

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

1	Introdução	1
2	Fundamentação Teórica	2
2.1	Problema Inverso	2
2.2	Inversão Sísmica	3
2.2.1	Aquisição Sísmica	3
2.3	Redes Neurais Convolucionais	5
2.3.1	Convolução	6
2.3.2	Filtros	9
2.3.3	Pooling	11
2.3.4	Propriedades das Redes Convolucionais	12
2.4	Resumo	14
	Bibliografia	15

Capítulo 1

Introdução

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 $F(\cdot)$ 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.

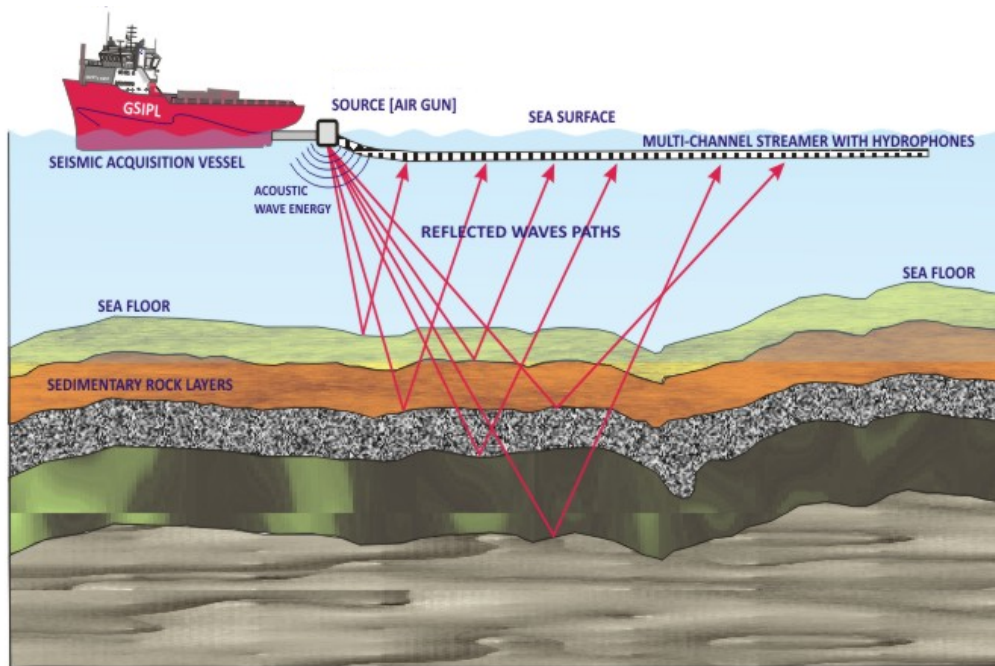


Figura 2.1: Método de sísmica de reflexão (GSIPL, 2017)

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):

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

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° , 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.

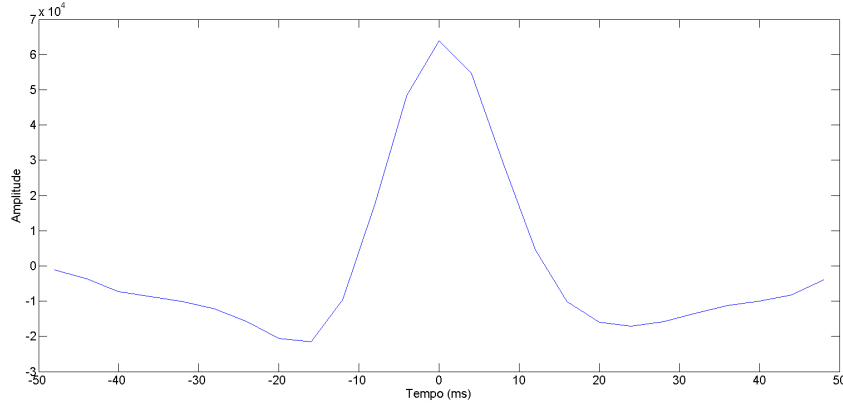


Figura 2.2: *Wavelet* extraída de dados reais

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 das quais 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.

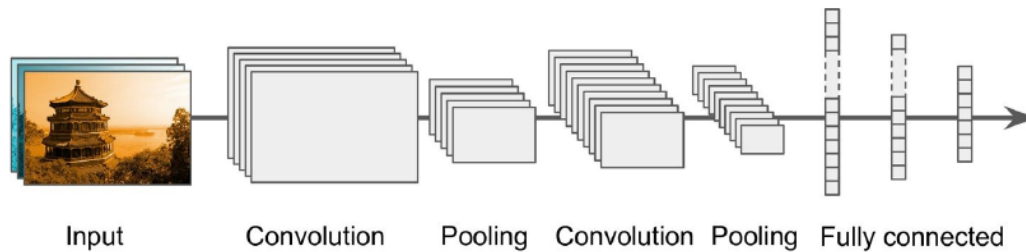


Figura 2.3: Arquitetura típica de uma rede neural convolucional. Fonte:Géron (2017)

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.

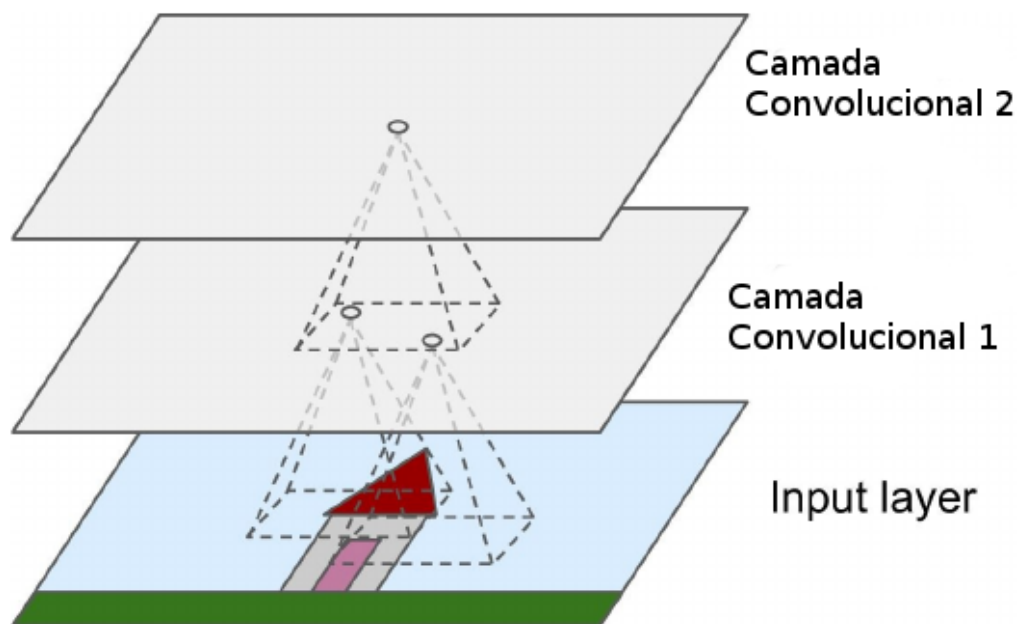


Figura 2.4: Camadas de uma CNN com campos receptivos retangulares.

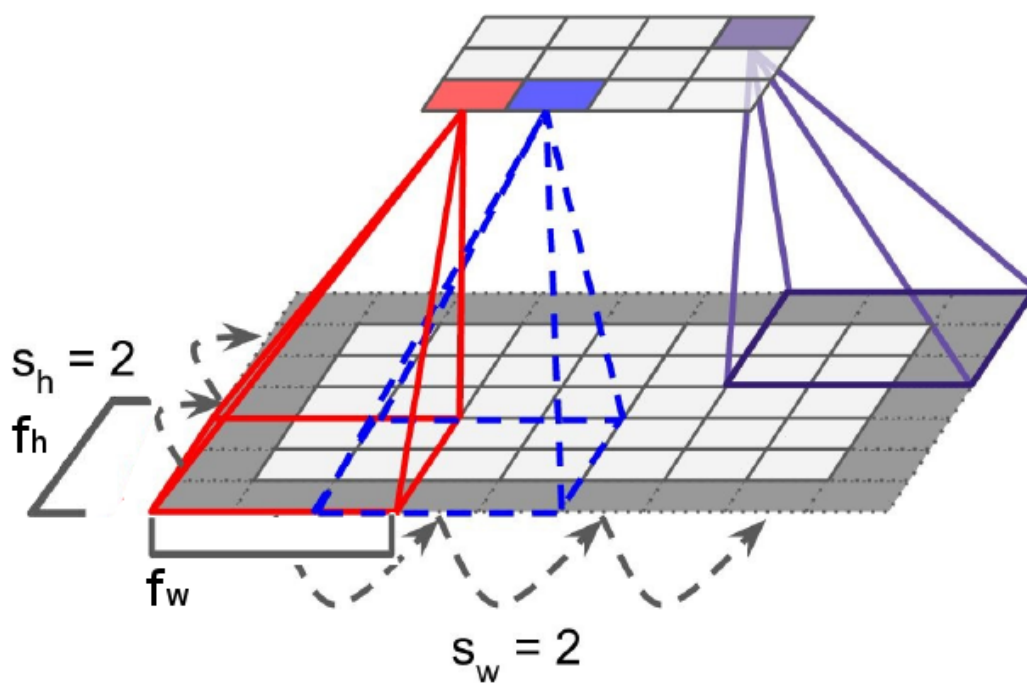


Figura 2.5: Conexão entre camadas com campo receptivo 3 x 3 e *strides* de tamanho 2.

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.

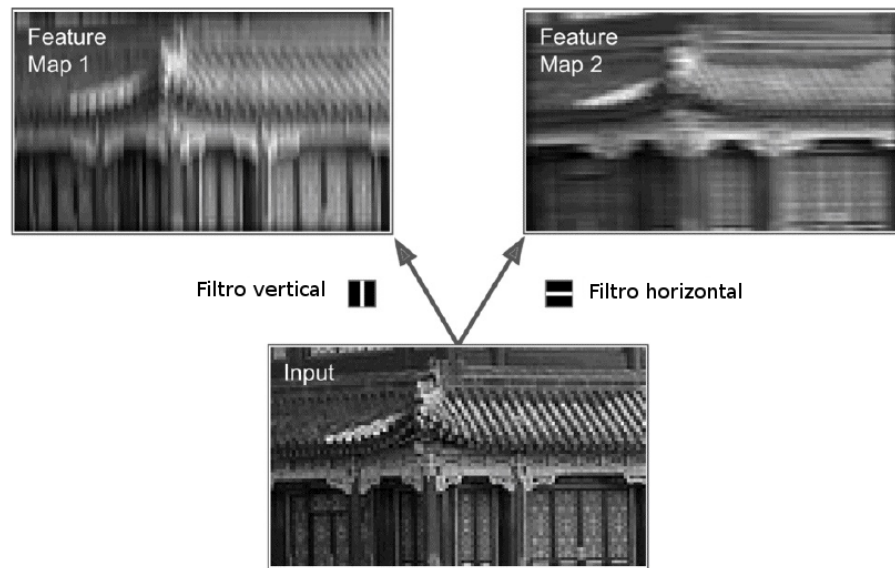


Figura 2.6: Aplicação de dois filtros diferentes para obter mapas de características.

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.

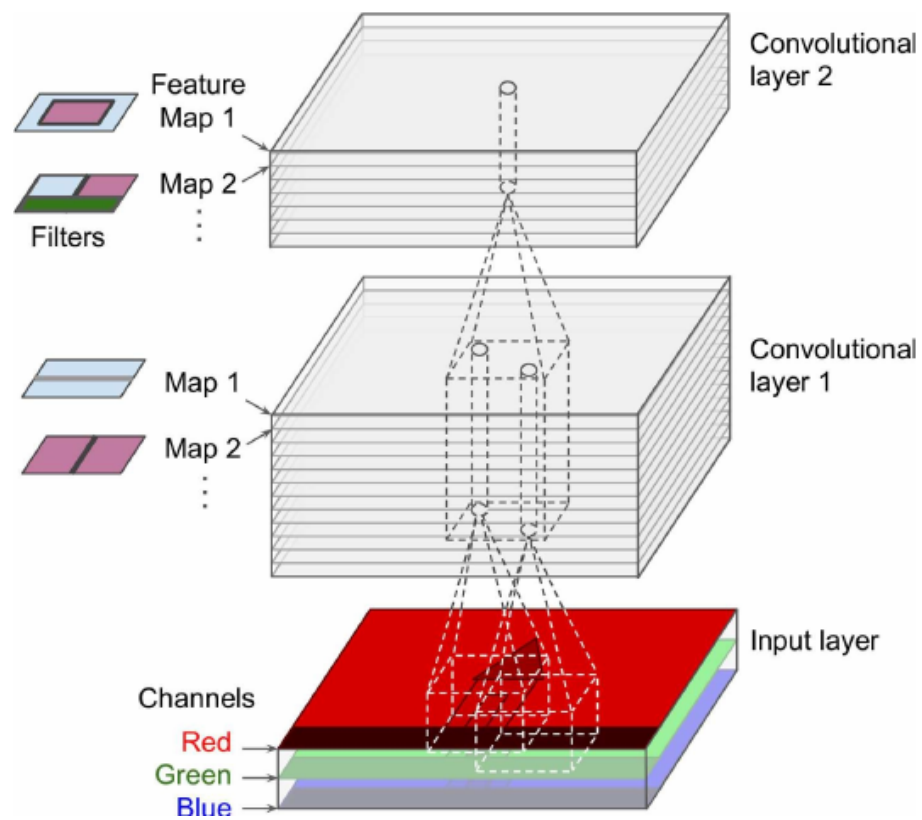


Figura 2.7: Camadas convolucionais com múltiplos mapas de características e imagens com três canais.

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.

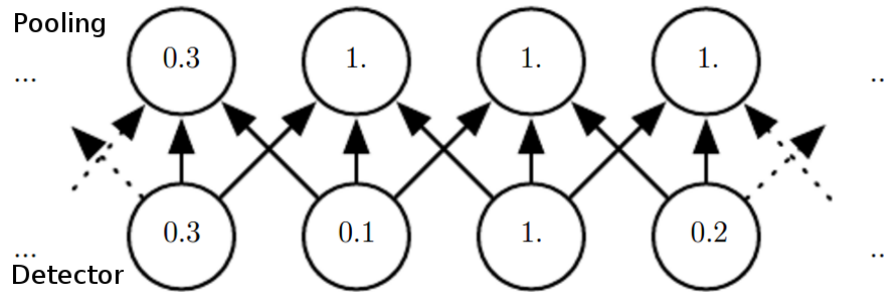


Figura 2.8: Operação de *pooling* com campo receptivo de tamanho 3. Nesta operação é selecionado o máximo valor de ativação da etapa de detecção.

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 da 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. Por exemplo, na classificação de imagens as entradas para a camada de classificação

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

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.

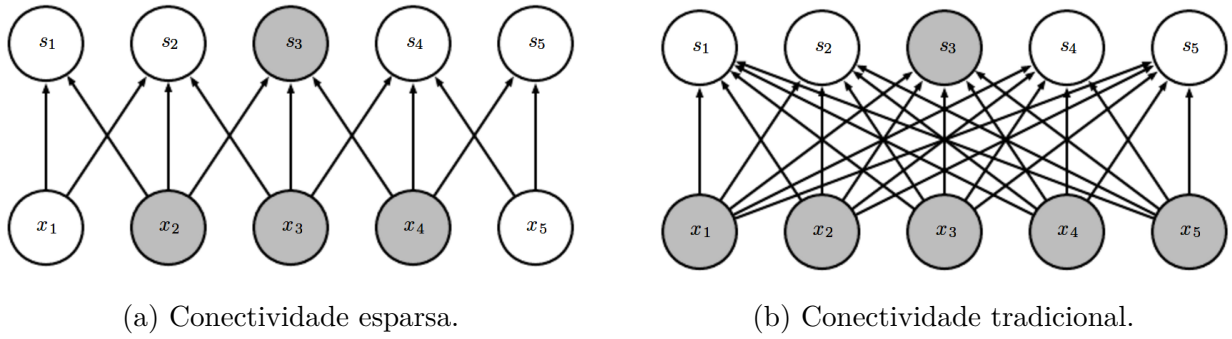
As função de *pooling* sumariza as respostas de vizinhanças separadas por k *pixels*, por isso, o tamanho do seu campo receptivo é menor que o campo receptivo da convolução. Isto aumenta a eficiência computacional da rede, pois a camada seguinte à *pooling* terá k vezes menos entradas para processar. 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 **O 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 na identificação 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

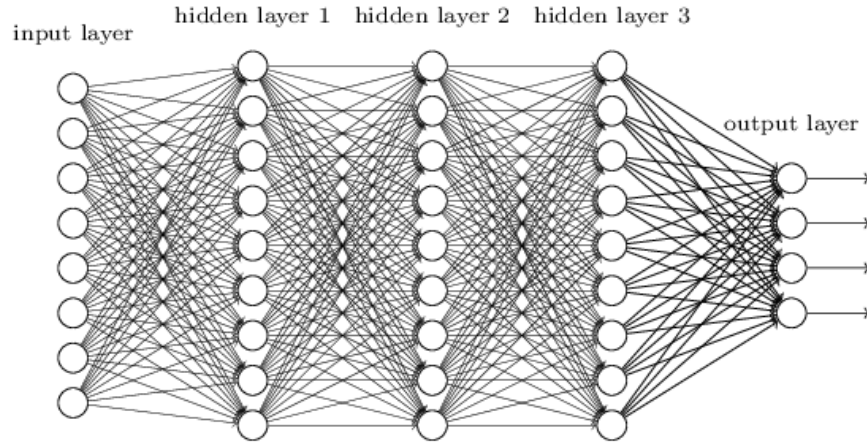


Figura 2.10: Organização de camadas de uma rede neural do tipo *feedforward*.

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, 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 convolvido 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

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.
- 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.
- 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.