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 revisão bibliográfica, problemática, proposta e resultados 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

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:

$$T = \sqrt{\frac{1}{5}H} + \frac{1}{340}H \tag{1.1}$$

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 admite muitas soluções, 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 de extrema importância 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 disciplina 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, e 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 relação aos métodos de 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).

No campo de processamento de imagens o método de *Deep Learning* de maior destaque nos dias atuais é conhecido como Redes Neurais Convolucionais (RNC). 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* (Dong et al., 2014). 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 pode ser treinada para realizar as etapas listadas anteriormente, de forma iterativa e para um conjunto de diferentes imagens. A proposta aqui discutida, sugere incorporar a abordagem das redes convolucionais ao final do processo de inversão. As imagens obtidas da inversão sísmica podem ser pós-processadas pelo modelo convolucional a fim de gerar 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 a hipótese de pesquisa deste trabalho. Nos capítulos posteriores serão apresentados os aspectos físicos, de implementação desta proposta e os resultados preliminares.

1.1 Hipótese

A super-resolução das imagens de inversão sísmica pode ser realizada com o aumento da alta frequência na aquisição e processamento do dado sísmico utilizado na inversão, entretanto, esta é uma tarefa difícil devido a fatores como atenuação da terra e ruído (Xiaoyu et al., 2012). Baseado neste fato, a hipótese de pesquisa deste trabalho é de que é possível obter ganho de resolução na inversão sísmica através da aplicação de um modelo de rede neural convolucional como método de pós-processamento.

1.2 Objetivos

Este trabalho investiga a problemática da super-resolução dos artefatos da inversão sísmica por meio de um modelo de redes neurais convolucionais. A abordagem se dá através da incorporação deste modelo de rede neural na última etapa do processo de inversão sísmica, 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.

1.2.1 Objetivos Específicos

Os objetivos específicos deste trabalho são:

- Conceber um modelo de rede neural convolucional para super-resolução de imagens de inversão.
- Entender as abstrações em cada camada da rede neural.
- Estudar o espectro de frequência das imagens pós-processadas com o modelo de rede convolucional.
- Definir uma estratégia de aplicação do modelo para diferentes tipos de dados de inversão (parametrização).
- Estudo de incerteza do processo de super-resolução.

Um objetivo secundário deste trabalho é utilizar o conhecimento acumulado para estudar estratégias que relacionem as redes convolucionais e a simulação geoestatística multiponto. Esta etapa de trabalho será desenvolvida mediante aprovação do processo de doutorado sanduíche, em tramitação no CNPq com número 202482/2017-0. Em caso de aprovação, esta atividade será realizada em cooperação com o Departamento de Ciências Geológicas, Universidade Stanford, sob orientação do Prof. Dr. Jef Karel Caers.

1.3 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 3apresenta o estado da arte relacionado à inversão acústica e redes convolucionais aplicadas na geração de imagens em super-resolução. O Capítulo 4 trata dos resultados preliminares, complementação da proposta e o plano de trabalho.

Capítulo 2

Fundamentação Teórica

Neste capítulo 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) cujos valores caracterizam completamente o sistema. A escolha dos parâmetros do modelo geralmente é não única, de modo que dois conjuntos de parâmetros diferentes podem ser equivalentes.

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.

$$d = F(m) \tag{2.1}$$

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:

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

onde, F é o sistema físico investigado, e relaciona os parâmetros do modelo $m = (m_1, m_2, ..., m_n) \subset R^n$ 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 a partir 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 chamadas de *facies* e possuem 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.

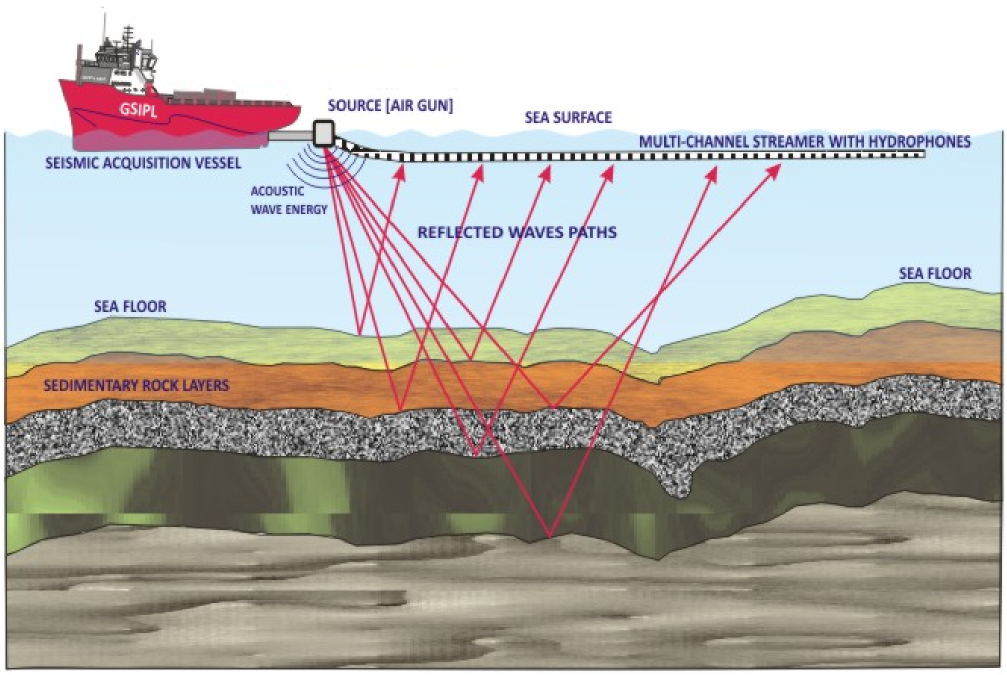


Figura 2.1: Método de sísmica de reflexão (GSIPL, 2017)

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):

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

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°, 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 $*$ da *wavelet* s com os coeficientes de refletividade r:

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

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 3-D. 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.

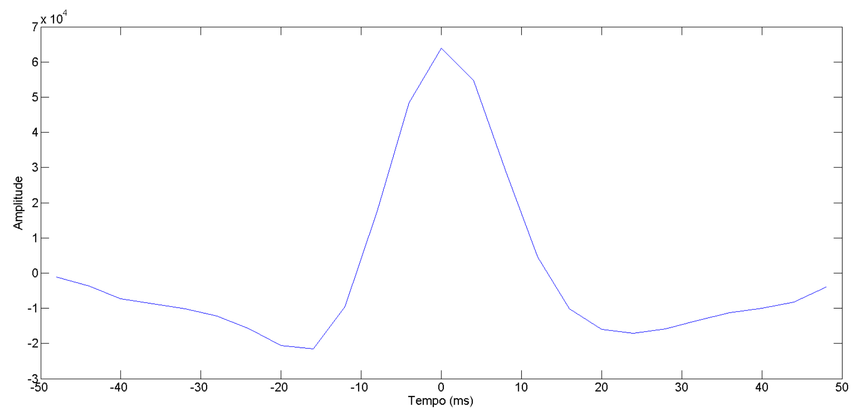


Figura 2.2: *Wavelet* extraída de dados reais

2.3 Redes Neurais Convolucionais

As Redes Neurais Convolucionais (RNC), 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 chamado 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:

$$s(t) = \int x(a)w(t-a)da. \tag{2.5}$$

A convolução costuma ser denotada com um asterisco e aplicada com o tempo t discretizado para valores inteiros, dada da seguinte forma:

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

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:

$$S(i,j) = (X * W)(i,j) = \sum_m \sum_n X(m,n)W(i-m, j-n). \tag{2.7}$$

Nas redes convolucionais há pelo menos duas estruturas básicas, a camada convolucional e a camada de *pooling*. A arquitetura típica de uma RNC compreende duas camadas convolucionais, cada uma 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 nerônios de uma camada estão conectados a todos os neurônios da camada seguinte.

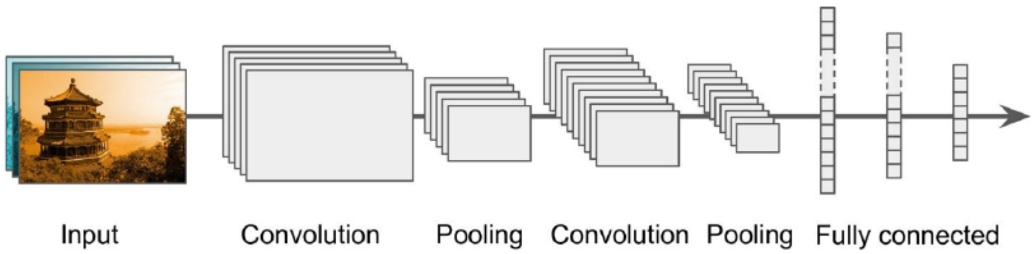


Figura 2.3: Arquitetura típica de uma rede neural convolucional. (Géron, 2017)

A camada convolucional é o elemento mais importante de uma RNC. 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.

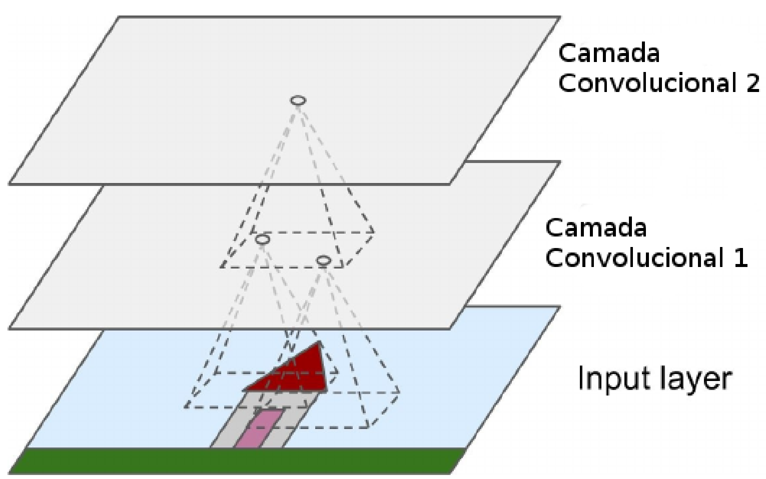


Figura 2.4: Camadas de uma RNC com campos receptivos retangulares.

A Figura 2.5 ilustra um exemplo de campo receptivo e das conexões entre duas camadas em 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 $i \times s_h$ até $i \times s_h + f_h - 1$, colunas $j \times s_w$ até $j \times s_w + f_w - 1$, 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.

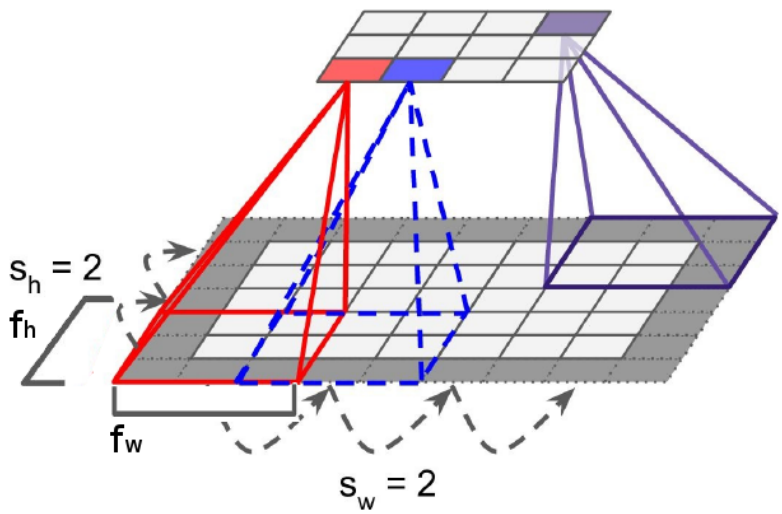


Figura 2.5: Conexão entre camadas com campo receptivo 3 x 3 e *strides* de tamanho 2.

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. Durante o treinamento de uma rede convolucional cada elemento dos filtros é otimzado. A Figura 2.6 ilustra dois filtros exemplos e o resolutado da convolução de cada um com uma certa imagem. 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. Na imagem da direita a convolução com o filtro horizontal destacou as linhas brancas horizontais, ao passo que o restante ficou borrado. Assim, quando uma característica é detectada por um neurônio, ela representa o tipo de padrão da entrada que causará a sua ativação. Estes padrões podem ser bordas, contornos ou estruturas com outras formas.

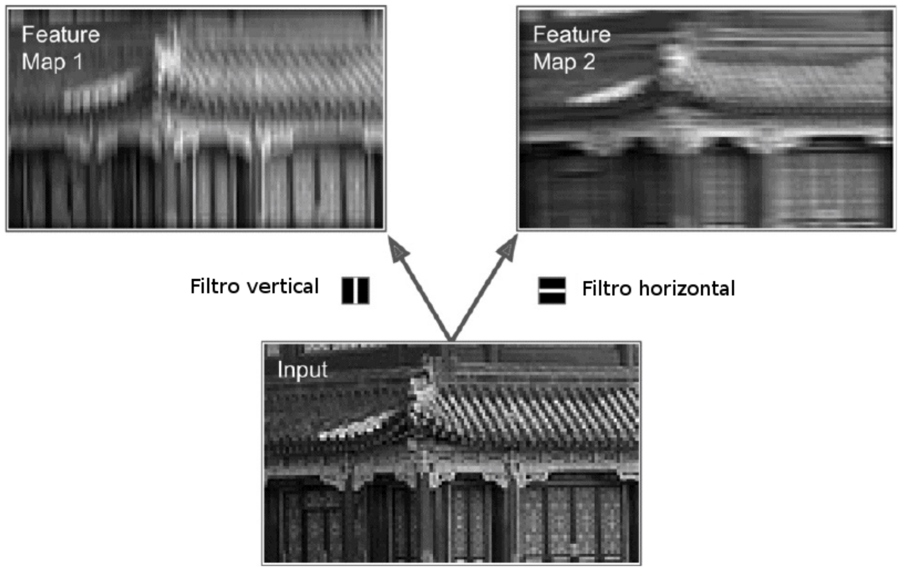


Figura 2.6: Aplicação de dois filtros diferentes para obter mapas de características.

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, de modo que, é possível imaginar que à medida que o número de imagens aumenta, a estrutura ilustrada se replica horizontalmente.

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 estágio de detecção, na qual cada ativação é submetida a uma função não-linear. O terceiro estágio é chamado de *pooling*, responsável por modificar a saída da camada convolucional para obter um sumário estatístico das saídas da convolução. Semelhante ao que ocorre na convolução, a região sobre a qual se aplica o *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

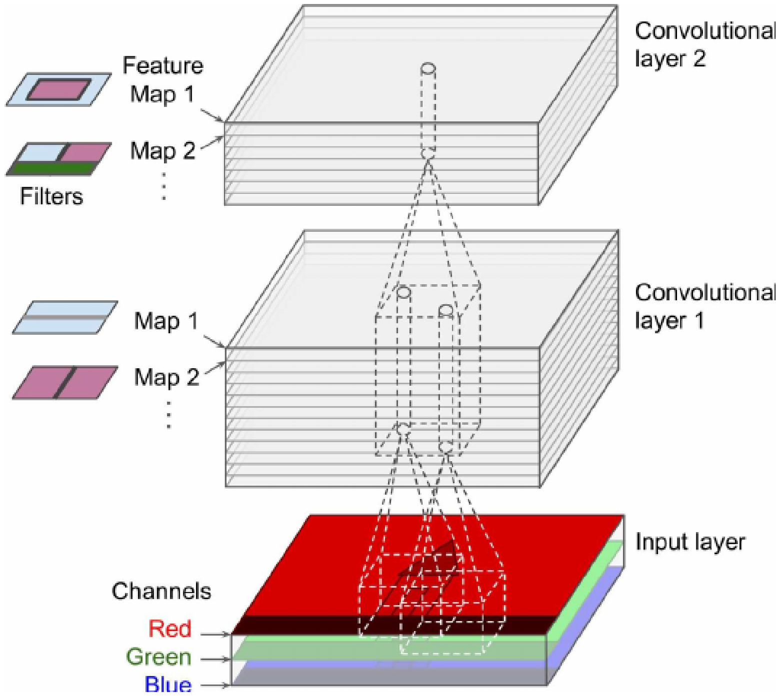


Figura 2.7: Camadas convolucionais com múltiplos mapas de características e imagens com três canais.

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 $L2$ de uma vizinhança, ou a média ponderada baseada na distância do *pixel* central.

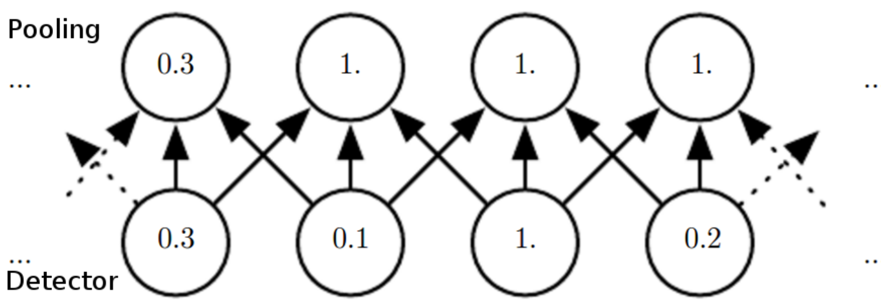


Figura 2.8: Operação de *pooling* com campo receptivo de tamanho 3. Nesta operação é selecionado o máximo valor de ativação da etapa de detecção.

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 [1]. Por outro lado, há contextos em que o local da característica é uma informação relevante e deve ser preservada. Por exemplo, 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. Na classificação de imagens, as entradas para a camada classificadora devem ter o mesmo tamanho. Assim, o *stride* entre regiões de *pooling* pode variar para que a camada classificadora 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 de *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 (Géron, 2017).

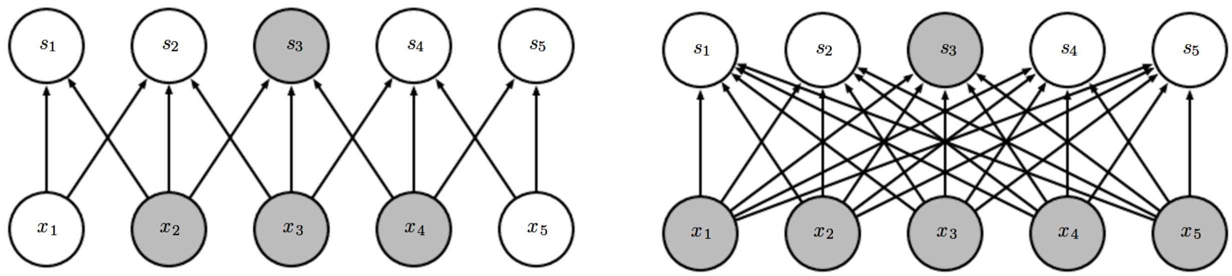
2.3.4 Propriedades das Redes Convolucionais

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 apresentados em maiores 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 *pixels*. Por exemplo, é possível identificar características de uma face humana no reconhecimento de pessoas, ou estruturas com significado geológico em um estudo geofísico. As Figuras 2.9a e 2.9b ilustram os modelos de conectividade esparsa e não-esparsa, respectivamente. É possível notar que na conectividade não-esparsa (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

[1] O rosto da figura pública Nestor Cerveró, por exemplo, seria facilmente identificável por uma rede convolucional com uma camada de *pooling*

destaque. O número de elementos que afetam o elemento em destaque na conectividade esparsa é definido pelo tamanho do filtro utilizado na convolução.



(a) Conectividade esparsa. (b) Conectividade tradicional.

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). Parece senso comum que estas informações

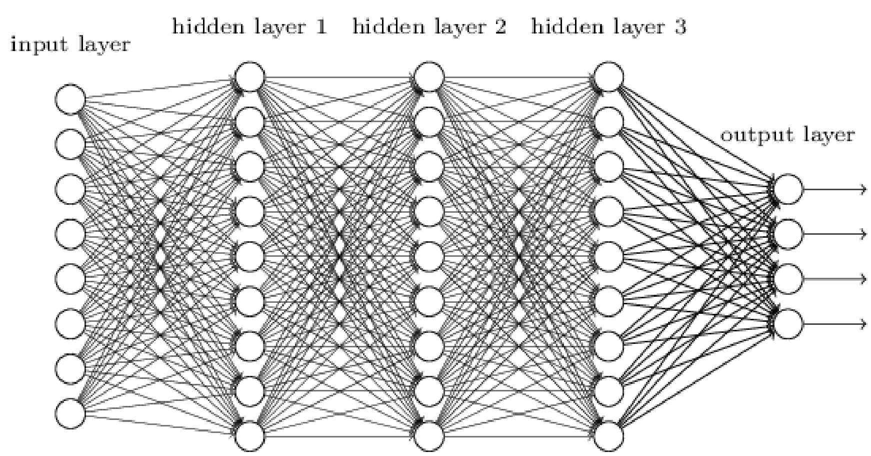


Figura 2.10: Organização de camadas de uma rede neural do tipo *feedforward*.

estruturais são muito relevantes em problemas geoestatísticos, afinal as imagens neste domínio representam estruturas geológicas.

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:

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

onde, σ 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 logo 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 RNC, 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. O Capítulo seguinte apresenta a revisão do estado da arte para inversão sísmica e para os modelos de RNC utilizados para Super-resolução.

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.

Deutsch, C. (2002), Geostatistical Reservoir Modeling, Applied geostatistics series, Oxford University Press.

Dong, C.; Loy, C. C.; He, K. e Tang, X. (2014), Learning a Deep Convolutional Network for Image Super-Resolution, Springer International Publishing, Cham.

Goodfellow, I.; Bengio, Y. e Courville, A. (2016), Deep Learning, MIT Press. `http://www.deeplearningbook.org`.

GSIPL (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.

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.

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.

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.