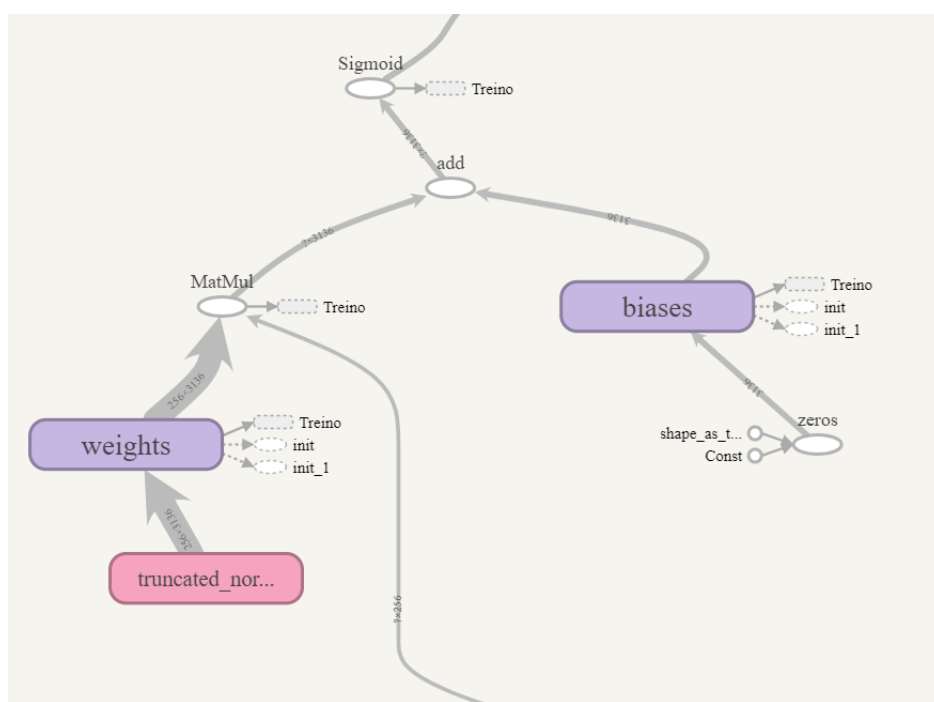


Figura 4: Bloco “saida” do grafo representando a última camada da rede de auto-codificação.

3.4 Rede

A rede de auto-codificação tem como base uma camada de entrada, responsável pelo recebimento das imagens, uma ou mais camadas escondidas, usadas para extrair informações relevantes no processo de codificação e decodificação dos dados, e por fim, uma camada de saída para a reconstrução dos dados. O algoritmo, de forma simplificada, pode ser ilustrado seguindo os processos abaixo.

1. inicia-se os pesos da rede aleatoriamente;
2. repita os seguintes processos de acordo com o número de iterações definido:
 - alimente a rede com o conjunto de treinamento;
 - calcule a saída com base nos pesos;
 - calcule o erro gerado pela saída;
 - Baseado no erro, faça o reajuste dos pesos;
 - Atualize os pesos;

3.4.1 Topologia

A rede de auto-codificação desenvolvida neste trabalho é caracterizada como rede direta (do inglês *feed-forward*) de camadas múltiplas. Durante os experimentos, diferentes quantidades de neurônios e camadas foram empregados com o objetivo de obter uma melhor compactação dos dados.

As duas configurações de rede a serem apresentadas, foram projetadas para receber uma entrada de 3.136 atributos, referentes aos 28x28 *pixels* de cada banda existente na imagem. A primeira rede foi delineada com 7 camadas na parte de codificação: 3.136-2.048-1.024-512-256-128-90 e com uma decodificação simétrica, com funções elu nas cinco primeiras camadas, uma função relu na penúltima camada e por fim, uma função linear na camada de compressão.

CAPÍTULO 3. RESULTADOS E DISCUSSÕES

Já a segunda rede foi definida para comprimir os dados em apenas 40 níveis de cinza mais representativos, seguindo a mesma topologia que a rede anterior, porém, composta apenas por 5 camadas em cada uma das duas partes (codificador e decodificador): 3.136-2.048-1.024-512-40.

3.4.2 Função de ativação

De modo geral, as funções de ativação são usadas com o propósito de fazer a rede aprender relações não lineares entre as variáveis (Ponti e Costa 2018). Na implementação da rede de auto-codificação foram usados 4 tipos diferentes de função de ativação: linear, Sigmoid, *Rectified Linear Unit* (Relu), *Exponential Linear Unit* (Elu).

A função linear conta com a própria transformação realizada pelo neurônios sobre a entrada, através dos pesos e viés. Já função sigmoid, vide Equação 3.1, é definida dentro do intervalo $[0, 1]$, e quanto menor for o valor de x , mais próximo do zero a função irá tender, da mesma forma, que para valores maiores de x , mais próximo do um o resultado será, conforme a Figura 5.

$$\sigma(x) = \frac{1}{1 + e^x} \quad (3.1)$$

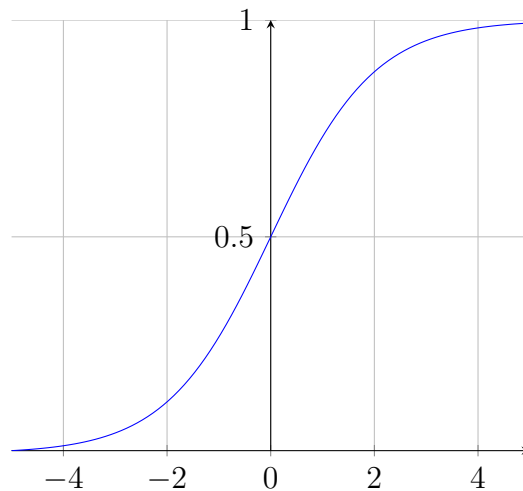


Figura 5: Comportamento da função de ativação sigmoid.

CAPÍTULO 3. RESULTADOS E DISCUSSÕES

A função de ativação Relu, de acordo com a Equação 3.2, é definida como 0 para valores de x menores que 0 e função identidade, caso contrario. Por fim, a função Elu, vide Equação 3.3, é uma das variações da função do tipo Relu, diferenciando apenas em e^x quando x é menor que 0. Na Figura 6 é possível visualizar o comportamento das respectivas funções dentro do intervalo $[-3, 3]$.

$$f(x) = \begin{cases} 0 & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases} \quad (3.2)$$

$$f(x) = \begin{cases} e^x - 1 & \text{for } x < 0 \\ x & \text{for } x \geq 0 \end{cases} \quad (3.3)$$

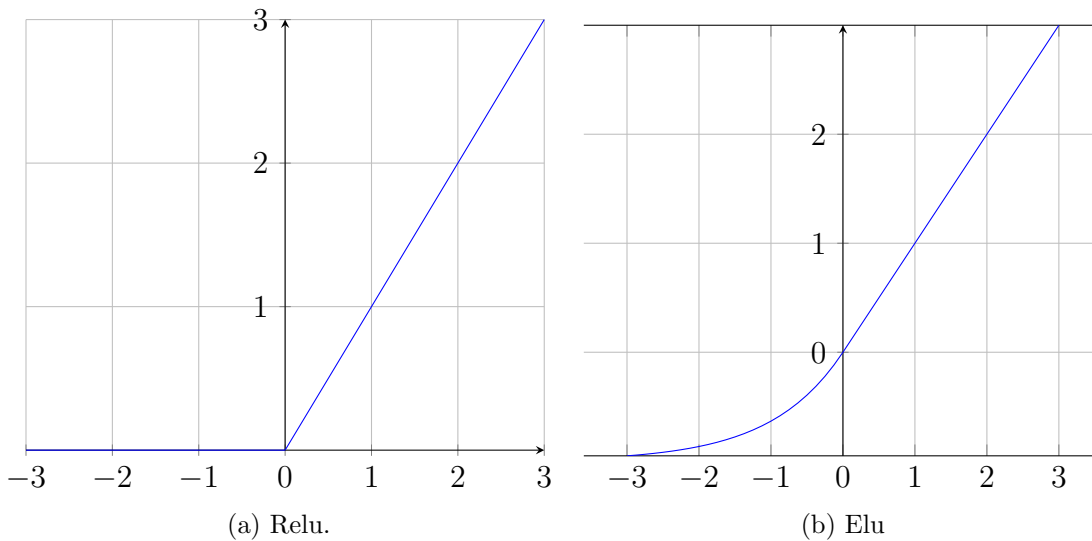


Figura 6: Função de ativação Relu e Elu.

3.4.3 Treinando a rede

Para o processo de treinamento foi utilizado 15.000 imagens, escolhidas aleatoriamente da base original como o conjunto de treino, sendo a rede configurada com tamanho de mini-batch de 100 amostras, taxa de aprendizagem igual a 1×10^{-4} e percorrendo um total de 100 épocas. Figura 7 mostra o

CAPÍTULO 3. RESULTADOS E DISCUSSÕES

resultado obtido usando a primeira rede, com uma compactação equivalente a 97,1% aproximadamente sobre dados os dados originais.

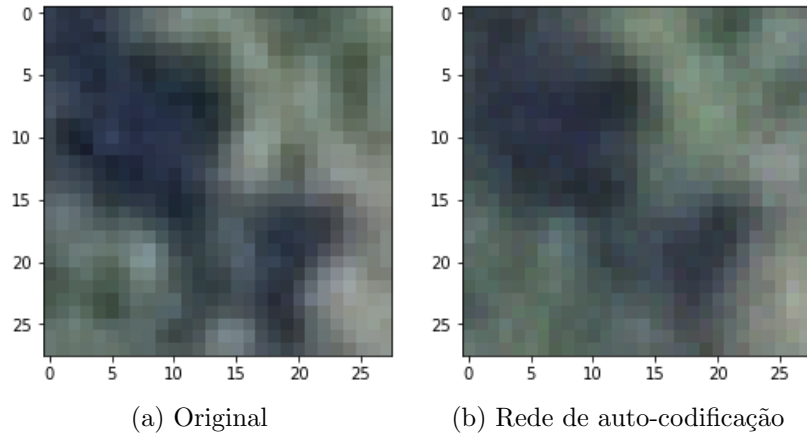


Figura 7: Comparação entre a figura original e o resultado obtido através da recuperação da codificação de 97,1% aproximadamente do dado de entrada.

Já a Figura 8 é apresentado a compressão feita através da segunda rede, neste caso o algoritmo conseguiu comprimir os dados em aproximadamente 98,7%. Em ambos casos, o erro quadrático médio gerado foi de 0.0038 e 0.0037, respectivamente.

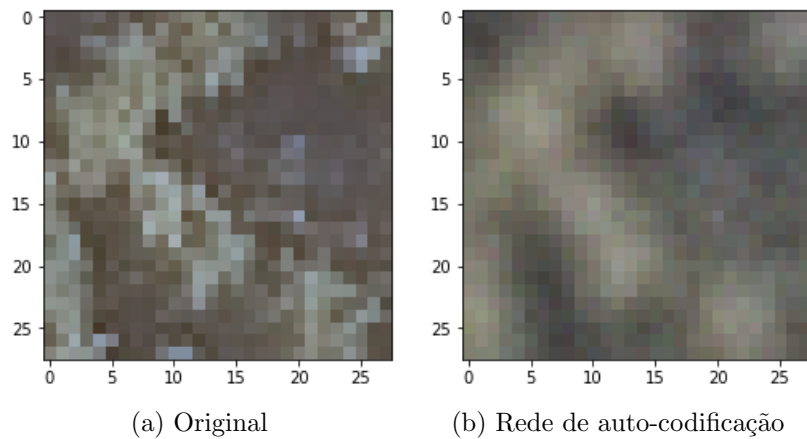


Figura 8: Comparação entre a figura original e o resultado obtido através da recuperação da codificação de 98,7% aproximadamente do dado de entrada.

3.5 Aplicação

A necessidade da compressão de imagens para as mais diversas finalidades, tais como videoconferência, armazenamento e plataformas web, vêm incentivando estudos no campo e com isso, lidando com avanços, superações de limitações e surgimento de novas abordagens. Dentre as inovações, pode-se citar aplicações com redes neurais, que vem alcançando bons resultados dentro do processamento de imagem e visão computacional (Jiang et al. 2017).

Atualmente, o Instituto Nacional de Pesquisas Espaciais (INPE) oferece ferramentas que trabalham com a aquisição de imagens, pode-se citar como exemplos o banco de imagens da DGI/INPE e o projeto Pauliceia 2.0 que utilizam sítios na internet para a disponibilização do conteúdo. Além disso, programas como o Satélite Sino-Brasileiro de Recursos Terrestres (CBERS), responsável por realizar coletas de dados para o monitoramento de fenômenos naturais na superfície terrestre, são designados para armazenar e enviar as imagens capturadas (Pesquisas Espaciais 2018).

Neste contexto, tendo em mente os resultados que podem ser obtidos e visando o aperfeiçoamento de abordagens com redes neurais em processos de compressão de dados, tem-se a importância da utilização de tal método, principalmente em aplicação voltadas a web, onde a representação degradada das imagens originais é tolerável.

Capítulo 4

Conclusão

A compressão de dados de satélite é essencial para o futuro das técnicas de sensoriamento remoto, dada a relação entre a velocidade de aquisição dos dados, contra a velocidade de transmissão. Desta maneira, é interessante realizar um estudo profundo nos algoritmos de próxima geração de formar a ter competitividade em relação aos projetos internacionais.

Devido aos altos custos de desenvolvimento de sistemas embarcados, abordagens que permitam atualizações, e manutenção são mais adequadas, sendo portanto, a FPGA como o dispositivo mais adequado para a implementação física das técnicas adotadas.

Finalmente, a adoção de especificações abertas e de tecnologias bem estabelecidas garante uma maior margem de segurança para o projeto.

As técnicas baseadas em redes de auto-codificação são as sementes para uma completa nova geração de algoritmos, e sua aplicação imediata, a nível de satélite não é adequada, todavia, possibilidades de utilização ao nível da internet irão surgir nos próximos anos.

Bibliografia

- [1] Paul R. Baumann. *HISTORY OF REMOTE SENSING, SATELLITE IMAGERY, PART II*. 2009. URL: <https://www.oneonta.edu/faculty/baumanpr/geosat2/RS%20History%20II/RS-History-Part-2.html> (acesso em 15/07/2018).
- [2] Toby Berger. “Rate Distortion Theory and Data Compression”. Em: *Advances in Source Coding*. Vienna: Springer Vienna, 1975, pp. 1–39. ISBN: 978-3-7091-2928-9. DOI: 10.1007/978-3-7091-2928-9_1. URL: https://doi.org/10.1007/978-3-7091-2928-9_1.
- [3] Chris Crawford. *DeepSat (SAT-4) Airborne Dataset*. 2018. URL: <https://www.kaggle.com/crawford/deepsat-sat4> (acesso em 26/08/2018).
- [4] E. Delp e O. Mitchell. “Image Compression Using Block Truncation Coding”. Em: *IEEE Transactions on Communications* 27.9 (set. de 1979), pp. 1335–1342. ISSN: 0090-6778. DOI: 10.1109/TCOM.1979.1094560.
- [5] Frederic Dufaux, Gary J. Sullivan e Touradj Ebrahimi. “The JPEG XR image coding standard [Standards in a Nutshell]”. Em: 26 (dez. de 2009), pp. 195–199, 204.
- [6] Steve Graham. *Remote Sensing: Introduction and History*. 1999. URL: <https://earthobservatory.nasa.gov/Features/RemoteSensing> (acesso em 15/07/2018).
- [7] G.E. Hinton e R.R. Salakhutdinov. “Reducing the Dimensionality of Data with Neural Networks”. Em: 313 (ago. de 2006), pp. 504–7.

BIBLIOGRAFIA

- [8] Tom Hope, Yehezkel S. Resheff e Itay Lieder. *Learning TensorFlow: A Guide to Building Deep Learning Systems*. 1st. O'Reilly Media, Inc., 2017. ISBN: 1491978511.
- [9] David A. Huffman. "A Method for the Construction of Minimum-Redundancy Codes". Em: *Proceedings of the Institute of Radio Engineers* 40.9 (set. de 1952), pp. 1098–1101.
- [10] Feng Jiang et al. "An end-to-end compression framework based on convolutional neural networks". Em: *IEEE Transactions on Circuits and Systems for Video Technology* (2017).
- [11] Tom M Mitchell et al. "Machine learning. 1997". Em: *Burr Ridge, IL: McGraw Hill* 45.37 (1997), pp. 870–877.
- [12] Andres Munoz. "Machine learning and optimization". Em: (2014).
- [13] J. B. O'Neal. "Predictive Quantizing Systems (Differential Pulse Code Modulation) for the Transmission of Television Signals". Em: *Bell System Technical Journal* 45.5 (), pp. 689–721. DOI: 10.1002/j.1538-7305.1966.tb01052.x. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/j.1538-7305.1966.tb01052.x>. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1002/j.1538-7305.1966.tb01052.x>.
- [14] Instituto Nacional de Pesquisas Espaciais. *CBERS-3 e 4*. 2018. URL: <http://www.cbers.inpe.br/sobre/cbers3-4.php> (acesso em 05/02/2018).
- [15] Moacir Antonelli Ponti e Gabriel B Paranhos da Costa. "Como funciona o Deep Learning". Em: *arXiv preprint arXiv:1806.07908* (2018).
- [16] D. Salomon. *Data Compression: The Complete Reference*. Springer Berlin Heidelberg, 2012. ISBN: 9783642860928. URL: <https://books.google.com.br/books?id=c90oCAAAQBAJ>.

BIBLIOGRAFIA

- [17] C. E. Shannon. “A Mathematical Theory of Communication”. Em: *Bell System Technical Journal* 27.3 (), pp. 379–423. DOI: 10.1002/j.1538-7305.1948.tb01338.x. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/j.1538-7305.1948.tb01338.x>. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1002/j.1538-7305.1948.tb01338.x>.
- [18] Claude E. Shannon. *Claude Elwood Shannon: Collected Papers*. Ed. por Neil J. A. Sloane e Aaron D. Wyner. Piscataway, NJ, USA: IEEE Press, 1993. ISBN: 0-7803-0434-9.
- [19] J. Sneyers e P. Wuille. “FLIF: Free lossless image format based on MANIAC compression”. Em: *2016 IEEE International Conference on Image Processing (ICIP)*. Set. de 2016, pp. 66–70. DOI: 10.1109/ICIP.2016.7532320.
- [20] Carlos Oscar Sánchez Sorzano, Javier Vargas e A Pascual Montano. “A survey of dimensionality reduction techniques”. Em: *arXiv preprint arXiv:1403.2877* (2014).
- [21] Xiaojin Zhu e Andrew B Goldberg. *Introduction to Semi-Supervised Learning: Morgan & Claypool*. 2009.