

MACHINE LEARNING NO MONITORAMENTO DE EXSUDAÇÕES E DERRAMES DE ÓLEO NA SUPERFÍCIE DO MAR: UMA PROPOSTA PARA O GOLFO DO MÉXICO

Miranda, F.P.¹; Carvalho, G.A.²; Matias I.O.³

¹Petrobras/CENPES; ²UFRJ/COPPE; ³PUC-/Rio

RESUMO: De acordo com o *Hype Cycle* elaborado pela empresa Gartner, as tecnologias ligadas a *machine learning* encontram-se no ponto máximo de expectativas da indústria e academia. Entretanto, a utilização de tais ferramentas de modo produtivo na exploração e produção de petróleo depende ainda de altos investimentos em pesquisa aplicada. Este é o grande desafio a ser enfrentado nos próximos 5 anos, segundo a citada empresa, e a razão de ser do estudo aqui proposto. Por sua vez, a indústria do petróleo em águas profundas a hiperprofundas demanda novas tecnologias para o aumento de suas reservas e para a gestão ambiental de seus ativos. Uma possível consequência das operações de produção e transporte é o vazamento a partir da cabeça de poços, plataformas, dutos ou embarcações. Ademais, o óleo pode chegar até o assoalho oceânico por processos geológicos, formando uma exsudação. Como o petróleo apresenta densidade menor que a água do mar, manchas de óleo permanecem na superfície oceânica por alguns dias e podem ser detectadas por imagens de satélite. Para monitorar tais feições, emprega-se o Radar de Abertura Sintética (SAR – *Synthetic Aperture Radar*). Apesar das capacidades do SAR, a identificação do tipo de óleo na superfície do mar não é tarefa trivial. O objetivo da pesquisa, portanto, é analisar medidas obtidas a partir do SAR a bordo do satélite RADARSAT-2 para discriminar dois tipos de manchas de óleo (*oil slicks*) observados na superfície do mar: exsudações (*oil seeps*) e derrames artificiais (*oil spills*). A aplicação de *machine learning* nas informações armazenadas em uma base de dados da Baía de Campeche (Golfo do México) pode trazer novos *insights* para a solução do problema. Tal acervo inclui 277 cenas SAR obtidas no período de 2008 a 2012. Especialistas interpretaram nessas imagens cerca de 4.916 polígonos representativos de *oil slicks*, subdivididos em *oil spills* (2.895) e *oil seeps* (2.021). Aproximadamente 80% dos *oil spills* (2.315) foram validados *in situ* pela companhia petrolífera PEMEX, enquanto 20% foram considerados *orphan spills* (580). Por sua vez, 1.438 *oil seeps* (71%) formaram agrupamentos sistemáticos em locais específicos e com repetição no tempo, enquanto 583 deles (29%) foram classificados como *orphan seeps*. Os tipos de atributos associados a cada um dos polígonos são: (a) informação contextual (e.g., classe (*oil seeps* ou *oil spills*); localização geográfica e batimétrica); (b) descritores da cena SAR (e.g., modo de operação, ângulos de incidência); (c) informação de tamanho (e.g., geometria, forma e dimensão); (d) assinatura de retroespalhamento SAR (σ_0 , β_0 e γ_0). No total, cada polígono poderá possuir cerca de 500 diferentes atributos. Técnicas convencionais de Análise Multivariada de Dados, tais como Correlação (*R-mode*), Análise de Principais Componentes (PCA) e Função Discriminante, foram anteriormente utilizadas para diferenciar os tipos de *oil slicks*, alcançando 70% de acurácia total. Pretende-se nesta pesquisa verificar a possibilidade de melhorar tal resultado com o emprego das seguintes tecnologias: Azure (biblioteca de *machine learning* da Microsoft), TensorFlow (biblioteca de *machine learning* da Google), NVIDIA *Deep Learning* SKD (para aplicações em GPU) e Watson/IBM (para inteligência artificial).

PALAVRAS-CHAVE: MACHINE LEARNING, OIL SLICK, RADARSAT-2