

Universidade Federal de Santa Catarina  
Departamento de Informática e Estatística  
Programa de Pós-Graduação em Ciência da Computação

Isaac Leonardo Santos Sacramento

Texto entregue como requisito para defesa do  
Exame de Qualificação de Doutorado, contendo re-  
visão bibliográfica, problemática, proposta e resul-  
tados prévios.

Orientador: Mauro Roisenberg

Florianópolis  
2017

# Resumo

O processo de caracterização de reservatórios de hidrocarbonetos consiste na determinação tridimensional e quantitativa da estrutura e das propriedades petrofísicas das rochas da área de interesse.

Palavras chave: Inversão Sísmica; Modelagem de Incerteza; Inversão Geoestatística; Redes Neurais Convolucionais.

# Abstract

The characterization process of hydrocarbon reservoirs entails in determining the 3D structure and petrophysical properties of the rocks at the area of interest.

Keywords: Seismic Inversion; Uncertainty Modeling; Geostatistical Inversion; Convolutional Neural Networks.

# Sumário

<b>1</b>	<b>Introdução</b>	<b>1</b>
1.1	Objetivo . . . . .	4
1.2	Organização do Texto . . . . .	5
<b>2</b>	<b>Fundamentação Teórica</b>	<b>6</b>
2.1	Problema Inverso . . . . .	6
2.2	Inversão Sísmica . . . . .	7
2.2.1	Aquisição Sísmica . . . . .	7
2.3	Redes Neurais Convolucionais . . . . .	9
2.3.1	Convolução . . . . .	10
2.3.2	Filtros . . . . .	13
2.3.3	Pooling . . . . .	15
2.3.4	Propriedades das Redes Convolucionais . . . . .	16
2.4	Resumo . . . . .	18
<b>3</b>	<b>Revisão da Literatura</b>	<b>19</b>
3.1	Métodos de Inversão Sísmica . . . . .	19
3.1.1	Inversão Sísmica Linear e Não Linear . . . . .	20
3.1.2	Máximo <i>a posteriori</i> . . . . .	21
3.2	Métodos de Super-resolução de Imagens . . . . .	23
3.2.1	Super-resolução por CNN . . . . .	23
3.3	Resumo . . . . .	26
	<b>Bibliografia</b>	<b>27</b>

# Capítulo 1

## Introdução

Um aspecto importante nas ciências físicas é poder inferir sobre parâmetros físicos a partir de dados. Em geral, as leis da física disponibilizam os artefatos necessários para calcular valores de dados, a partir de um modelo. Este procedimento é conhecido como problema direto (*forward problem*). A modelagem direta, portanto, inicia com um modelo, sobre o qual um experimento ou processo é simulado matematicamente. Se o modelo estiver correto, a resposta obtida deve parecer com dados reais. O processo de inversão faz exatamente o contrário, consiste em utilizar as medidas efetuadas para inferir os valores de parâmetros que caracterizam o sistema (Tarantola, 2005).

Para entender o problema inverso é conveniente explicar o problema direto antes. Considere o seguinte exemplo: suponha que uma pedra é lançada em um poço de água. Após determinado tempo um som é ouvido. É esperado que haja uma relação entre a profundidade do poço e o tempo entre soltar a pedra e ouvir o som do impacto. Da física, ocorre a existência de uma relação causa-efeito para este evento, dada por:

$$T = \sqrt{\frac{1}{5}H} + \frac{1}{340}H \quad (1.1)$$

onde a profundidade  $H$  é a causa e o tempo  $T$  é o efeito. Neste caso, o problema direto é calcular o tempo  $T$  para ouvir o som, dada profundidade  $H$ . A solução pode ser determinada inserindo um valor pra  $H$  na equação 1.1 e calcular o valor de  $T$ . O problema inverso é uma abordagem mais difícil, pois se deseja saber a profundidade  $H$ , dado apenas o tempo  $T$ .

No exemplo citado, ambos os problemas, direto e inverso, possuem solução. Entretanto, a maioria dos problemas inversos recai sobre duas características comuns que tornam sua solução não-trivial. Primeiro, a não unicidade de solução (problema não-determinístico), na qual

o mesmo conjunto de medidas observáveis pode resultar de mais de uma configuração de parâmetros. No exemplo citado, seria como obter a mesma altura  $H$  para diferentes tempos de queda  $T$  da pedra. Segundo, a natureza mal-posta do problema inverso, isto é, uma pequena mudança arbitrária nos valores observados pode causar uma mudança grande na solução fonte equivalente. Em um paralelo com o exemplo do poço, é como obter uma grande variação na profundidade, dado uma pequena variação no valor do tempo de queda.

Por conta da sua característica mal-posta, o problema inverso possui muitas soluções possíveis, de modo que representaria um equívoco considerar apenas uma solução como a mais correta. Via de regra, ao final do processo de inversão é comum realizar um processo de amostragem dentro do conjunto das possíveis soluções a fim de obter um estudo sobre elas. Este estudo pode ser uma análise de incerteza em torno da média de um conjunto de soluções do problema inverso.

O problema inverso possui um papel extremamente importante em diferentes áreas do conhecimento como Matemática, Medicina, Física e Geoestatística. Geoestatística é a aplicação de métodos estatísticos nas ciências da terra. Esta é a ciência que trata, dentre outros assuntos, da modelagem e caracterização de reservatórios, cujo tema é de amplo interesse para a indústria de óleo e gás. Por caracterização de reservatório se entende o processo para obter um modelo de propriedades petrofísicas (tais como, tipos de contato entre rochas, porosidade e permeabilidade), em 3-D e alta resolução, que seja consistente com os dados de que se dispõe Deutsch (2002). O processo de caracterização de reservatórios possui diferentes etapas que podem ser descritas em alto nível como na abordagem sequencial a seguir:

1. A primeira etapa envolve definição da geometria e estratigrafia dos intervalos do reservatório a ser modelado. Ainda, o desenvolvimento de um modelo conceitual de continuidade para propriedades como fácies, porosidade e permeabilidade .
2. Modelagem dos tipos de contato entre rochas, mais conhecidos como *facies*.
3. Modelagem da propriedade porosidade com base nos tipos de facies. A modelagem de porosidade costuma ser realizada antes da permeabilidade, devido à disponibilidade de dados sísmicos e dados amostrais localizados, também chamados de poços.
4. Os modelos 3-D para permeabilidade são atrelados à porosidade e facies anteriormente estabelecidos.

5. Múltiplas realizações, igualmente prováveis, são realizadas por repetição de todo o processo. Embora todas as realizações sejam equiprováveis, há realizações mais similares a outras, de modo que a classe à qual pertencem possui maior probabilidade.
6. Os modelos são usados como entrada em um simulador ou visualizados e usados como suporte na tomada de decisão.

Embora a abordagem anterior utilize alguns termos que, à primeira vista, pareçam incompreensíveis, sua apresentação contextualiza o ponto do processo de modelagem de reservatório em que a inversão sísmica acontece (etapa 3). Porque os dados medidos são obtidos por sísmica de reflexão, este método é chamado de inversão sísmica. A inversão sísmica na modelagem de reservatórios disponibiliza artefatos que são modelos de propriedades de rocha (propriedades petrofísicas) a partir principalmente, mas não exclusivamente, da sísmica disponível e de modelos construídos com dados amostrais. Tais artefatos são visualizados na forma de imagens, de modo que é possível supor que quanto maior o nível de resolução destas imagens, mais contundente será a justificativa para a tomada de decisão (etapa 6).

Atualmente, em Aprendizagem de Máquina, *Deep Learning* é o tema em maior evidência. *Deep Learning* é toda solução que permite aos computadores aprender a partir da experiência e entender o mundo em termos de hierarquia de conceitos. Assim, nos algoritmos de *Deep Learning*, o aprendizado por experiência (supervisionado) evita a interferência humana no sentido de especificar formalmente o conhecimento que o computador necessita. Adicionalmente, a hierarquia de conceitos permite aos computadores aprenderem conceitos complicados a partir de conceitos mais simples (Goodfellow; Bengio e Courville, 2016).

Muitos algoritmos representam o estado da arte em diferentes áreas como processamento de imagens, sinais e linguagem natural, e reconhecimento de padrões. No campo de processamento de imagens o método de *Deep Learning* de maior destaque nos dias atuais é conhecido como Redes Neurais Convolucionais (CNN). Seu surgimento data da década de 1980, com aplicação essencialmente no reconhecimento de imagens. Entretanto, com o advento das Unidades Gráficas de Processamento (GPU) e a maior disponibilidade de dados para treinamento, as redes convolucionais são empregadas com sucesso em serviços de busca de imagens, carros auto-dirigíveis, sistemas de classificação de imagens em video, entre outras aplicações complexas (Buduma, 2015, p. 50).

O processo de super-resolução é um método de processamento de imagens e visão computacional. Consiste em recuperar uma imagem de alta resolução a partir de uma imagem de baixa

resolução. Semelhante ao problema da inversão sísmica, o problema da super-resolução também é mal-posto, uma vez que pode haver múltiplas soluções em alta resolução para uma dada imagem em baixa resolução. Em ambos os problemas, uma forma de lidar com esta questão é restringir o espaço de soluções com informações *a-priori*. Considere  $Y$  uma imagem de baixa resolução, por exemplo, uma imagem interpolada. O objetivo da super-resolução é recuperar, a partir de  $Y$ , uma imagem  $F(Y)$  que é o mais similar possível a uma imagem de alta resolução  $X$  considerada como imagem que representa a verdade. De um ponto de vista conceitual, o mapeamento  $F$  consiste essencialmente em realizar três operações:

1. Para a imagem de baixa resolução  $Y$ , extração de mapas de características. Um mapa de características pode ser imaginado como um conjunto de sub-imagens com determinadas características da imagem original.
2. Mapeamento não-linear do mapa de características. Após a operação não-linear, cada característica mapeada passa a ter uma representação em alta resolução.
3. Reconstrução da imagem a partir das representações de alta resolução citadas. Esta nova imagem deve ser similar à imagem que representa a verdade  $X$ .

Um modelo de rede neural convolucional pode ser capaz de realizar as etapas listadas anteriormente, de forma iterativa e para um conjunto de diferentes imagens. Esta abordagem pode ser incorporada ao final do processo de inversão, de modo que as imagens obtidas da inversão sísmica sejam pós-processadas pelo modelo convolucional a fim de obter imagens com maior riqueza de detalhes. Como mencionado, o ganho de resolução nas imagens de propriedades petrofísicas pós-inversão pode conferir maior confiabilidade na interpretação da solução inversa e, consequentemente na tomada de decisão.

A seção seguinte apresenta os objetivos deste trabalho. Nos capítulos posteriores serão apresentados os aspectos físicos e de implementação. Serão apresentados também os resultados preliminares da pesquisa.

## 1.1 Objetivo

Este trabalho ataca a problemática da super-resolução dos artefatos da inversão sísmica baesiana por meio de um modelo de *Redes Neurais Convolucionais*. A abordagem se dá através da incorporação deste modelo de rede neural á última etapa do processo de inversão



sísmica bayesiana, as realizações, para alcançar um maior nível de resolução das imagens de propriedades petrofísicas. Resultados prévios indicam que o modelo baseado em redes neurais convolucionais é capaz de agregar informações de alta frequência às inversões sísmicas.

Outro objetivo deste trabalho é desenvolver um modelo baseado em redes neurais convolucionais que permita a realização de simulação geoestatística multiponto. Esta etapa de trabalho será desenvolvida em cooperação com o Departamento de Ciências Geológicas, Universidade Stanford, sob orientação do Prof. Dr. Jef Karel Caers.

## **1.2 Organização do Texto**

Este documento está organizado da seguinte forma. Após esta breve introdução, o Capítulo 2 apresenta a fundamentação teórica para os processos de inversão sísmica, as redes neurais convolucionais e a super-resolução. O Capítulo 3 apresenta o estado da arte em geração de imagens em super-resolução a partir de imagens de baixa resolução, bem como as justificativas para a realização desta pesquisa. O Capítulo 4 trata da implementação desta proposta e dos resultados preliminares referente ao modelo de super-resolução treinado e aplicado às imagens de impedância pós-inversão.

# Capítulo 2

## Fundamentação Teórica

Neste capítulo serão apresentados os conceitos abordados neste trabalho. O problema inverso será apresentado em linhas gerais e a inversão sísmica será abordada em maiores detalhes. Serão apresentados os conceitos relacionados a *Deep Learning*, assim como os elementos de redes neurais convolucionais. Esta fundamentação teórica é relevante para o entendimento de como o modelo de rede neural convolucional pode ser adotado para obter ganho qualitativo e quantitativo no pós-processamento da inversão sísmica.

### 2.1 Problema Inverso

A teoria de inversão é utilizada em diversas áreas para inferir os valores de parâmetros relacionados com processos físicos a partir de um conjunto de dados medidos, os quais são chamados dados experimentais. É possível descrever o problema inverso como o processo de obter informações de um sistema parametrizado, a partir de dados que podem ser medidos por meio de algum experimento físico e das relações teóricas com os parâmetros desejados, mas que não são passíveis de medição. Frequentemente, algum conhecimento *a priori* é incorporado ao modelo.

Um sistema físico depende do domínio em estudo. Pode ser uma galáxia para um astrofísico, pode ser a Terra para um geofísico ou uma partícula quântica para um físico quântico. Em comum, o fato de que, para ser estudado, um sistema físico segue três passos básicos: a parametrização do sistema, a modelagem direta e a modelagem inversa (Tarantola, 2005). A parametrização do sistema se refere à definição do conjunto mínimo de elementos (parâmetros, variáveis) cujos valores caracterizam completamente o sistema.

A modelagem direta significa prever os valores dos parâmetros observáveis (dados  $d$ ), que

correspondem a um dado modelo (conjunto de parâmetros  $m$ ). Esta predição pode ser denotada pela Eq. 2.1. Onde  $F(.)$  é chamado operador direto.

$$F(m) = d \quad (2.1)$$

Por sua vez, a modelagem inversa se refere ao uso de resultados atuais das medições dos parâmetros físicos observáveis, para inferir os valores atuais dos parâmetros do modelo (não-observáveis). O problema inverso pode ser descrito em uma forma discreta como:

$$m = F^{-1}(d) \quad (2.2)$$

onde,  $F$  é o sistema físico investigado, e relaciona os parâmetros do modelo  $m = (m_1, m_2, \dots, m_n) \subset R^n$  estimado com os dados observados  $d \in R^s$ . Como mencionado no Capítulo 1, um problema inverso possui múltiplas soluções, de modo que o modelo  $m$  pertence a um conjunto de modelos  $M$  admissíveis. Na prática,  $d$  pode ser uma função no domínio do tempo e/ou espaço, ou pode ser uma coleção de observações discretas.

## 2.2 Inversão Sísmica

Os métodos geofísicos frequentemente envolvem a solução e avaliação de problemas inversos, pois permitem inferir a distribuição das propriedades físicas na subsuperfície da Terra usando observações da superfície. A inversão sísmica tem um papel fundamental na solução de problemas geofísicos, em especial na caracterização de reservatórios (Bosch; Mukerji e Gonzalez, 2010; Srivastava e Sen, 2009). Do ponto de vista prático, as soluções para o problema de inversão sísmica melhoram a exploração e o gerenciamento na indústria petrolífera, uma vez que os dados sísmicos estimados possuem forte correlação com as propriedades petrofísicas (porosidade, densidade, etc.) das rochas da subsuperfície (Passos de Figueiredo et al., 2014). Para facilitar o entendimento da inversão sísmica, considere a subsuperfície como sendo formada por camadas sobrepostas de diferentes tipos de rochas. As regiões onde ocorrem as transições entre tipos diferentes de rochas são comumente chamadas de *facies* e podem ter espessuras diferentes.

### 2.2.1 Aquisição Sísmica

O dado sísmico é o principal parâmetro observável utilizado na inversão sísmica. A aquisição destes dados se dá por meio da sísmica de reflexão. Este método utiliza pulsos sísmicos de uma

fonte artificial controlada e monitora a resposta em função do tempo. Neste sistema, cada região de contato entre dois tipos de rochas diferentes gera reflexão e refração do pulso sísmico, como demonstrado na Figura 2.1. De um ponto de vista bastante elementar, é possível intuir que a parte refletida da onda se propaga em todas as direções, de modo que os componentes horizontal e vertical podem ser medidos. O componente horizontal (*s-wave*), referente à reflexão horizontal da onda, é utilizado no processo de inversão conhecido como inversão elástica. Por outro lado, o componente vertical da onda (*p-wave*), referente à reflexão vertical do pulso emitido, é utilizado no processo conhecido como inversão acústica.

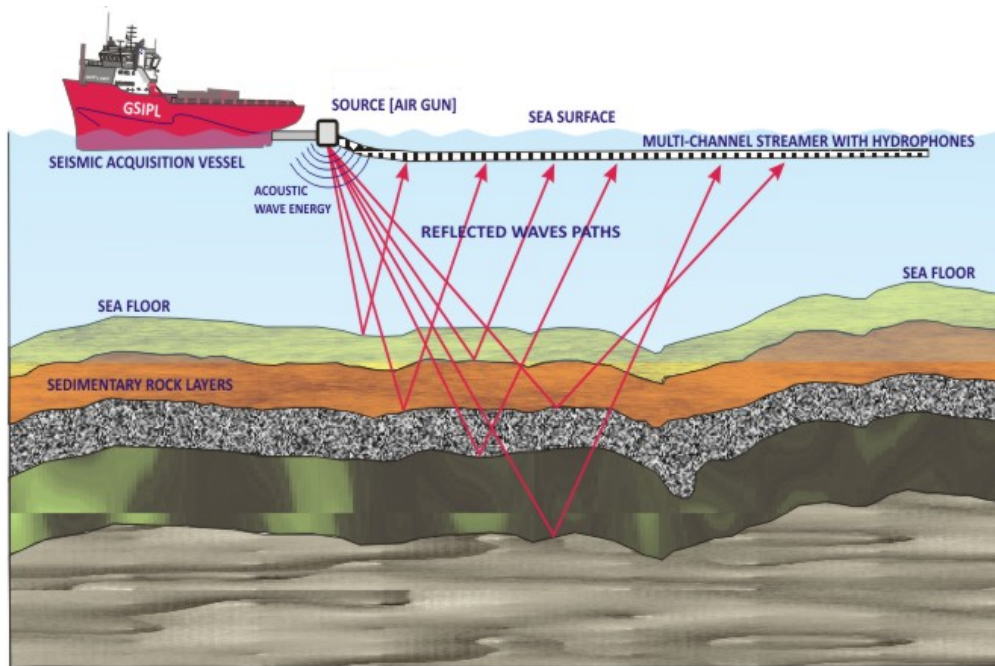


Figura 2.1: Método de sísmica de reflexão (GSIP, 2017)

O pulso de onda emitido durante a aquisição possui um formato próprio, uma identidade, conhecido como *wavelet*. Assim, a resposta sísmica medida é composta em parte por esta identidade e, em parte, pela característica da interface entre duas camadas de rochas diferentes, na qual o pulso reflete. Esta característica é chamada de coeficiente de refletividade (equação 2.3):

$$r(t) = \frac{z(t + \delta t) - z(t)}{z(t + \delta t) + z(t)} \quad (2.3)$$

onde,  $z(t)$  é a impedância acústica no tempo  $t$  definida por  $z(t) = \rho(t)v(t)$ , onde  $\rho(t)$  é a densidade da rocha e  $v(t)$  a velocidade de propagação da onda acústica. O dado sísmico utilizado na inversão acústica, portanto, é uma aproximação da resposta da camada terrestre. Pode ainda

ser definido como a convolução entre a *wavelet* de aquisição e o valor de refletividade entre as camadas, com ângulo de incidência e reflexão de  $90^\circ$ , respectivamente. Por este motivo, este modelo é chamado convolucional. Com os coeficientes de reflexão e a discretização da medida de tempo, é possível modelar o dado sísmico  $d(t)$  aplicando a convolução  $\otimes$  da *wavelet*  $s$  com os coeficientes de refletividade  $r$ :

$$d(t) = s(\tau) \otimes \sum_{j=1}^N r(t - t_j) \delta(t - t_j) + e_d(t) \quad (2.4)$$

onde  $N$  é o número total de camadas,  $e_d(t)$  representa o ruído aleatório em função do tempo e cada  $d_{xy}$  é chamado de traço sísmico. Um conjunto de traços sísmicos também é chamado de imagem, seção ou cubo, no caso de um levantamento 3D. A *wavelet* ideal seria um pulso tipo delta contendo todas as frequências, entretanto, na prática as *wavelets* são pulsos de banda limitada entre  $6Hz$  e  $65Hz$ , o que limita a frequência da sísmica e sua resolução (Sen, 2006, p. 11). Como consequência, as imagens resultantes do processo de inversão também terão o seu espectro de frequência limitado. A Figura 2.2 ilustra uma *wavelet* típica extraída de dados reais.

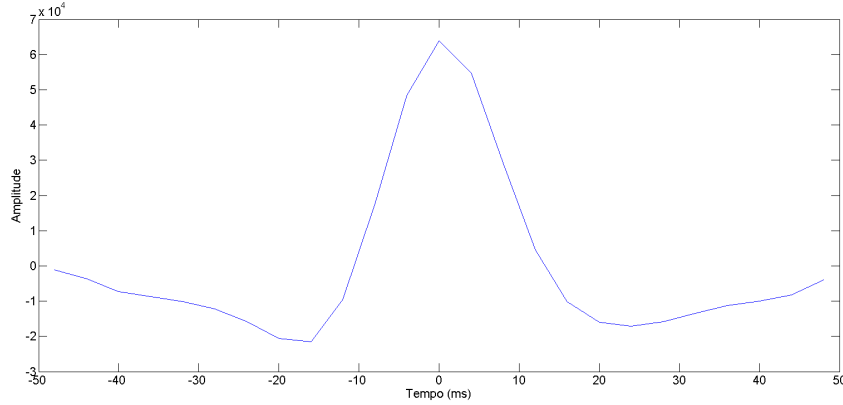


Figura 2.2: *Wavelet* extraída de dados reais

## 2.3 Redes Neurais Convolucionais

As Redes Neurais Convolucionais (CNN), também chamadas de redes convolucionais, são um tipo de rede neural especializada em processamento de dados que possuem uma topologia conhecida e em forma de grade (Goodfellow; Bengio e Courville, 2016). Exemplos deste tipo de dado são as séries temporais, que podem ser vistas como uma grade em uma dimensão (1-D)

com amostras em intervalos regulares de tempo, e dados de imagem, que podem ser vistos como uma grade (2-D) de *pixels*. Este modelo de rede neural é chamada convolucional, pois emprega a operação de convolução no lugar de multiplicação comum entre matrizes, em pelo menos uma de suas camadas.

### 2.3.1 Convolução

A operação de convolução é definida como a integral do produto de duas funções após uma delas sofrer um certo deslocamento. Considere o exemplo em que se deseja rastrear a localização de uma nave espacial com um sensor a laser. O sensor disponibiliza uma saída  $x(t)$  referente à posição da nave no tempo  $t$ . Ambos,  $x$  e  $t$ , são valores reais, de modo que uma saída diferente pode ser obtida em qualquer instante de tempo. Considerando que o sensor possui um certo ruído, para realizar uma estimativa mais precisa da posição da nave é preciso ponderar várias medidas de posição juntas. Como os valores medidos mais recentemente são mais relevantes, se estima uma função peso  $w(a)$ , onde  $a$  é o tempo de medição. Se esta média ponderada for aplicada a todos os instantes, a estimativa de posição da nave será suavizada:

$$s(t) = \int x(a)w(t-a)da \quad (2.5)$$

Esta operação é chamada convolução e pode ser definida para quaisquer funções, às quais a integral da equação 2.5 esteja definida. A convolução costuma ser denotada com um asterisco e aplicada com o tempo  $t$  discretizado para valores inteiros:

$$s(t) = (x * w)(t) = \sum_{a=-\infty}^{\infty} x(a)w(t-a) \quad (2.6)$$

No contexto das redes convolucionais,  $x$  se refere ao conjunto de imagens de entrada e  $w$  é denominado *kernel* ou filtros. As imagens de entrada são uma sequência multidimensional de dados, enquanto os filtros são uma sequência multidimensional de parâmetros a serem otimizados pelo algoritmo de aprendizagem. Nos casos em que o problema compreende imagens  $X$  e filtros  $W$  utilizados em duas dimensões a convolução ganha o seguinte formato:

$$S(i, j) = (X * W)(i, j) = \sum_m \sum_n X(m, n)W(i-m, j-n) \quad (2.7)$$

Nas redes convolucionais há pelo menos duas estruturas básicas, a camada convolucional e a camada de *pooling*. A arquitetura típica de uma CNN compreende duas camadas convolucionais,

cada uma é seguida por uma camada *pooling*, como ilustrado na Figura 2.3. À medida que as imagens progridem ao longo da rede, suas dimensões diminuem, entretanto, elas se tornam mais profundas em termos de hierarquia de conceitos extraídos. No topo da pilha de camadas da rede se adiciona camadas completamente conectadas, sendo que, na última camada ocorre a saída prevista. Esta estrutura de camadas completamente conectadas é a mesma utilizada nas redes neurais tradicionais do tipo *feedforward*, nas quais todos os neurônios de uma camada estão conectados a todos os neurônios da camada seguinte.

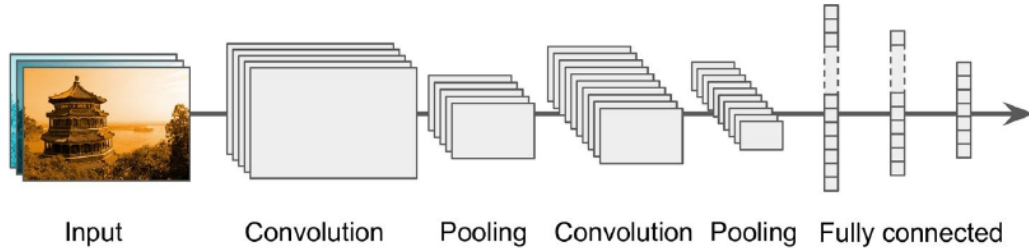


Figura 2.3: Arquitetura típica de uma rede neural convolucional. Fonte:(Géron, 2017)

A camada convolucional é o elemento mais importante de uma CNN. Esta camada é estruturada de modo a fazer com que cada um dos seus neurônios esteja conectado a um pequeno grupo de *pixels* da camada de entrada (Figura 2.4) e não a todos os pixels, como ocorre em redes neurais tradicionais. Cada neurônio da camada seguinte se conecta apenas aos neurônios contidos em uma pequena região da camada anterior e assim sucessivamente. Esta região que define o grupo de neurônios conectados ao neurônio da próxima camada é chamada **campo perceptivo**. Este formato permite o aprendizado de características de baixo nível na primeira camada e de características de mais alto nível nas camadas seguintes.

A Figura 2.5 ilustra a conexão entre as camadas de uma rede convolucional. Considere um neurônio localizado na linha  $i$  e coluna  $j$  de uma dada camada. Este neurônio estará conectado às saídas dos neurônios da camada anterior localizados nas linhas  $\times i s_h$  até  $\times i s_h + f_h - 1$ , colunas  $\times j s_w$  até  $\times j s_w + f_w - 1$ , onde  $f_h$  e  $f_w$  são a altura e a largura do campo receptivo,  $s_h$  e  $s_w$  são os deslocamentos vertical e horizontal ao longo das imagens da camada anterior. O tamanho destes deslocamentos é chamado de passo ou *stride* e quanto maior o *stride*, menor será a imagem resultante na camada seguinte. Para *stride* de tamanho 0 a camada seguinte terá as mesmas dimensões da camada anterior.

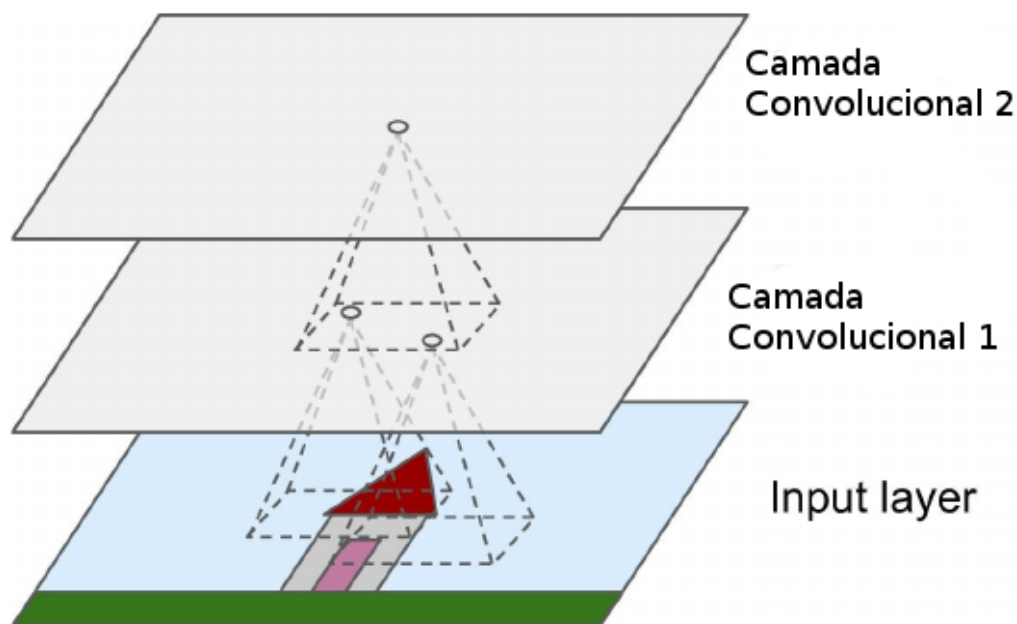


Figura 2.4: Camadas de uma CNN com campos receptivos retangulares.

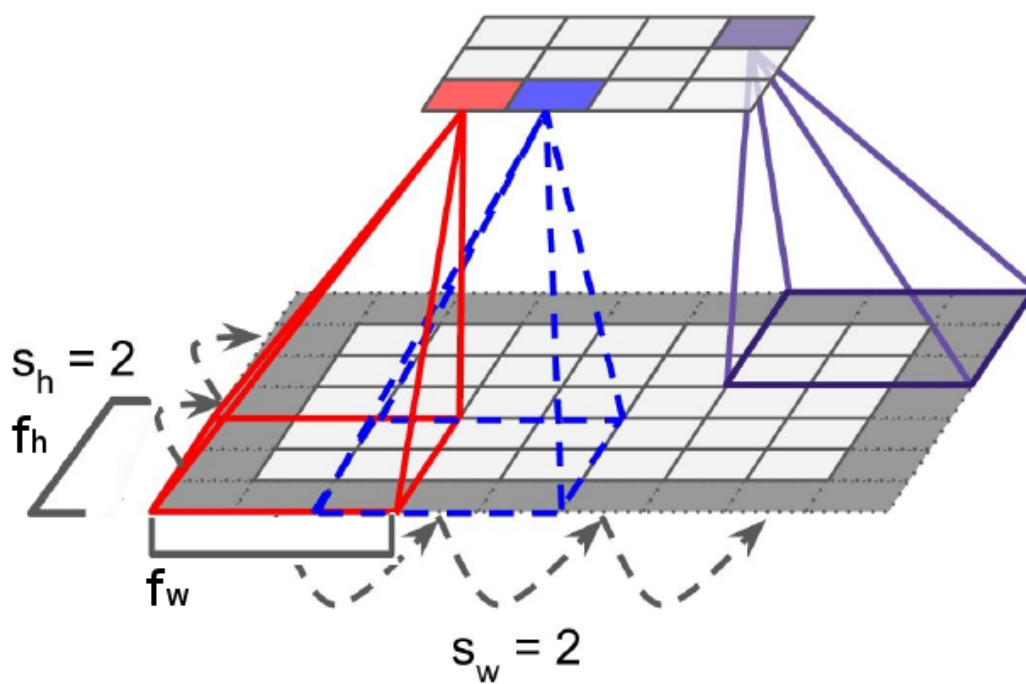


Figura 2.5: Conexão entre camadas com campo receptivo 3 x 3 e *strides* de tamanho 2.



### 2.3.2 Filtros

Os filtros (pesos) em uma camada convolucional são representados como uma pequena imagem com as mesmas dimensões do campo receptivo. São eles os elementos convolvidos com a imagem de entrada para obter o resultado da camada convolucional. A Figura 2.6 ilustra dois conjuntos de filtros possíveis. O primeiro filtro é um quadrado preto (*pixels* de valor 0) contendo uma coluna central branca (*pixels* com valor 1). Analogamente, o segundo filtro é um quadrado preto contendo uma linha central branca. É possível notar na imagem da esquerda que as linhas verticais brancas se tornaram mais evidentes, enquanto as outras partes da imagem se tornaram mais borradas. Igualmente, na imagem da direita a convolução com o filtro horizontal destacou as linhas brancas horizontais, ao passo que o restante ficou borrado. Assim, uma característica detectada por um neurônio representa o tipo de padrão da entrada que causará a sua ativação. Estes padrões podem ser bordas, contornos ou estruturas com outras formas.

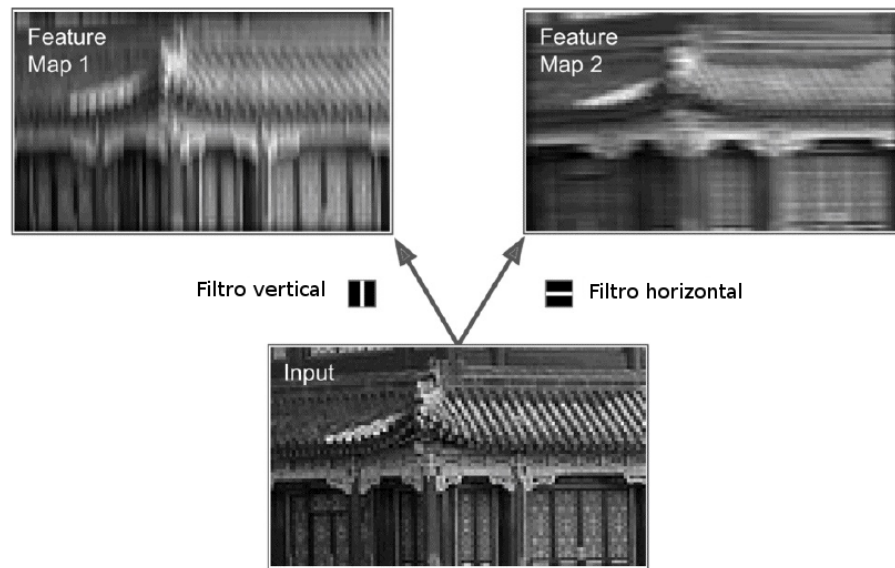


Figura 2.6: Aplicação de dois filtros diferentes para obter mapas de características.

Em situações reais, a camada convolucional possui muitos mapas de características, resultando em uma representação em 3-D como ilustrado na Figura 2.7. Os mapas de características de uma camada convolucional são o resultado da convolução de uma das imagens de entrada com os diversos filtros específicos desta camada. Na Figura 2.7 estão ilustrados os mapas para a convolução com apenas uma imagem, de modo que é possível imaginar que, à medida que o número de imagens aumenta, a estrutura ilustrada se replica horizontalmente.

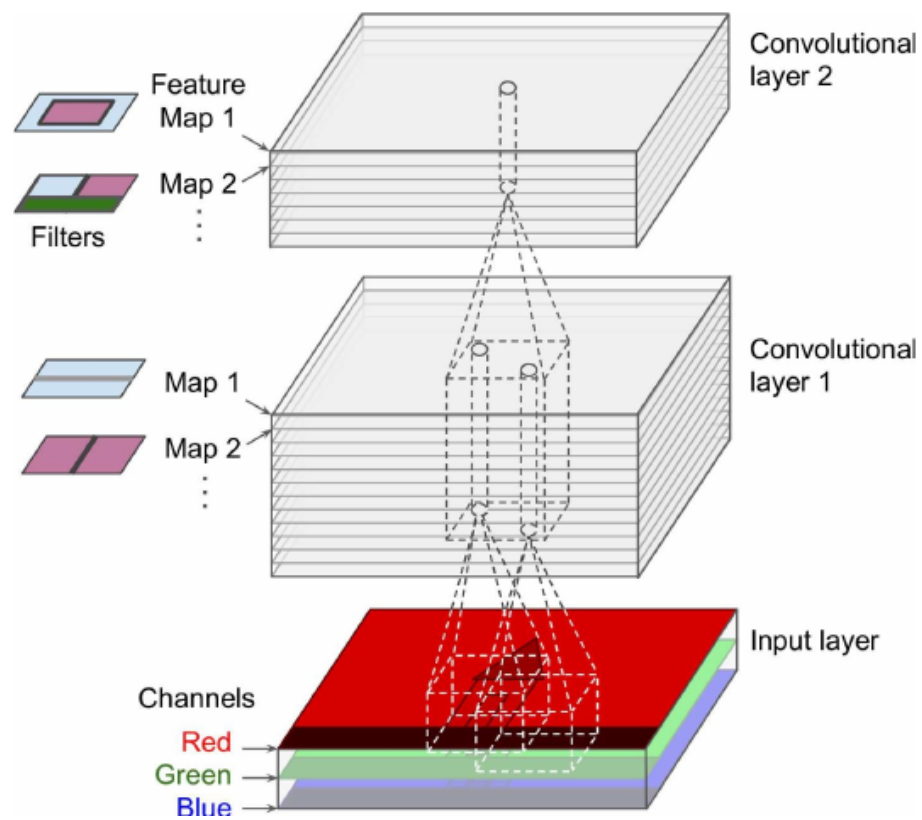


Figura 2.7: Camadas convolucionais com múltiplos mapas de características e imagens com três canais.

### 2.3.3 Pooling

O processamento ao longo de uma rede convolucional ocorre em três estágios. No primeiro estágio, acontecem as convoluções entre as imagens e os filtros para produzir um conjunto de ativações lineares. O segundo estágio é chamado etapa de detecção, na qual cada ativação é submetida a uma função não-linear. A terceira etapa é chamada de *pooling*, responsável por modificar a saída da camada convolucional para obter um sumário estatístico das saídas em uma determinada vizinhança. Semelhante ao que ocorre na convolução, a região sobre a qual se aplica *pooling* é definida por um campo receptivo e o deslocamento é definido por um *stride*. O *pooling* permite tornar invariante pequenas translações no conjunto de entrada, ou seja, ainda que haja pequenas translações na entrada, os valores da maioria das saídas após o *pooling* permanecem iguais. A Figura 2.8 ilustra o funcionamento da função de *pooling* máximo, na qual o máximo valor de ativação dentro de uma vizinhança é selecionado. Outras funções de *pooling* incluem o valor médio dentro de uma região retangular, a normalização  $L^2$  de uma vizinhança, ou a média ponderada baseada na distância do *pixel* central.

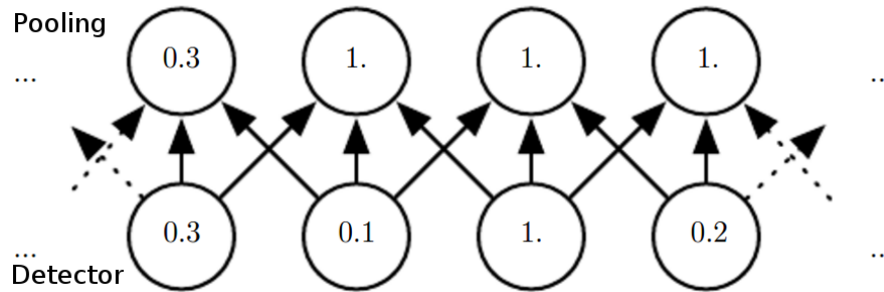


Figura 2.8: Operação de *pooling* com campo receptivo de tamanho 3. Nesta operação é selecionado o máximo valor de ativação da etapa de detecção.

A propriedade de invariância é útil quando a existência de uma característica é mais relevante que o local exato onde ela ocorre. Por exemplo, para determinar se o rosto de uma pessoa ocorre em uma certa imagem, não é necessário saber com precisão o local dos olhos, basta saber se há um olho do lado esquerdo do rosto e outro olho do lado direito <sup>1</sup>. Por outro lado, há contextos em que o local da característica é uma informação relevante e deve ser preservada. Em modelagem de reservatórios, a detecção de bordas referentes a uma facie selante sobre uma região de reservatório. Adicionalmente, a operação de *pooling* permite lidar com entradas de tamanho variável. Por exemplo, na classificação de imagens as entradas para a camada

<sup>1</sup>O rosto da figura pública Nestor Cerveró, por exemplo, seria facilmente identificável por uma rede convolucional com uma camada de *pooling*

de classificação devem ter o mesmo tamanho. Assim, o *stride* entre regiões de *pooling* pode variar para que a camada de classificação receba o mesmo número de sumários estatísticos, independente do tamanho das imagens.

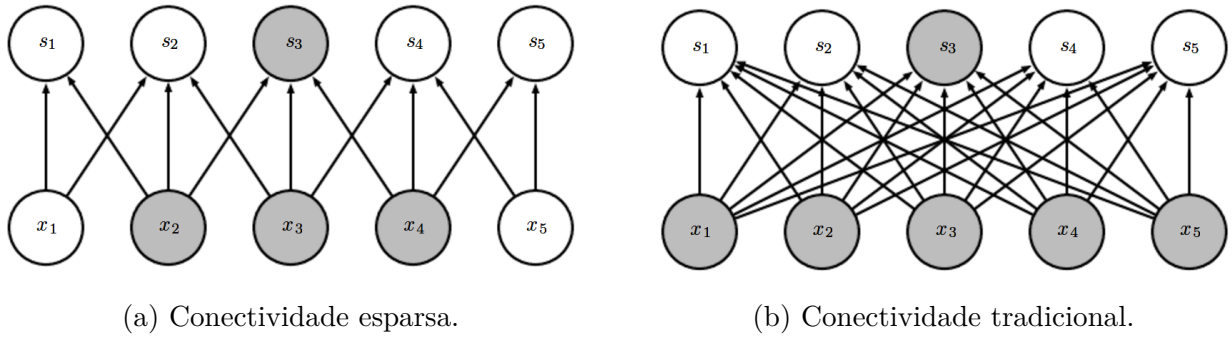
As função de *pooling* sumariza as respostas de vizinhanças separadas por  $k$  *pixels*, por isso, o tamanho do seu campo receptivo é menor que o campo receptivo da convolução. Isto aumenta a eficiência computacional da rede, pois a camada seguinte à *pooling* terá  $k$  vezes menos entradas para processar. Quando o número de filtros da camada seguinte é função do tamanho da sua entrada, a redução promovida pela função de *pooling* pode resultar em maior eficiência estatística e redução da quantidade de memória.

### 2.3.4 Propriedades das Redes Convolucionais

Por conta da sua arquitetura, as redes convolucionais se sustentam sobre três pilares: interações esparsas, compartilhamentos de parâmetros e representações equivariantes. As propriedades de interação esparsa e compartilhamento de pesos serão apresentadas com maior nível de detalhes nesta seção, embora já tenham sido introduzidos de forma intuitiva nas seções anteriores.

As **interações esparsas**, também chamadas de conectividade esparsa ou pesos esparsos, ocorre quando os filtros possuem dimensão menor que a entrada, ou seja a dimensão do campo receptivo é menor que a dimensão das imagens de entrada. De um ponto de vista prático, a imagem de entrada pode ter milhares de *pixels*, entretanto, é possível detectar apenas pequenas regiões com características de maior relevância na imagem de entrada com filtros que compreendam apenas algumas dezenas ou centenas de *pixels*. Por exemplo, é possível identificar características de uma face humana no reconhecimento de pessoas, ou estruturas com significado geológico em um estudo geofísico. Como consequência, menos parâmetros são armazenados e há um ganho na eficiência estatística do modelo. As Figuras 2.9a e 2.9b ilustram os modelos de conectividade esparsa e tradicional, respectivamente. É possível notar que na conectividade tradicional (Figura 2.9b) todos os elementos da camada inferior afetam o elemento em destaque  $s_3$  da camada seguinte, enquanto na conectividade esparsa (Figura 2.9a) apenas três elementos afetam o elemento em destaque. O número de elementos que afetam o elemento em destaque na conectividade esparsa é definido pelo tamanho do filtro utilizado na convolução.

O **compartilhamento de parâmetros**, também chamado de **pesos amarrados**, se refere ao uso do mesmo parâmetro para mais de uma função no modelo. Como já mencionado, nas redes neurais tradicionais todos os neurônios de uma camada são conectados a todos os neurônios



da camada anterior e cada neurônio possui um *bias*, como ilustrado na imagem 2.10. Entretanto, este modelo é pouco eficiente, pois não tira vantagem da estruturas espaciais das imagens de entrada (Goodfellow; Bengio e Courville, 2016). Estas informações estruturais são muito

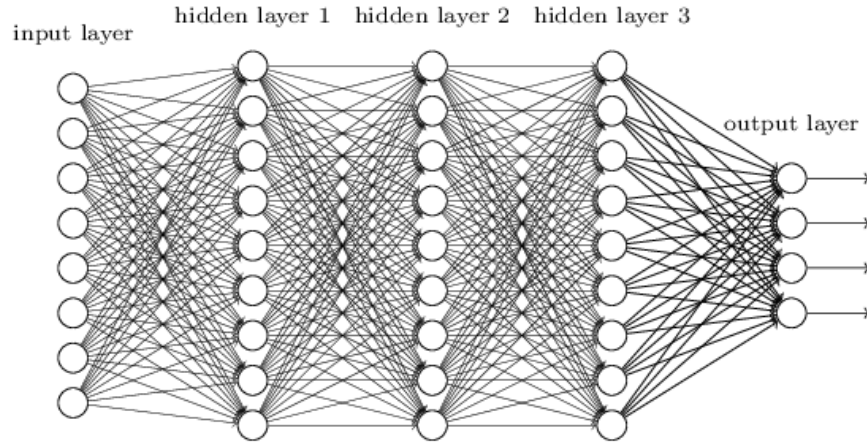


Figura 2.10: Organização de camadas de uma rede neural do tipo *feedforward*.

relevantes quando o problema em estudo é geostatístico. Por outro lado, no compartilhamento de pesos a saída de cada neurônio de uma camada depende apenas do conjunto de neurônios de uma pequena região definida pelo campo receptivo da camada anterior:

$$\sigma \times \left( b + \sum_m \sum_n w_{m,n} a_{i+m,j+n} \right) \quad (2.8)$$

onde,  $\sigma$  é uma função de ativação,  $b$  é o valor compartilhado do *bias*,  $w_{m,n}$  é uma matriz de pesos compartilhados (filtros) e  $a_{i+m,j+n}$  denota a entrada  $a_{x,y}$  na posição  $x, y$ . Como o mesmo filtro é convolucionado ao longo da imagem, os mesmos pesos e *bias* aprendem diferentes características da imagem. Deste modo, cada conjunto de pesos e *bias* é compartilhado por diferentes regiões em cada imagem e o número de pesos conectados ao neurônio da camada seguinte diminui em relação ao modelo tradicional. Isto faz com que a convolução seja mais eficiente que a

multiplicação de matriz do ponto de vista de requisitos de memória e eficiência estatística.

O compartilhamento de pesos confere às redes convolucionais a propriedade de **equivariância** de translação. Se uma função é equivariante, significa que se a entrada muda, a saída muda igualmente. Matematicamente, a função  $f(x)$  é equivariante à função  $g$  se  $f(g(x)) = g(f(x))$ . No caso da convolução, se  $g$  é uma função que translada a entrada, então a convolução será equivariante a  $g$ . A convolução com imagens cria um mapa 2-D dos locais onde certas características aparecem na entrada. A propriedade de equivariância permite rastrear objetos transladados na entrada. Se um objeto aparece em uma determinada posição e, em seguida, aparece em outra posição, sua representação irá mover a mesma quantidade na saída. É importante frisar que, nas CNN, a propriedade de equivariância é aplicável apenas para a translação, de modo que a convolução não é equivariante para transformações de escala e rotações na imagem.

## 2.4 Resumo

Este Capítulo detalhou os principais conceitos abordados neste trabalho. O problema inverso foi introduzido e a inversão sísmica apresentada em maiores detalhes. Foram apresentados os elementos que compõem as redes neurais convolucionais: a convolução, as camadas convolucionais, os filtros e a camada de *pooling*. Foram apresentadas também as propriedades das camadas convolucionais: conectividade esparsa, compartilhamento de parâmetros e equivariância de translação.

# Capítulo 3

## Revisão da Literatura

Neste capítulo serão apresentadas as revisões sistemáticas relacionadas ao método de inversão acústica e dos modelos de super-resolução de imagens. Esta revisão evidenciou o potencial de pesquisa desta proposta, pois apresenta uma lacuna em métodos de pós-processamento da inversão sísmica.

### 3.1 Métodos de Inversão Sísmica

É importante ter em mente que, durante a inversão, as operações são realizadas sobre dois espaços de representações diferentes: o espaço do modelo e o espaço de dados. No contexto da inversão sísmica, os dados sísmicos  $d$  são representados no espaço dos dados e a propriedade de impedância acústica das rochas é representada no espaço do modelo  $m$ . A escolha dos parâmetros do modelo geralmente é não única, de modo que dois conjuntos de parâmetros diferentes podem ser equivalentes. Entretanto, para uma abordagem quantitativa, uma parametrização precisa ser definida Tarantola (2005) e, no contexto da inversão acústica, o parâmetro adotado é a impedância acústica. Para obter informações sobre os parâmetros do modelo, é necessário realizar observações através de experimentos físicos, como por exemplo, a aquisição sísmica. Este conjunto de dados representa os parâmetros observáveis do sistema, ou espaço de dados, e representa o ponto de partida para a inversão.

Sob um olhar ingênuo é possível questionar por quê não definir a função inversa da modelagem direta e calcular, de forma imediata, os parâmetros do modelo a partir dos espaço dos dados. No entanto, os métodos de inversão direta sofrem de instabilidades devido ao ruído e características do problema (Sen, 2006, p. 50). Outra opção é utilizar tentativa e erro para ajustar os parâmetros até conseguir uma resposta semelhante aos dados experimentais. Formal-

mente isto é automatizado utilizando métodos de otimização. Para tanto, é preciso definir uma função de custo, ou função objetivo, que mede o ajuste dos dados produzidos pelos parâmetros do modelo (dato sintético) ao dato medido.

### 3.1.1 Inversão Sísmica Linear e Não Linear

O conteúdo desta seção apresenta as suposições de linearidade necessárias que tornam a inversão acústica um processo analítico e computacionalmente eficiente. O detalhamento matemático pode ser consultado em Passos de Figueiredo et al. (2014).

Para entender o processo de inversão sísmica, é conveniente ter em mente que os problemas inversos podem ser classificados de acordo com a natureza do relacionamento entre os dados e o modelo, e de acordo com o comportamento da função objetivo. Assim, eles podem ser: linear, fracamente não-linear, quasi-linear e não-linear. Na maioria dos problemas geofísicos o operador direto  $G$  é não-linear. Como nos algoritmos de aprendizagem de máquina, na inversão sísmica a não-linearidade implica em uma função de custo com forma complicada, possivelmente com mínimos locais. Por outro lado, se o operador  $G$  for aproximadamente linear, a função de erro se tornará convenientemente quadrática em relação a perturbações no espaço do modelo. A maior parte da teoria de inversão é baseada em problemas de inversão linear e, em muitas aplicações, ela é adequada para representar a natureza do sistema Sen (2006).

O modelo sísmico direto pode ser representado pelo modelo convolucional dado por:

$$d(t) = \int_{-\infty}^{\infty} s(\tau)r(t - \tau)d\tau + e_d(t) \quad (3.1)$$

onde  $d(t)$  é o traço sísmico,  $s(t)$  é a *wavelet*,  $e(t)$  é um ruído aleatório e  $r(t)$  é o refletividade. A representação discreta para o modelo convolucional do dato sísmico é dado pela operação matricial:

$$\mathbf{d} = \mathbf{S}\mathbf{r} + \mathbf{e} \quad (3.2)$$

onde  $\mathbf{S}$  é uma matriz convolucional construída utilizando uma *wavelet* e  $\mathbf{r}$  é a matriz de refletividades.

Como já mencionado, a relação entre o pulso sísmico e a propriedade de impedância acústica é não-linear. Para escapar da problemática da não-linearidade do operador direto, é necessário aproximar linearmente o pulso sísmico da impedância acústica. Para isto, duas medidas são necessárias: a primeira é admitir a refletividade como o logaritmo da impedância acústica (equação 3.3).



$$r(t) = \frac{1}{2} \Delta \ln(z(t)) \quad (3.3)$$

Esta aproximação é válida para valores de refletividade menores que 0.3. A segunda medida, é adotar um operador diferencial  $\mathbf{D}$ . Assim, se define o operador linear  $\mathbf{G} = (\mathbf{1}/2)\mathbf{SD}$  e o modelo  $m = \ln(z)$ . Com isto, a relação entre o dado sísmico e o parâmetro do modelo (impedância acústica) se torna linear por:

$$\mathbf{d} = \mathbf{G}\mathbf{m} + \mathbf{e} \quad (3.4)$$

Em teoria, o ruído é uma interferência aleatória que não se tem controle, na prática se considera ruído tudo que não é explicado pela função  $G$ , e.g. imprecisões no modelo físico e problemas com filtragem e processamento dos dados. Com isto, o problema da inversão acústica se torna não-linear e passa a ter representação matricial.

### 3.1.2 Máximo *a posteriori*

A teoria mais simples e genérica possível é obtida quando se usa uma abordagem probabilística (Tarantola, 2005). Na solução para a inversão sísmica, os parâmetros do modelo convolucional 3.4 podem ser representados em termos de suas distribuições de probabilidade. No modelo estocástico proposto por Passos de Figueiredo et al. (2014), as distribuições são consideradas normais e multivariadas e são denotadas por  $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ . Assim, assumindo que o ruído  $\mathbf{e}$  respeita uma distribuição gaussiana, as distribuições de probabilidade para o vetor dos dados sísmicos experimentais  $\mathbf{d}$ , para a *wavelet*  $\mathbf{w}$  e para o vetor dos parâmetros do modelo  $\mathbf{m}$  são definidos, respectivamente, pelas distribuições 3.5, 3.6 e 3.7.

$$p(\mathbf{d}|\boldsymbol{\mu}_d, \boldsymbol{\Sigma}_d) = N(\boldsymbol{\mu}_d, \boldsymbol{\Sigma}_d) \quad (3.5)$$

Onde  $\boldsymbol{\mu}_d = \mathbf{G}\mathbf{m}$  é o vetor com a sísmica sintética e  $\boldsymbol{\Sigma}_d$  é a matriz de covariância do ruído da sísmica, a qual é definida conforme a confiabilidade que o especialista tem no dado sísmico ou seu nível de ruído.

$$p(\mathbf{s}|\boldsymbol{\mu}_s, \boldsymbol{\Sigma}_s) = N(\boldsymbol{\mu}_s, \boldsymbol{\Sigma}_s), \quad (3.6)$$

Onde o valor esperado da *wavelet*  $\boldsymbol{\mu}_s$  é definido como um vetor nulo. Para que o método possa ser aplicado para a inversão acústica, é necessário estimar uma *wavelet* que possa ser aplicada no

modelo convolucional. Esta estimativa é realizada aplicando este mesmo processo de inversão na região de ocorrência de um poço, onde a refletividade pode ser calculada diretamente (Passos de Figueiredo et al., 2014). O algoritmo de Gibbs então é utilizado para amostrar na distribuição posterior da *wavelet* e a o valor médio e a incerteza são calculados.

$$p(\mathbf{m}|\boldsymbol{\mu}_m, \boldsymbol{\Sigma}_m) = N(\boldsymbol{\mu}_m, \boldsymbol{\Sigma}_m), \quad (3.7)$$

Neste ponto é possível inserir no método de inversão informações *a priori* que eventualmente estejam disponíveis. Por exemplo,  $\boldsymbol{\mu}_m$  pode ser uma matriz de baixas frequências gerada a partir da interpolação da impedância acústica observada em dois poços já perfurados (Passos de Figueiredo et al., 2014).

A inversão por Máximo *a posteriori* (MAP) (Buland e Omre, 2003; Figueiredo et al., 2014) é realizada para cada traço individualmente. As distribuições condicionais e o modelo convolucional apresentados anteriormente são as estruturas necessárias para realizar a inversão acústica. O ponto de partida é a aplicação do próprio método para estimar a *wavelet*, com ela é possível estimar as distribuições de probabilidades envolvidas no modelo. Em seguida, basta calcular a exponencial do modelo convolucional para obter a distribuição posterior para o parâmetro do modelo, que no caso em questão é a impedância acústica. Esta distribuição é dada por:

$$p(\mathbf{m}|\mathbf{d}_o, \mathbf{s}, \boldsymbol{\mu}_m, \sigma_d^2, \sigma_m^2) = N(\boldsymbol{\mu}_{m|}, \boldsymbol{\Sigma}_{m|}), \quad (3.8)$$

Neste arcabouço a média e variância posterior para cada traço podem ser calculadas analiticamente via (Figueiredo et al., 2014):

$$\boldsymbol{\mu}_{m|} = \boldsymbol{\mu}_m + \boldsymbol{\Sigma}_m \mathbf{G}^T (\mathbf{G} \boldsymbol{\Sigma}_m \mathbf{G}^T + \boldsymbol{\Sigma}_d)^{-1} (\mathbf{d}_o - \mathbf{G} \boldsymbol{\mu}_m), \quad (3.9)$$

$$\boldsymbol{\Sigma}_{m|} = \boldsymbol{\Sigma}_m - \boldsymbol{\Sigma}_m \mathbf{G}^T (\mathbf{G} \boldsymbol{\Sigma}_m \mathbf{G}^T + \boldsymbol{\Sigma}_d)^{-1} \mathbf{G} \boldsymbol{\Sigma}_m. \quad (3.10)$$

onde o cálculo da matriz inversa acima pode ser aproveitado para vários traços de uma região de interesse em certos casos, ou seja, quando as matrizes de covariância possam ser assumidas iguais para todos os traços da sísica da região. Desta forma alteram-se a sísica  $\mathbf{d}_0$  e a baixa frequência  $\boldsymbol{\mu}_m$  obtendo-se a média posterior para o traço desejado.

A solução para o método de inversão MAP é expressa em termos da covariância e do valor esperado. A matriz de covariância posterior indica a incerteza presente no resultado, não é necessário definir a tolerância de ajuste aos dados explicitamente, mas é preciso definir a matriz

de covariância *a priori* do resultado esperado, ou seja, é preciso ter conhecimento, mesmo que de forma grosseira, das correlações espaciais e variâncias que se espera do resultado. Por ser o valor esperado, as imagens de impedância acústica obtidas se caracterizam por serem suavizadas, principalmente na região de transição entre camadas. Durante a convolução, a *wavelet* funciona como uma modeladora, de modo que as altas frequências são filtradas e nível de detalhes das imagens se torna limitado.

## 3.2 Métodos de Super-resolução de Imagens

Super-resolução é o processo de obter uma ou mais imagens de alta resolução a partir de uma ou mais imagens de baixa-resolução através do aumento no número de pixel por unidade de área. Os algoritmos de super-resolução têm aplicação nas mais diferentes áreas, tais como, processamento de imagens aéreas e de satélite, reconhecimento de íris, holografia digital, melhoramento de imagens faciais e de texto, entre outras.

Os modelos de super-resolução podem ser classificados de acordo com diferentes fatores. Este fatores incluem o domínio de aplicação, o número de imagens de baixa resolução aplicadas e o método de reconstrução Nasrollahi e Moeslund (2014). Métodos baseados em interpolação são fáceis de implementar e amplamente utilizados, entretanto estes métodos sofrem de falta de expressividade, uma vez que modelos lineares não são capazes de expressar dependências complexas entre as entradas e as saídas (Hou e Andrews, 1978). Na prática tais métodos falham na tentativa de prever adequadamente detalhes de alta frequência levando a saídas de alta resolução borradas. Efeito semelhante ocorre durante a inversão sísmica, na qual as imagens resultantes apresentam resolução limitada e contornos borrados.

**A revisão da literatura foi realizada com as seguintes palavras-chaves: A revisão foi realizada sistematicamente nos seguintes periódicos:** O uso de métodos de *Deep Learning* para resolver problemas de super-resolução é uma abordagem recente. O modelo que representa o estado da arte em redes convolucionais para super-resolução é a apresentado na subseção a seguir.

### 3.2.1 Super-resolução por CNN

Não há evidência, na literatura, de trabalhos que abordem o problema de aumento de resolução de imagens de propriedades petrofísicas pós-inversão através de métodos de aprendizado supervisionado, como as redes neurais convolucionais. De acordo com Xiaoyu et al. (2012) para

melhorar a resolução da inversão sísmica é necessário adicionar alta frequência na aquisição e processamento do dado sísmico. Entretanto, expandir a reflexão de alta frequência é uma tarefa difícil por conta de fatores como atenuação da terra, ruído de alta frequência, entre outros. Além disso, como já mencionado na seção 3.1.2, o próprio processo de inversão acústica, modula a sísmica durante a convolução. Assim, a estratégia sugerida neste trabalho objetiva a inserção de faixas de alta frequência no pós-processamento da impedância invertida.

Os algoritmos de super-resolução realizam buscas por fragmentos de estruturas e os combinam para criar detalhes de alta-frequência (Freeman; Jones e Pasztor, 2002; Huang; Singh e Ahuja, 2015). Outra abordagem caminha no sentido de melhorar métodos de interpolação simples através da construção de dicionários de filtros pré-treinados e selecionar os fragmentos por algum algoritmo de *Hashing* (Romano; Isidoro e Milanfar, 2017). Os algoritmos citados possuem o foco na velocidade de inferência, as redes convolucionais, por outro lado, focam na construção das imagens de alta resolução, de modo a obter magnitudes cada vez maiores de detalhes.

A aplicação de redes convolucionais para realizar super-resolução é uma abordagem recente na literatura. Os primeiros trabalhos neste sentido datam do ano de 2014 e visaram a super-resolução de imagens únicas no domínio espacial Dong et al. (2014). As redes convolucionais extraem implicitamente múltiplas camadas de abstração através da otimização dos seus filtros. Elas são capazes de modelar a distribuição conjunta sobre uma imagem  $x$  como o produto de distribuições condicionais (van den Oord et al., 2016):

$$p(x) = \prod_{i=1}^{n^2} p(x_i | x_1, \dots, x_{i-1}) \quad (3.11)$$

onde,  $x_i$  é o pixel modelado. A imagem é percorrida linha por linha e cada pixel depende apenas dos *pixels* localizados em uma vizinhança pré-determinada. Para garantir esta propriedade se define uma máscara para os filtros da convolução, ou seja, é atribuído valor 0 para os pesos fora da zona de interesse. A amostragem ocorre sequencialmente, cada pixel é reinserido para a rede para que o próximo pixel seja previsto, deste modo cada *pixel* depende fortemente dos *pixels* anteriores sob uma perspectiva não-linear. O uso de convolução permite que a previsão para todos os *pixels* seja realizada de forma paralela durante o treinamento.

O modelo condicional sofreu melhoria para o reconhecimento de estruturas mais complexas. Entre as camadas convolucionais do modelo condicionado foram adicionadas unidades

multiplicativas com a seguinte ativação:

$$y = \tanh(W_{k,f} * x) \odot \sigma(W_{k,g} * x) \quad (3.12)$$

onde  $\sigma$  é a função sigmoide,  $k$  é o número da camada,  $*$  é o operado convolucional e  $\odot$  é a produto membro a membro. Este modelo é conhecido como *Gated PixelNN* (van den Oord et al., 2016).

O modelo de super-resolução condicional leva em consideração vetores de descrição  $\mathbf{h}$ , de modo que o modelo da distribuição condicional é dada por:

$$p(x|\mathbf{h}) = \prod_{i=1}^{n^2} p(x_i|x_1, \dots, x_{i-1}|\mathbf{h}) \quad (3.13)$$

Desta forma, as ativações das camadas convolucionais dependem de  $\mathbf{h}$ , antes de passarem pela função de não-linearidade. Se  $\mathbf{h}$  for informação referente a uma classe, todas as camadas terão um *bias* que determina a dependência desta classe. Entretanto, esta dependência não estará relacionada à localização do *pixel*. Por outro lado, se  $\mathbf{h}$  for mapeado para uma representação espacial  $\mathbf{s} = m(\mathbf{h})$ , onde  $m$  é uma rede deconvolucional, as camadas convolucionais terão *biases* dependentes da localização das estruturas contidas em  $\mathbf{h}$ , presentes na imagem. Assim, a equação 3.12 ganha a seguinte forma:

$$y = \tanh(W_{k,f} * x + V_{k,f} * \mathbf{s}) \odot \sigma(W_{k,g} * x + V_{k,g} * \mathbf{s}) \quad (3.14)$$

O modelo de rede condicional proposto por Dahl; Norouzi e Shlens (2017) representa o estado da arte em modelos convolucionais para super-resolução de múltiplas imagens. O modelo é composto de uma rede condicionante, do tipo tipo ResNet (He et al., 2016) e uma rede *prior*, do tipo PixelCNN (van den Oord et al., 2016). A rede condicionante realiza o mapeamento de uma imagem de baixa resolução para uma estrutura probabilística de alta resolução. Assim, ela permite compor a estrutura da imagem através da distribuição de probabilidade marginal dos *pixels*. A rede *prior* adiciona detalhes de alta resolução para tornar a saída mais realísticas.

Para treinar um modelo que mapeie uma imagem  $x$  de baixa resolução em uma imagem  $y$  de alta resolução, dada uma imagem  $y^*$  considerada a realidade desejada, é preciso otimizar os parâmetros  $\theta$  da distribuição condicional  $p_\theta(\mathbf{y}|\mathbf{x})$  de modo a maximizar a função objetivo

condicional dada por (Dahl; Norouzi e Shlens, 2017):

$$O(\theta|\mathcal{D}) = \sum_{(\mathbf{x}, \mathbf{y}^*) \in \mathcal{D}} \log p(\mathbf{y}^*|\mathbf{y}) \quad (3.15)$$

onde  $\mathcal{D} \equiv \{(\mathbf{x}^{(i)}, \mathbf{y}^{*(i)})\}_{i=1}^N$  denota o conjunto de treinamento da rede, composto pelos pares de imagens de baixa resolução e de alta resolução que representa a realidade observada.

Considere uma imagem  $\mathbf{x} \in \mathbb{R}^L$ ,  $A_i(\mathbf{x}) : \mathbb{R}^L \rightarrow \mathbb{R}^K$  representa a rede condicionante capaz de prever um vetor de valores que correspondem a  $K$  valores possíveis que o  $i$ -ésimo *pixel* de saída pode assumir. Analogamente,  $B_i(\mathbf{y}_{<i}) : \mathbb{R}^{i-1} \rightarrow \mathbb{R}^K$  representa a rede *prior* capaz de prever um vetor de valores do  $i$ -ésimo *pixel*. A previsão da distribuição sobre o  $i$ -ésimo pixel de saída é obtida pela adição dos dois conjuntos de saída e aplicação do operado de *softmax*:

$$p(y_i|\mathbf{x}, \mathbf{y}_{<i}) = \text{softmax}(A_i(\mathbf{x}) + B_i(\mathbf{y}_{<i})) \quad (3.16)$$

O algoritmo Gradiente Descente Estocástico é usado para otimizar os parâmetros  $A$  e  $B$ , a fim de maximizar a *log-likelihood* da equação 3.15. O aprendizado da rede ocorre pela otimização da função de custo entre as previsões do modelo (equação 3.16) e os valores discretos da imagem que representa a realidade  $y_i^* \in \{1...K\}$ :

$$O = \sum_{(\mathbf{x}, \mathbf{y}^*) \in \mathcal{D}} \sum_{i=1}^M \left( \zeta[\mathbf{y}_i^*]^T (A_i(\mathbf{x}) + B_i(\mathbf{y}_{<i}^*)) - lse(A_i(\mathbf{x}) + B_i(\mathbf{y}_{<i}^*)) \right) \quad (3.17)$$

Mais recentemente, os avanços das pesquisas do Google em *Deep Learning* disponibilizaram ferramentas de implementação de diferentes algoritmos de aprendizagem de máquina. Dentre estas ferramentas está o *Framework* de *Deep Learning* TensorFlow, no qual os modelos de redes convolucionais podem ser implementados e testados.

### 3.3 Resumo

Neste capítulo foram revisados os estados da arte em inversão sísmica acústica e modelos de rede convolucional para super-resolução. Pontos críticos dos métodos foram considerados e identificados para pesquisa futura. O próximo capítulo irá definir a proposta de pesquisa, apresentar o plano de trabalho e concluir com as perspectivas de contribuição.

# Bibliografia

- Bosch, M.; Mukerji, T. e Gonzalez, E. F. (2010), Seismic inversion for reservoir properties combining statistical rock physics and geostatistics: A review, *Geophysics* **75**(5), 75A165–75A176.
- Buduma, N. (2015), *Fundamentals of Deep Learning*, Academic Press, O'Reilly Media.
- Buland, A. e Omre, H. (2003), Bayesian linearized avo inversion, *Geophysics* **68**(1), 185–198.
- Dahl, R.; Norouzi, M. e Shlens, J. (2017), Pixel recursive super resolution, *CoRR* .
- Deutsch, C. (2002), *Geostatistical Reservoir Modeling*, Applied geostatistics series, Oxford University Press.
- Dong, C.; Loy, C. C.; He, K. e Tang, X. (2014), Image super-resolution using deep convolutional networks, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **38**(2), 295–307.
- Figueiredo, L. P.; Santos, M.; Roisenberg, M.; Neto, G. e Figueiredo, W. (2014), Bayesian framework to wavelet estimation and linearized acoustic inversion, *Geoscience and Remote Sensing Letters, IEEE* **PP**(99), 1–5.
- Freeman, W. T.; Jones, T. R. e Pasztor, E. C. (2002), Example-based super-resolution, *IEEE Computer Graphics and Applications* **22**(2), 56–65.
- Goodfellow, I.; Bengio, Y. e Courville, A. (2016), *Deep Learning*, MIT Press. <http://www.deeplearningbook.org>.
- GS IPL (2017), Seismic surveys, <http://geostar-surveys.com/methodology%20-%20High%20Resolution%20Seismic%20surveys.html>.
- Géron, A. (2017), *Hands-on Machine Learning with Scikit0Learn and TensorFlow*, Academic Press, O'Reilly Media.

- He, K.; Zhang, X.; Ren, S. e Sun, J. (2016), Deep residual learning for image recognition, 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770–778.
- Hou, H. e Andrews, H. (1978), Cubic splines for image interpolation and digital filtering, IEEE Transactions on Acoustics, Speech, and Signal Processing **26**(6), 508–517.
- Huang, J. B.; Singh, A. e Ahuja, N. (2015), Single image super-resolution from transformed self-exemplars, 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 5197–5206.
- Nasrollahi, K. e Moeslund, T. B. (2014), Super-resolution: a comprehensive survey, Machine Vision and Applications **25**(6), 1423–1468.
- Passos de Figueiredo, L.; Santos, M.; Roisenberg, M.; Schwedersky Neto, G. e Figueiredo, W. (2014), Bayesian framework to wavelet estimation and linearized acoustic inversion, Geoscience and Remote Sensing Letters, IEEE **11**(12), 2130–2134.
- Romano, Y.; Isidoro, J. e Milanfar, P. (2017), Rair: Rapid and accurate image super resolution, IEEE Transactions on Computational Imaging **3**(1), 110–125.
- Sen, M. K. (2006), Seismic Inversion, Society of Petroleum Engineers, Richardson, TX, USA.
- Srivastava, R. P. e Sen, M. K. (2009), Fractal-based stochastic inversion of poststack seismic data using very fast simulated annealing, Journal of Geophysics and Engineering **6**(4), 412.
- Tarantola, A. (2005), Inverse Problem Theory and Methods for Model Parameter Estimation, Society for Industrial and Applied Mathematics.
- van den Oord, A.; Kalchbrenner, N.; Vinyals, O.; Espeholt, L.; Graves, A. e Kavukcuoglu, K. (2016), Conditional image generation with pixelcnn decoders, CoRR **abs/1606.05328**.
- Xiaoyu, X.; Yun, L.; Desheng, S.; Xiangyu, G. e Huifeng, W. (2012), Studying the effect of expanding low or high frequency on post-stack seismic inversion, pp. 1–5.