

Deteção semi-automática **de áreas verdes** **permanentes**

João Albuquerque

Aluno Mestardo em Engenharia Informatica

Departamento de Informatica

Faculdade de Ciencias da Universidade de Lisboa

jc.albuquerque@campus.fct.unl.pt

Paulo Pires

Tenente Coronel de Cavalaria

Engenheiro Informático

Centro de informação Geoespacial do Exército

ppires@igeoe.pt

Fernando Birra

Professor Doutor

Departamento de Informatica

Faculdade de Ciencias da Universidade de Lisboa

fpb@fct.unl.pt

Carlos Damásio

Professor Doutor

Departamento de Informatica

Faculdade de Ciencias da Universidade de Lisboa

cd@fct.unl.pt

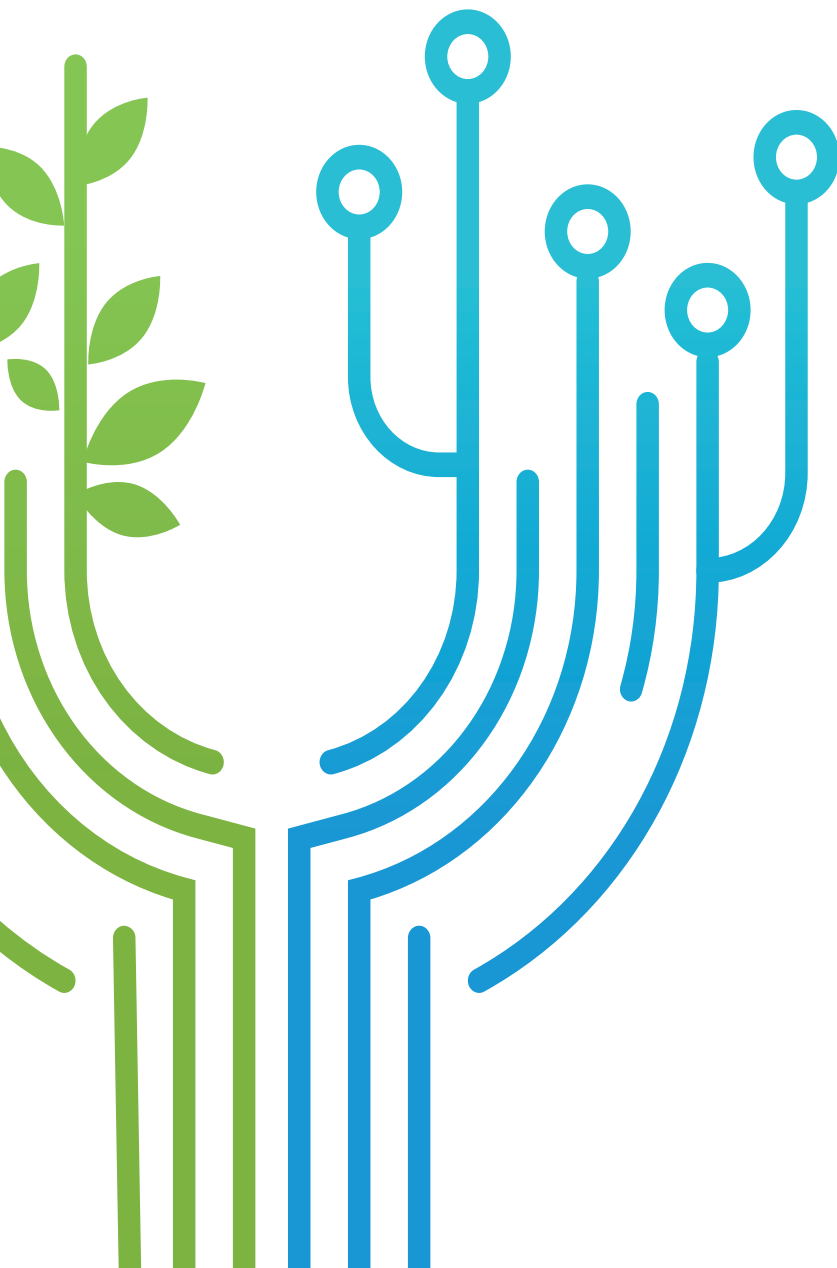
João Moura Pires

Professor Doutor

Departamento de Informatica

Faculdade de Ciencias da Universidade de Lisboa

jmp@fct.unl.pt





O Centro de Informação Geoespacial do Exército (CIGeoE) cartografa o território nacional através da estereorrestituição dos objetos a partir de fotografias aéreas, com posterior validação no terreno, por operadores de fotogrametria qualificados para esse efeito. Um processo demasiado moroso, o que pode afetar o desfaseamento das cartas do exército, em relação à realidade observada.

Para este estudo, dentro de uma área limitada, foram implementadas duas metodologias principais de classificação, desenvolvendo-se para esse efeito três classificadores com excelentes resultados, nomeadamente *XGBoost*, *Random Forest* e *Support Vector Machines*, utilizando diferentes abordagens de classificação. Uma das metodologias implementadas envolve um produto de *Sentinel-1* e *2* da mesma data da classificação original ou com uma diferença temporal muito pequena e uma outra metodologia que envolve o uso de séries temporais com features para a classificação, mantendo todas as normas necessárias garantindo que os objetos adquiridos possuem o mesmo grau de detalhe e rigor. Para o âmbito deste projeto escolheu-se a deteção semiautomática de vegetação permanente (oito classes principais). O algoritmo desenvolvido apresentou resultados muito satisfatórios nas regiões de estudo, constituindo um passo muito importante na automatização da classificação automática da vegetação em Portugal. Com a implementação deste algoritmo na Cadeia de Produção do CIGeoE, todo o processo de aquisição e classificação da vegetação ficará muito mais célere, contribuindo deste modo para a diminuição do desfaseamento temporal entre as fotografias aéreas e a disponibilização da informação aos utilizadores.

Motivação e Contexto

O Centro de Informação Geoespacial do Exército (CIGeoE) provê com informação geoespacial o Exército e outras entidades, bem como desenvolve ações de investigação científica e tecnológica. Constituir-se como referência nacional de excelência no fornecimento de informação geográfica, em particular na produção da Carta Militar de Portugal, na escala 1:25 000 a Carta Base de Portugal. A produção da informação geoespacial da Carta Militar tem por base uma aquisição global dos objetos do mundo real, sendo um dos temas a vegetação, adquirida através de fotografia aérea sendo posteriormente interpretada e vetorizada manualmente por operadores de fotogrametria que classificam a informação geográfica pretendidas de acordo com as Normas de Aquisição, num processo designado por restituição¹.

Devido ao grande volume de informação a processar manualmente pelos operadores, e ao facto de esta informação ser validada no terreno face ao desfaseamento temporal entre as fotografias aéreas e realização dos trabalhos de campo, este processo caracteriza-se como sendo bastante demorado e exigente.

Pretende-se, com a elaboração deste projeto, o desenvolvimento de ferramentas que, recorrendo a algoritmos de aprendizagem automática, permitam a classificação automática da vegetação permanente, a partir de imagens multiespectrais de *Sentinel*, de acordo com as Normas de Aquisição do CIGeoE. A inclusão deste processo na cadeia de produção da Carta Militar, visa permitir a redução drástica do tempo de aquisição dos dados, assim como um maior grau de atualização face aos dados de validação provenientes do trabalho de campo.

A vegetação permanente é definida de acordo com estas normas, seja ela natural ou cultivada, que mantém a sua área vegetacional durante a mudança de estações, sem a presença de

¹ Estas normas são confidenciais, e consequentemente de acesso restrito, sendo-me disponibilizadas para efeitos auxiliares na construção de uma metodologia eficiente.



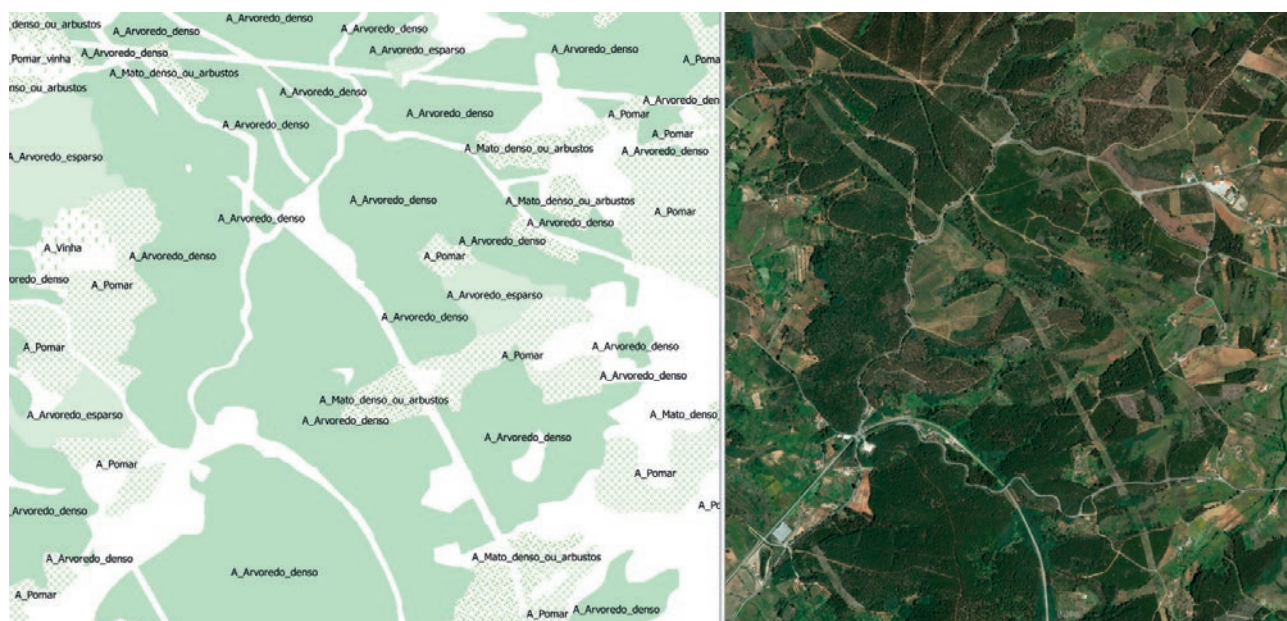


Figura 1 – Exemplo da classificação da vegetação permanente feita pelo CIGeoE

fatores exteriores que possam afetar consideravelmente a sua dimensão (incêndios, desflorestação, etc.). As suas características poderão ser influenciadas ligeiramente com o decorrer do tempo, mas no geral manter-se-ão relativamente consistentes com o decorrer do ano. Estas características tornam este tipo de vegetação relevante para a catalogação cartográfica.

A Figura 1, pretende ilustrar o produto final do CIGeoE comparativamente às ortofotos utilizadas neste projeto para a classificação da vegetação. O objetivo deste artigo é automatizar este processo, desde a deteção remota como a figura da direita à aquisição e classificação automaticamente na figura da esquerda.

Abordagem

DADOS DE ESTUDO

Como *ground truth* o CIGeoE forneceu 4 *shapefiles* relativos a 4 regiões distintas de Portugal. Estes *shapefiles* têm 10 000 por 16 000 metros de área, por *shapefile*, com uma classificação da vegetação permanente de acordo com o contexto da carta militar. Todos estes *shapefiles* têm a sua classificação proveniente da estereorrestituição efetuada a partir de fotografias aéreas obtidas em maio de 2015. Estas regiões são Fundão, Monchique, Sendim do Douro e Praia da Tocha, tendo sido previamente escolhidas por conterem vegetação muito diversificada contribuindo deste modo para uma variabilidade do algoritmo, um excelente ponto de partida para um estudo inicial.

Para a utilização desta informação como dados de input para treino e para uma classificação *pixel* a *pixel* foi necessário adicionar-se uma feature numérica a estes *shapefiles* que correlaciona o tipo de vegetação a um algarismo, funcionando assim com um *id* único de cada classe. Deste modo foi possível rasterizar-se estes *shapefiles* utilizando o seu *id*, no CRS EPSG:32629, tendo sido usados



Figura 2 - Áreas de Interesse

pixels de 10m x 10m. Esta resolução constitui um decréscimo significativo comparativamente à resolução espacial das fotografias aéreas utilizadas pelo CIGeoE, no entanto esta resolução garante que a resolução espacial dos produtos de *Sentinel-1* e *Sentinel-2* sejam alinhados com a resolução espacial das fotografias aéreas.

PRÉ-CLASSIFICAÇÃO

Anteriormente à classificação foi estudada a distribuição das classes nos produtos referentes à *groundtruth*.

Na Figura 3, está representada a distribuição de classes através das quatro regiões de interesse. A distribuição destas classes é muito heterogênea, com uma larga incidência nas classes de No Data (referente a *pixéis* que, ou não são vegetação, ou são vegetação que sai fora do contexto do interesse do estudo) e também nas classes de Arvoredo Denso e Arvoredo Esparso. Por outro lado, as clas-

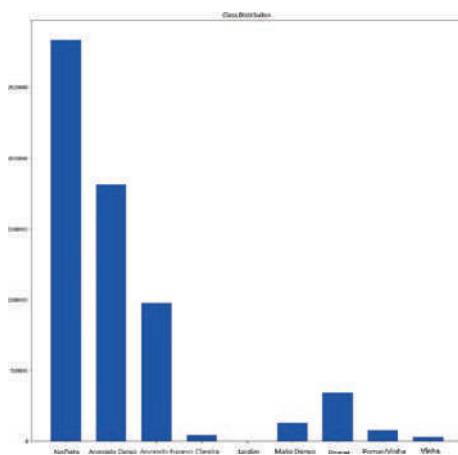


Figura 3 - Distribuição de classes pelas quatro folhas de classificação

ses Mata, Estufa e Sebe ou Valado que inicialmente estava presente nas classes a classificar, não têm qualquer número de amostras nestas folhas sendo assim ignoradas neste trabalho. No entanto, contrariamente ao ponto anterior, surge uma classe previamente desconhecida, a Clareira que, de acordo com as especificações do CIGeoE, representa áreas de vegetação rasteira que se encontram circundadas de Arvoredo. A classe Jardim ou Horta, em mais de quatro milhões de amostras, apenas tem representação em cerca de uma centena dessas amostras e, conseqüentemente, foi também desconsiderada para o algoritmo de classificação.

A Tabela 1, mostra a representação das classes principais do CIGeoE nas amostras contidas na *groundtruth*. Conseqüentemente das 10 classes principais da classificação do CIGeoE serão trabalhadas sete classes com uma classe adicional, a Clareira, perfazendo um total de oito classes a classificar.

A Figura 4, é um exemplo de um produto raster

Classes Principais do CIGeoE	Classes a Utilizar
Sem Representação	No Data
Arvoredo Denso	Arvoredo Denso
Arvoredo Esparso	Arvoredo Esparso
Sem Representação	Clareira
Mata	Sem Representação
Mato Denso ou Arbustos	Mato Denso ou Arbustos
Estufa	Sem Representação
Pomar	Pomar
Vinha	Vinha
Pomar/Vinha	Pomar/Vinha
Sebe ou Valado	Sebe ou Valado
Jardim ou Horta	Representação Muito Diminuta (Descartada)

Tabela 1 - Classes principais do CIGeoE e a sua representação nos dados de estudo

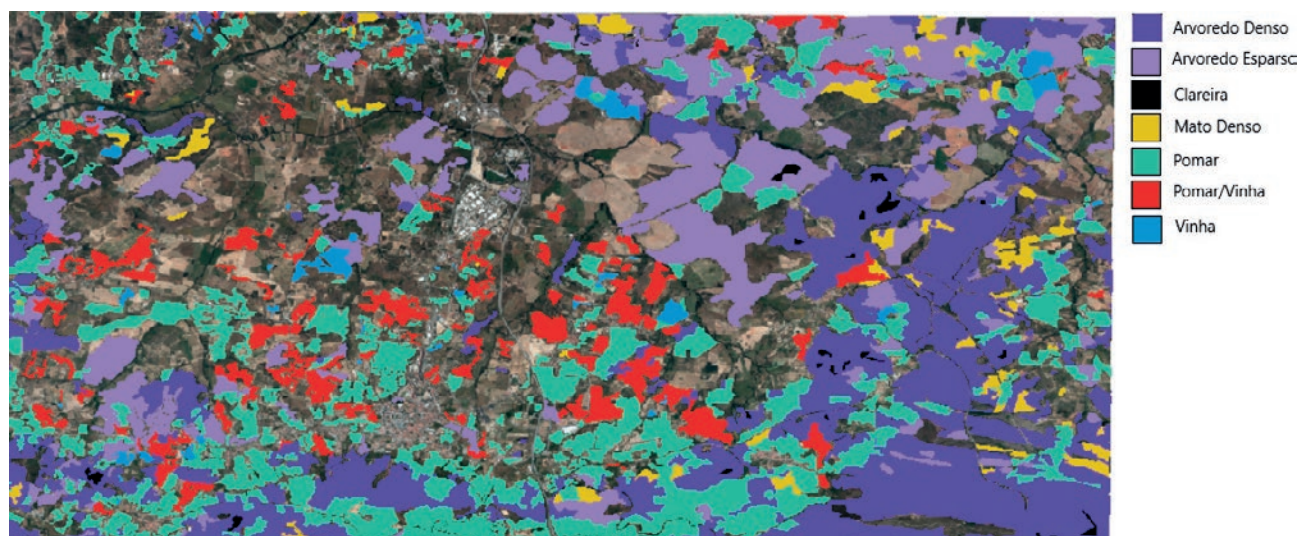


Figura 4 - Produto *raster* resultante da classificação do CIGeoE

resultante da classificação manual efetuada pelo CIGeoE. Cada uma das zonas coloridas representa um tipo de vegetação permanente relevante para a classificação em análise. Esta imagem fornece também uma ajuda visual para verificar a disparidade entre o número de amostras de cada classe e a sua distribuição espacial em cada folha.

VISÃO GERAL DO ALGORITMO

Como fonte de dados de deteção remota foram utilizados produtos de *Sentinel-1* (Imagens SAR) e produtos de *Sentinel-2* (Imagens Espectrais). Estes produtos apresentam excelente resolução espacial (cerca de 10m em algumas bandas espectrais), como um número aceitável de bandas espectrais, constituindo assim um conjunto robusto de características de treino.

No que diz respeito à aprendizagem automática os classificadores utilizados foram o *Support Vector Machines* (SVM), *Random Forest* e *XGBoost*. Estes classificadores apresentam resultados muito consistentes e coerentes em literatura semelhante e por esses mesmos motivos foram os classificadores escolhidos.

A metodologia de classificação de modo geral segue os seguintes passos:

- **Carregamento e Processamento de Dados**

Os dados, tanto as features como a *ground-truth* são carregados em memória utilizando o GDAL, sendo as imagens carregadas a partir do ficheiro e em seguida transformadas em matrizes que cada valor é representativo de um *pixel* dessa mesma imagem. A partir de um ficheiro, no caso das features, cada entrada na matriz representava todos as bandas disponíveis associadas ao *pixel* respetivo. Cada imagem é carregada individualmente.

- **Validação dos Modelos** Inicialmente todos os classificadores foram otimizados utilizando o algoritmo de validação de modelos *Random Search*. O algoritmo de *Random Search* foi implementado com 100 iterações, e utilizando o *Cohen Kappa* com métrica de validação, foi-lhe também transmitido o conjunto total de dados a utilizar pois este algoritmo utiliza internamente *Stratified K-Fold* para dividir o conjunto de dados. Após esta otimização, os modelos ótimos de cada classificador são mantidos para os testes posteriores, evitando assim que se tenha de correr à otimização deste algoritmo sempre que seja preciso proceder à classificação de dados tornando o processo de classificação mais rápido e menos pesado.

- **Treino** Posteriormente à obtenção dos parâmetros ótimos para cada classificador, 20% dos dados relativos a cada imagem são utilizados para treino do modelo de classificação. Estes dados estão estratificados, o que significa que o modelo de treino vai receber 20% dos dados de cada classe, mantendo uma distribuição idêntica à do conjunto total, evitando assim a ausência de classes com fraca representação no conjunto de *input*.
- **Previsão da Imagem** Por fim, os restantes dados respetivos a cada imagem são, individualmente, fornecidos ao algoritmo para a previsão da imagem na totalidade. Após a previsão, uma nova imagem raster é gerada com as previsões geradas pelo algoritmo. São também calculadas métricas de validação para perceber melhor o desempenho geral de cada algoritmo.

SÉRIE ESTÁTICA VS SÉRIES TEMPORAIS

Neste algoritmo foram utilizadas as metodologias distintas temporal estática e séries temporais. A distinção entre estas duas metodologias é que nas séries estáticas apenas um produto de *Sentinel-1* e *Sentinel-2* é usada para o treino e classificação dos dados, enquanto nas séries temporais foram usados diversos produtos de datas diferentes, de um intervalo de um ano, para treino e classificação dos dados.

Para a implementação da metodologia de séries temporais foram criadas quatro variantes procurando uma implementação ideal com resultados excecionais que pudesse ser implementada na linha de produção de cartas militares do CIGeoE. Estas variantes serão brevemente descritas de seguida.

- **Série Temporal Convencional** Esta variante utiliza indiscriminadamente todas as suas *features* no processo de classificação, mantendo as características dos dados o mais próximo do original possível.
- **Série Temporal com Métricas Estatísticas** Esta variante troca os *inputs* fornecidos pelos produtos originais por métricas estatísticas como a mediana, a moda, quantis, calculados a partir dos valores originais, diminuindo assim um número possivelmente absurdo de amostras e mascarando alguns outliers que possam enviesar o treino.
- **Série Temporal com Informação Temporal Incorporada (Tuplo e Peso)** Estas duas variantes procuram integrar os dados temporais referentes a cada produto, com as suas

características espectrais. Para isso duas variantes foram concebidas, uma que integra estes dados como tuplo, e outra que integra a informação temporal como peso.

VALIDAÇÃO DA CLASSIFICAÇÃO

Dada a importância da validação na viabilização do algoritmo, foi implementada uma técnica de validação *pixel a pixel* utilizando métricas muito conhecidas como a matriz de confusão, o *Kappa*, *Accuracy*, *Precision*, *Recall*. Aliada a estas métricas é feita uma avaliação objetiva, não só aos níveis das métricas, mas também ao nível visual do produto raster resultante da classificação.

Nesta técnica, a validação *pixel a pixel*, são extraídos os valores de cada *pixel* do produto utilizado como *groundtruth* e estes valores são comparados com o produto final resultante da classificação. Este é um processo de fácil implementação e compreensão e fornece uma ideia geral do comportamento do algoritmo na classificação de vegetação.

Apresentação de Resultados

Nesta secção serão apresentados e discutidos os resultados das diferentes metodologias usadas para testar e avaliar o algoritmo. Primeiramente serão discutidos os resultados da série temporal estática, depois serão avaliadas e comparadas as diferentes variantes da metodologia de séries temporais, e por fim será feita uma análise à capacidade de generalização deste algoritmo.

RESULTADOS SÉRIE ESTÁTICA

O resultado da classificação está expresso na Tabela 2 utilizando um único produto de *Sentinel-1* e *Sentinel-2* e considerados resultados aceitáveis, onde os classificadores de *Random Forests* (RF) e *XGBoost* apresentam resultados muito semelhantes.

Podemos observar que o classificador de RF apresenta resultados ligeiramente superiores aos

Métricas	<i>Random Forest</i>	<i>XGBoost</i>	<i>SVMs</i>
Overall Accuracy	89%	85%	64%
Overall Precision	93%	89%	29%
Overall Recall	76%	75%	27%
Overall F1Score	83%	82%	26%
Overall Kappa	83%	78%	43%

Tabela 2 - Métricas resultantes da classificação de série temporal estática

resultados do classificador *XGBoost*. Ambos os classificadores apresentam um recall com valores visivelmente inferiores aos valores das outras métricas, o que, por conseguinte, significa que estes classificadores têm problemas com a elevada discrepância no número de amostras das diferentes classes referido anteriormente, na secção “*Pré-Classificação*”, resultando, numa elevada percentagem, as classes minoritárias ficarem mal classificadas.

Por outro lado, o classificador *Support Vector Machines* (SVMs) apresenta resultados muito fracos, e, por conseguinte, será descartado em avaliações futuras. Estes fracos resultados são consequência de uma fraca capacidade em lidar com um número tão elevado de amostras (cerca de 1,5 milhões por produto).

Após o estudo destes resultados decidiu-se, estudar o desempenho do algoritmo (utilizando o algoritmo *Random Forest* devido às suas métricas relativamente superiores) por classes individuais. Deste modo possibilita-se o estudo, mais aprofundado do comportamento deste algoritmo.

Classes	Precision	Recall	F1Score
NoData	88%	96%	92%
Arvoredo Denso	89%	92%	90%
Arvoredo Esparso	91%	74%	80%
Clareira	99%	60%	75%
Mato Denso	97%	65%	78%
Pomar	86%	81%	84%
Pomar/Vinha	95%	66%	78%
Vinha	97%	78%	87%

Tabela 3 - Classificação por classe utilizando o classificador *Random Forest* com série estática

Com o estudo adicional da Tabela 3, pode-se afirmar que apesar de os resultados serem satisfatórios existe uma diferença significativo na métrica recall em algumas classes. Isto deve-se à vasta diferença entre número de amostras entre as classes com maior número de amostras e classes com menor número de amostras (descritas em “*Pré-Classificação*”). As classes Clareira, Mato Denso e Pomar/Vinha têm um número de amostras muito diminuto comparativamente a classes como Arvoredo Denso e Arvoredo Esparso. Esta disparidade introduz confusão na classificação pois grande parte destas amostras estão a ser classificadas como classes maioritárias, justificando assim valores de recall mais reduzidos. Adicionalmente observa-se que as classes No Data,

Arvoredo Denso e Pomar/Vinha são as únicas classes com valores de precision inferiores a 90%, pois existem amostras de outras classes que são classificadas consistentemente como amostras destas classes, refletindo mais uma vez o problema relatado anteriormente, da dificuldade na distinção entre classes com números muito discrepantes de amostras.

RESULTADOS SÉRIES TEMPORAIS

Com a adição de séries temporais, conseguiu-se uma melhoria significativa face aos resultados obtidos na metodologia temporal estática. Para isso fez-se uma análise qualitativa das quatro variantes criadas para esta metodologia de modo a maximizar o desempenho.

Métricas	S.T. Convencional	S.T. com Métricas Estatísticas	S.T. Informação Temporal como Tuplo	S.T. Informação Temporal como Peso
Accuracy	94%	93%	94%	94%
Precision	97%	96%	96%	97%
Recall	86%	83%	85%	87%
F1Score	91%	89%	90%	91%
Kappa	91%	88%	90%	91%

Tabela 4 - Validação dos diferentes modelos de séries temporais

Com a comparação dos resultados obtidos na Tabela 4, observa-se que todas as variantes de metodologia de séries temporais apresentam uma melhoria muito significativa relativamente à metodologia convencional de série estática. No entanto duas variantes destacam-se das outras em termos de desempenho: a série temporal convencional e a série temporal convencional usando informação temporal como peso. Estas duas variantes apresentam resultados muito semelhantes no entanto existe uma pequena vantagem em usar a variante de informação temporal como peso, pois esta pode dar mais peso a produtos com datas com menor incidência de produtos viáveis, (o mês de Janeiro costuma apresentar uma maior cobertura de nuvens e por isso um número inferior de ima-

Métricas	Random Forest	XGBoost
Overall Accuracy	95%	97%
Overall Precision	97%	97%
Overall Recall	90%	95%
Overall F1Score	95%	97%
Overall Kappa	93%	96%

Tabela 5 - Métricas resultantes da classificação de séries temporais

gens espectrais viáveis) característica que não se apresenta numa metodologia de séries temporais convencional.

Quando se utiliza esta metodologia para a classificação da vegetação das regiões de interesse, observam-se resultados muito satisfatórios. Segundo a Tabela 5 a utilização de diversos produtos de *Sentinel-1* e *Sentinel-2* constitui uma melhoria muito significativa na classificação de vegetação. Todas as métricas apresentam uma melhoria muito positiva, mais notavelmente na métrica recall, subindo para valores acima dos 90%, o que significa que a confusão entre classes é muito inferior e apesar de ainda estar presente, tem uma presença muito diminuta na classificação de vegetação.

Comparando os dois classificadores, o *XGBoost* tem um desempenho relativamente superior ao classificador Random Forest, no entanto este classificador apresenta um tempo de execução significativamente inferior com cerca de 57 minutos face ao tempo de 269 minutos por parte do algoritmo *XGBoost* (reduzível para 89 minutos utilizando GPU multi-threading). Estes resultados oferecem uma dinâmica interessante de trade-off entre desempenho e viabilidade, com o algoritmo *XGBoost* apresentando tempo de execução superior, mas resultados ótimos, e o classificador Random Forest apresentando resultados bastante bons, mas ligeiramente inferiores ao *XGBoost* mas com um tempo de execução muito reduzido.

Para o estudo do desempenho por classe, desta vez utilizou-se o classificador *XGBoost* devido às suas métricas muito promissoras.

Com a observação e estudo da Tabela 6 podemos observar que a confusão entre classes maioritárias e minoritárias muito presente na Tabela 3 é muito reduzida utilizando séries temporais. Apesar de se verificar um comportamento semelhante à Tabela 3 em relação à distribuição de valores na generalidade existe um aumento muito significativo nos valores em cima, nomeadamente a métrica recall mantém-se sempre acima dos 90%. Isto significa

Classes	Precision	Recall	F1Score
NoData	97%	98%	98%
Arvoredo Denso	97%	98%	97%
Arvoredo Esparso	97%	97%	97%
Clareira	96%	91%	94%
Mato Denso	98%	93%	95%
Pomar	96%	95%	96%
Pomar/Vinha	97%	93%	95%
Vinha	98%	95%	96%

Tabela 6 - Classificação por classe utilizando o classificador *XGBoost* com séries temporais

que com séries temporais mais amostras de classes minoritárias são introduzidas o que permite ao algoritmo discernir e classificar as classes pretendidas com maior facilidade.

utilizados pelo CIGeoE na classificação da vegetação permanente.

Conclusão

O presente artigo pretendeu descrever de uma forma sucinta todo o desenvolvimento de uma ferramenta para a aquisição e classificação da vegetação permanente no contexto da carta militar do CIGeoE.

O tema vegetação consiste em 8 classes principais distribuídas inicialmente por quatro folhas correspondendo às regiões de interesse Fundão, Monchique, Sendim do Douro e Tocha. As classes têm uma distribuição altamente desequilibrada constituindo um obstáculo importante no treino e classificação desta vegetação.

Com esse objetivo adquiriram-se produtos de

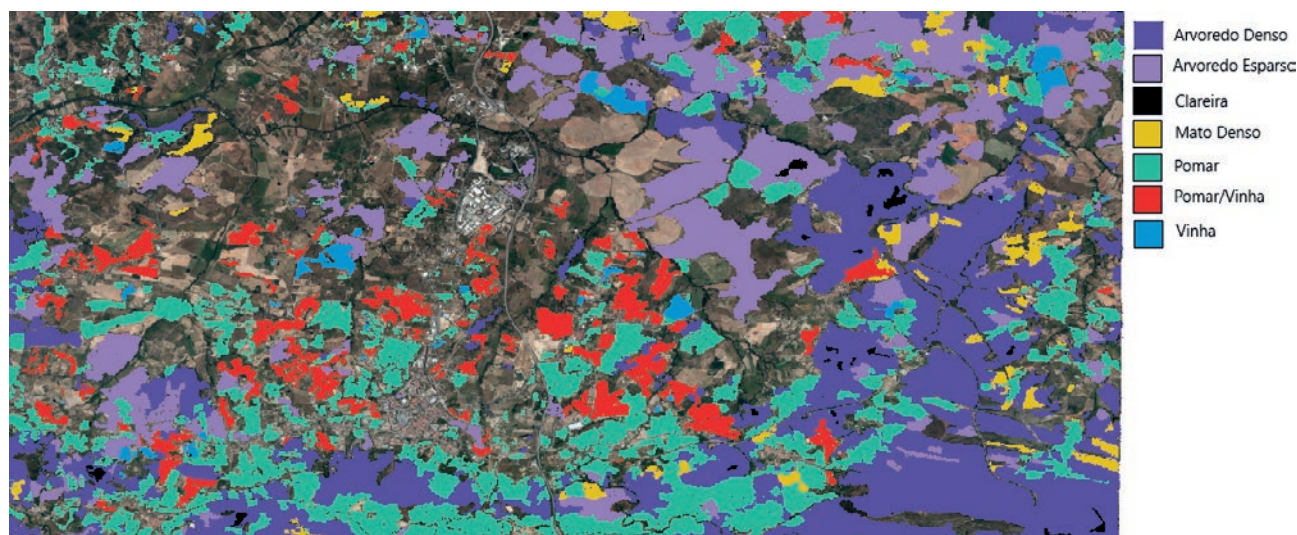


Figura 5 - Produto raster resultante da classificação automática usando metodologia de séries temporais e classificador *XGBoost*

No estudo da Figura 5 e a sua subsequente comparação com a Figura 4 podemos observar que as imagens são muito idênticas na sua classificação, reforçando e salientando o bom desempenho que esta metodologia teve na classificação de vegetação permanente. Existem, no entanto, algumas áreas a serem melhoradas pois ainda existe algum ruído em zonas de vegetação pertencentes a classes minoritárias.

Esta imagem representa um exemplo de um produto final resultante do algoritmo utilizado. Este é um produto raster com as mesmas características descritas em 2.1. Após a criação deste produto é ainda necessária a sua vectorização criando-se assim os *shapefiles* finais, que irão ser

Sentinel-1 e *Sentinel-2* e desenvolveram-se três classificadores que apresentam muito bons resultados nas técnicas de literatura, nomeadamente *XGBoost*, *Random Forest* e *Support Vector Machines*, utilizando diferentes abordagens de classificação.

Foram implementadas duas metodologias principais de classificação. Uma metodologia envolvendo apenas um produto de *Sentinel-1* e 2 da mesma data da classificação original ou com uma diferença temporal muito pequena. A segunda metodologia envolve o uso de séries temporais com features para a classificação, para isso foram utilizados diversos produtos de *Sentinel-1* e 2 de diversas datas ao longo dos anos de 2015 e 2016.

Numa avaliação inicial revelou-se que a utilização apenas um produto de *Sentinel-1* e *Sentinel-2* era insuficiente para se obter resultados suficientemente precisos, devido ao elevado desequilíbrio entre o número de amostras de diferentes classes provocando confusão significativa na classificação. Os classificadores RF e *XGBoost* apresentaram um desempenho muito semelhante sem grande diferenciação por parte de ambos.

Contrariamente, o classificador SVM revelou-se incapaz de lidar com um número tão elevado de dados, apresentando métricas muito inferiores aos seus classificadores concorrentes.

Consequentemente com o acréscimo significativo de dados na metodologia de séries temporais o classificador SVM foi desconsiderado.

Após a implementação e validação da metodologia de séries temporais, verificou-se uma melhoria muito positiva no desempenho dos classificadores nomeadamente o *XGBoost* que, apesar de ter um tempo muito superior para treino e classificação, apresenta métricas muito satisfatórias e que estão de acordo com os padrões do CIGeoE. Comparativamente com outros trabalhos de classificação automática de vegetação, este algoritmo apresenta resultados acima da média principalmente quando é considerada a diversidade de classes de vegetação que estão a ser estudadas nesta classificação.

Conclui-se, com este trabalho, que o algoritmo desenvolvido tem resultados muito satisfatórios nas regiões de estudo constituindo um primeiro passo muito importante na automatização da classificação automática da vegetação em Portugal. Com a implementação deste algoritmo na Cadeia de Produção do CIGeoE, todo o processo de aquisição e classificação da vegetação ficará muito mais célere, contribuindo deste modo para a diminuição do desfasamento temporal entre as fotografias aéreas e a disponibilização da informação aos utilizadores.

RECOMENDAÇÕES PARA TRABALHO FUTURO

- Generalização Espacial Apesar de resultados muito promissores nestas regiões de estudo, após alguns estudos preliminares na generalização para áreas maiores revelaram-se obstáculos maiores na classificação de vege-

tação permanente. Nomeadamente uma grande variabilidade nas classes do CIGeoE. Consequentemente um dos principais pontos de partida para melhorias futuras é conseguir-se obter uma classificação proveniente de generalização temporal aceitável.

- Segmentação de Imagem Com segmentação de imagem pode-se, não só mitigar o problema de desequilíbrio de entre as classes, mas também reduzir o ruído provocado por uma classificação *pixel a pixel*. Por isto mesmo estudar um processo de segmentação das imagens é recomendado.
- Produtos de Satélite com Maior Resolução Espacial Com maior resolução espacial será possível aumentar o número de amostras no geral, e consequentemente possivelmente diminuir a percentagem de dados para treino, como obter informação mais rica e mais fácil de catalogar e classificar.
- Distinguir Estruturas Urbanas Será interessante utilizar estes classificadores para diferentes problemas ainda no contexto da Carta Militar como a distinção de estruturas urbanas ou artificiais, com o intuito de automatizar o máximo possível o processo de criação de cartas militares.
- Utilização de features Adicionais Incluir features adicionais como texturas e modelos digitais do terreno no processo de classificação utilizando séries temporais. Devido à falta de memória RAM da máquina utilizada, não foi possível incluir estas features nesta classificação, no entanto estas features tiveram uma influência muito positiva na classificação temporal estática de vegetação e por isso será interessante o estudo destas mesmas features numa classificação de séries temporais.
- Estudar as capacidades de generalização temporal dos classificadores Caso a generalização espacial não tenha resultados satisfatórios, seria interessante estudar a capacidade que os classificadores teriam em generalizar a classificação treinando com imagens de um ano e classificando imagens de anos a seguir.