



guidebook

16° WORKSHOP DE

WORCAP




COMPUTAÇÃO

2016
PLICADA

25 e 26 de outubro de 2016

Auditório Fernando de Mendonça - LIT
Instituto Nacional de Pesquisas Espaciais

www.lac.inpe.br
/worcap2016

 Fototerra



/worcap2016

A detecção de bordas é uma das operações mais importantes para a análise visual e automática de imagens. É a abordagem mais usada para detecção de descontinuidades nos valores ou atributos de uma imagem.

Diferentes técnicas de detecção de bordas geram diferentes resultados e não é possível afirmar que uma única borda seja consistentemente melhor em todos os pontos para todos os problemas.

A aprendizagem profunda (Deep Learning) envolve desenvolver sistemas para reconhecer padrões, muitas vezes complexos e abstratos, alimentando grandes quantidades de dados através de redes sucessivas de neurônios artificiais e refinar a forma como essas redes respondem à entrada. Essa abordagem tem se mostrado muito eficaz para o reconhecimento de fala ou outros tipos de áudio, ou ainda para classificar informações visuais.

A nossa proposta objetiva, criar um arcabouço baseado em Aprendizagem Profunda que permita extrair informações das imagens de radar polarimétricas produzindo uma borda mais completa oriunda dessas informações. O desafio consiste em ser capaz de fazer essa extração de informações ponderando de forma quantitativa e qualitativa a contribuição de cada fonte produzindo um sistema amigável e adaptado à prática dos usuários de imagens de sensoriamento remoto.

SENTINEL - UM AMBIENTE VIRTUAL PARA GERENCIAMENTO E MONITORAMENTO DE APLICAÇÕES CIENTÍFICAS

Bruno Leonor, Walter dos Santos, Reinaldo Rosa

A computação vem revolucionando o modo como diversas áreas científicas tem conduzido suas pesquisas. Para realizar simulações/pesquisas os sistemas utilizados lidam com uma enorme quantidade de variáveis que podem exigir uma grande quantidade de cálculos, o que tem contribuído com o aumento na demanda por recursos computacionais cada vez mais poderosos. O crescente número de projetos científicos e um explosivo aumento do volume de dados tem motivado a adoção de abordagens em eScience em diversas partes do mundo como apoio na resolução de problemas práticos e de pesquisa. A proposta deste trabalho é apresentar uma solução de eScience para apoiar cientistas no acompanhamento de suas pesquisas em um ambiente de nuvem, possibilitando acesso as suas aplicações computacionais onde poderão parametrizar, executar, visualizar e compartilhar os resultados obtidos transparentemente através de qualquer dispositivo que ofereça acesso a web. A infraestrutura computacional utilizada para demonstração deste projeto é constituída basicamente de microcomputadores Raspberry PI devido a seu baixo custo e recursos oferecidos, podendo ser associados outros recursos de hardware sempre que necessário. Cada uma das aplicações é executada em um ambiente isolado através de um container (Docker). Para armazenamento dos dados (entrada e saída) das aplicações optamos pelo uso do banco de dados MongoDB, um banco de dados orientado a documentos. Como estudo de caso abordaremos o programa EMBRACE (Estudo e Monitoramento Brasileiro do Clima Espacial), pois o conhecimento e previsão de diferentes fenômenos que afetam diretamente as atividades humanas, como a atividade solar por exemplo, têm grande importância para o bem estar da sociedade. Como um protótipo para a implementação, a técnica DFA (detrended fluctuation analysis) foi aplicada na análise das medições de radiação solar de 1978 a 2012. Discutimos também como novas aplicações podem ser adicionadas ao ambiente considerando o novo paradigma do Data Science/Big Data.

MOBILIDADE URBANA EM UM DIA TÍPICO - SÉRIE TEMPORAL DE PROPRIEDADES ESTATÍSTICAS

Jessica Santos, Leonardo Bacelar, Marcos Quiles, Elbert Macau

Um sistema complexo pode ser modelado através do conceito de rede complexa que, por meio de grafos, retrata um conjunto de conexões com características topológicas não completamente regulares nem mesmo completamente aleatórias. Na sociologia, na biologia, na mobilidade urbana e em muitas outras áreas, essas características apresentam semelhanças permitindo a reutilização das soluções a partir da generalização dos modelos. A modelagem utilizada neste estudo visa utilizar ferramentas tradicionalmente referentes a áreas como Física Estatística e Sistemas Complexos para analisar dados de mobilidade urbana em cenário de crescimento urbano não planejado que pode gerar impactos negativos, como congestionamentos, poluição do ar e falhas de infraestruturas. A cidade de São José dos Campos/SP detém um estudo para melhorar sua infraestrutura a partir de dados reais de mobilidade obtidos em um dia típico, pesquisa efetuada pelo IPPLAN, Instituto de Pesquisa, Administração e Planejamento. Neste estudo a cidade é dividida em 55 zonas de tráfego. Um grafo de origem e destino foi gerado com representação matricial em 3 dimensões, no qual cada vértice representa um centroide da zona de tráfego e as arestas são ponderadas pelo “sampling”, fluxo de pessoas, com variação no tempo. Para a análise estatística os dados foram agrupados numa janela de definida pelo pesquisador. Neste trabalho foi considerado um agrupamento a cada 1h. Os resultados - séries temporais de propriedades estatísticas - mostram que o maior pico de fluxo ocorre entre 12h e 13h, os valores mínimos no período das 2h e 6h, e os horários que representam a maior média de fluxo entre zonas de tráfegos ocorrem, respectivamente, entre 7h e 8h, 13h e 14h e 17h e 18h. As variáveis analisadas são ainda espacializadas, com base em Séries Temporais em Bancos de Dados Geográficos.

PROJETO E DESENVOLVIMENTOS DE UM BANCO DE DADOS DE AMOSTRAS PARA USO EM PROJETOS DE MAPEAMENTO DE USO E COBERTURA DA TERRA

Luan Melo, Lubia Vinhas

Para o monitoramento de uso e cobertura da Terra ou de mudanças de uso e cobertura são produzidos mapas que servem para diferentes aplicações. Os mapas são produzidos utilizando diferentes tipos de estratégias que variam de acordo com a natureza da classificação, a complexidade e a escala da área de estudo. Mapas de uso e cobertura da Terra são mapas temáticos que possuem uma legenda associada descrevendo as possíveis categorias que um pixel pode assumir. Não há um sistema padrão de classificação internacionalmente aceito e adotado pelas diferentes organizações e pesquisadores, mas toda vez que um mapa temático é criado, uma legenda foi adotada. Existem diversas metodologias para geração de mapas de uso e cobertura da Terra, e tradicionalmente os métodos de classificação são divididos em dois grupos: supervisionados e não-supervisionados. Métodos supervisionados baseiam-se na similaridade entre as amostras que representam cada classe e os pixels da imagem para rotulá-los de acordo com as classes conhecidas. Portanto, a utilização de uma amostra mais representativa de cada classe de interesse é fundamental para uma boa acurácia da classificação. Dessa forma, o objetivo deste artigo é estudar e projetar um banco de dados de amostras de uso e cobertura da Terra. É verificado se o uso de bancos de dados orientados a grafo para o armazenamento de amostras pode se tornar mais eficiente em aplicações de classificação.

TÉCNICAS COMPUTACIONAIS DE ANÁLISE MORFOLÓGICA DE GALÁXIAS

Rubens Sautter, Reinaldo Rosa, Reinaldo de Carvalho

A morfologia de galáxias é uma ferramenta eficiente para analisar a dinâmica de galáxias, permitindo caracterizar o efeito da interação entre galáxias e o seu meio. O principal desafio computacional deste tipo de análise é a grande quantidade de dados que devem ser analisados para uma descrição precisa da dinâmica de galáxias. Neste trabalho é apresentado o CyMorph, uma nova ferramenta desenvolvida em Python para a análise morfológica de imagens de galáxias. Esta nova ferramenta é proposto o refinamento dos parâmetros morfométricos já estabelecidos na literatura, como concentração, assimetria e suavização. Também é proposto a incorporação de novos parâmetros morfométricos, como a espiralidade e o novo coeficiente de assimetria gradiente. Para cada parâmetro é apresentado um conjunto de testes para determinar o melhor critério a ser utilizado. Para avaliar o desempenho desta ferramenta, são apresentados resultados preliminares da medida dos parâmetros em imagens simuladas de galáxias em três resoluções, apresentando casos com a presença e a falta de ruído na imagem.

MÉTODO IMEX PARA SOLUÇÃO DO SISTEMA BIDIMENSIONAL DE EQUAÇÕES DE BURGERS

Antonio Zarzur, Stephan Stephany, Saulo Freitas

Simulações computacionais baseiam-se em modelos matemáticos desenvolvidos para certas classes de fenômenos. A solução computacional das equações diferenciais parciais presentes nesses modelos requer a escolha de um método de integração temporal capaz de simular, de forma estável, a evolução do problema. Entre os diversos métodos existentes, tradicionalmente categorizados em esquemas explícitos ou implícitos, há alguns mais adequados para determinadas classes de fenômenos, não existindo um método geral que sirva adequadamente para todos os fenômenos. Esquemas explícitos são facilmente implementáveis, por utilizar apenas informações conhecidas para calcular o próximo passo de tempo, porém sua estabilidade é condicional devido ao teorema da equivalência de Lax. Consequentemente, os passos de tempo necessários para a evolução do problema podem ser restritos a valores demasiadamente pequenos. Esquemas implícitos podem levar a uma abordagem estável sem restrições nas discretizações temporais e espaciais, mas por outro lado podem resultar em sistemas de equações não lineares cuja solução é computacionalmente cara. Uma abordagem mais recente é a combinação desses dois métodos, resultando nos chamados esquemas IMEX (Implícito-Explícito). Tais métodos foram desenvolvidos para solucionar equações com escalas de tempo variáveis, de forma que os termos rápidos são resolvidos implicitamente e os mais lentos são resolvidos explicitamente. O esquema resultante visa otimizar o tempo de processamento evitando passos de tempo desnecessariamente pequenos para os termos rápidos. Neste trabalho, diferenças finitas são utilizadas para solucionar o sistema bidimensional de equações de Burgers, sendo a integração temporal realizada por cada um dos três esquemas previamente mencionados (explícito, implícito, IMEX). As abordagens são estudadas de forma comparativa com relação às suas propriedades e ao seu desempenho computacional. Os resultados numéricos são analisados contra a solução analítica da equação de Burgers, que é conhecida e amplamente utilizada na mecânica de fluidos computacional.

PROPOSTA PARA O PLANEJAMENTO DA EXPANSÃO DA CANA-DE-AÇÚCAR ATRAVÉS DE INFERÊNCIA BAYESIANA

Alexsandro Silva, Leila Garcia, Thales Korting

Métodos de inferência são utilizados para produzir novas informações espaciais a partir da integração de dados. A inferência Bayesiana utiliza o teorema de Bayes para atualizar o conhecimento prévio de um evento (probabilidade a priori) considerando uma nova evidência (probabilidade condicional), permitindo que se tenha um raciocínio plausível baseado no grau de confiança (probabilidade a posteriori). Redes Bayesianas são modelos que empregam a inferência Bayesiana e são definidas em termo de um grafo acíclico direcionado que representa as variáveis do modelo e suas dependências condicionais.