

Estimativa da geração instantânea (nowcasting) de painéis solares a partir de imagens do céu

Gustavo Friol Bento
Universidade Federal de Santa Catarina
Graduação em Engenharia Elétrica
Florianópolis, Brasil
gustavofriol@gmail.com

Resumo—O presente texto serve como relatório para o projeto final da disciplina de aprendizado de máquina do professor Danilo Silva.

Index Terms—Aprendizado de Máquina, CNN, SUNSET, Nowcast, Geração fotovoltaica

I. INTRODUÇÃO

A volatilidade das condições meteorológicas faz com que a previsão a curto prazo da geração solar seja extremamente desafiadora. Mudanças repentinas na cobertura de nuvens podem ocasionar variações significativas na geração solar numa escala de minutos, o que se torna um empecilho na inserção dessa fonte nas redes de energia ao redor do mundo. Em condições parcialmente nubladas, por exemplo, a potência gerada por painéis fotovoltaicos pode praticamente zerar em menos de um minuto.

Em [1], Sun et al. comprovam que imagens do céu capturadas localmente podem fornecer informações relevantes na previsão da geração de um sistema fotovoltaico. Dentro da imagem, a posição do sol se correlaciona diretamente com a insolação teórica. A distribuição, a cor e a transparência das nuvens afetam a absorção e a reflexão da luz. O movimento e a forma das nuvens até mesmo prenunciam futuras imagens do céu e, consequentemente, a futura produção de energia fotovoltaica. Tais informações são difíceis de extrair com modelos explícitos, o que tem motivado um grupo recente de pesquisadores a empregar técnicas de machine learning para capturar essas relações, visando a uma previsão a curto prazo mais confiável.

A. Trabalhos Relacionados

Em [1], Sun et al. propõem uma rede neural convolucional para o "nowcast" da geração solar: uma tarefa que consiste em estimar a potência gerada no instante t_0 com base em imagens do céu capturadas no mesmo momento. A CNN foi batizada como SUNSET (Stanford University Neural network for Solar Electricity Trend) e sua arquitetura é mostrada na Figura 1. Ela compartilha de um framework parecido com os de CNNs famosas como a AlexNet, VGG, Inception, e ResNet: a imagem de entrada é passada sequencialmente por várias estruturas de Convolução-Pooling (Conv-Pool), cada uma consistindo em um conjunto de várias camadas convolucionais e uma camada de pooling ao final. Após a última estrutura Conv-Pool, a entrada processada é achatada (flattened) em um

vetor e alimentada em várias camadas totalmente conectadas (fully connected). As camadas totalmente conectadas geram um resultado de regressão (para prever dados contínuos) ou um score de classificação (para prever dados discretos ou categóricos). No caso da SUNSET, logicamente, as camadas finais geram um resultado de regressão.

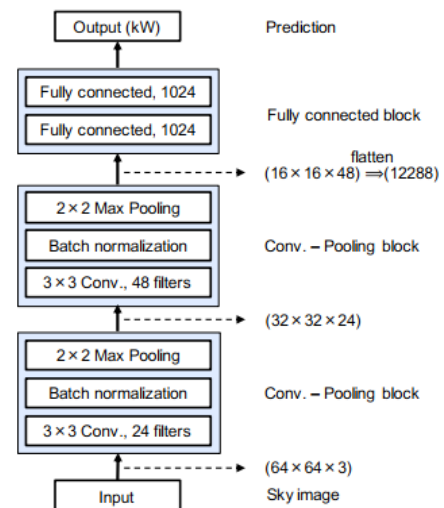


Figura 1: Arquitetura do modelo SUNSET - Nowcast [2]

Com a comprovação de que o modelo SUNSET era capaz de correlacionar geração solar com o estado atual do céu, o mesmo grupo de pesquisadores foi instigado a investigar a capacidade do modelo realizar uma previsão de curto prazo da geração solar dos mesmos painéis do primeiro estudo. Dessa forma, em [3] a arquitetura do SUNSET foi adaptada de forma que o modelo fosse capaz de realizar uma predição com 15 minutos de antecedência para a geração dos painéis. Agora, ao invés de receber apenas imagens do céu em um instante t_0 , o modelo também recebe os dados de geração dos 15 minutos anteriores a t_0 e, com isso, produz a sua predição em kW, obtendo um RMSE de teste de 2.51.

Já em 2022, com o objetivo de acelerar e incentivar a pesquisa de diferentes métodos e modelos para previsão da geração solar baseada em imagens do céu, o grupo decidiu criar um dataset de benchmark chamado SKIPP'D - SKy Images and Photovoltaic Power Generation Dataset. Em 2023,

o paper do dataset foi publicado pela Solar Energy [2] e apresentou os resultados do SUNSET para as tarefas de nowcast e forecast, porém com o modelo treinado e avaliado com o SKIPP'D. Para o nowcast, o SUNSET atingiu um RMSE de teste de 2.4, enquanto para o forecast o valor foi de 3.03.

II. MATERIAL E MÉTODOS

Como o título já diz, a proposta do projeto é avaliar o desempenho de uma CNN para a tarefa de nowcast. Em particular, o primeiro objetivo do projeto foi reproduzir os resultados para o nowcast apresentados em [2].

Nas seguintes subseções, será apresentado o conjunto de dados, bem como os métodos e modelos utilizados.

A. Conjunto de Dados

Como foi introduzido anteriormente, o conjunto de dados utilizado neste projeto foi o SKIPP'D (SKy Images and Photovoltaic Power Generation Dataset), um dataset público e padronizado criado para facilitar a pesquisa e o benchmarking de modelos de previsão de energia solar baseados em imagens. Os dados foram coletados pelo grupo de pesquisa Environmental Assessment and Optimization (EAO) na Universidade de Stanford, Califórnia, durante o período de março de 2017 a dezembro de 2019. A localização possui um clima mediterrâneo com verões quentes, caracterizado por longos períodos de céu claro e invernos mais curtos com céu parcialmente nublado.

O dataset possui uma versão com os dados brutos com imagens de alta resolução (2048x2048) para que o usuário possa fazer o próprio processamento das imagens e decidir qual resolução quer utilizar em seus experimentos. Contudo, para o projeto foi utilizada a versão de benchmark do dataset, que consiste em pares de imagens do céu e dados de geração FV com uma frequência de 1 minuto. As imagens de alta resolução são processadas e redimensionadas para uma resolução de 64x64, sendo capturadas por uma câmera fish-eye de 360 graus. Os dados de geração são concorrentes às imagens e foram coletados de um sistema de painéis solares com capacidade de 30.1 kW-DC, localizado a aproximadamente 125 metros da câmera.

1) *Pré-processamento*: Para criar a versão de benchmark, os dados brutos passaram pela seguinte pipeline de processamento: inicialmente, as imagens de alta resolução foram extraídas de vídeos com uma frequência de 1 minuto e, posteriormente, redimensionadas para a resolução final de 64x64. Durante o processo, registros anormais de geração FV, como valores negativos ou falhas longas, foram filtrados, assim como imagens que se repetiam devido a um erro no decodificador de vídeo. Ao final, as imagens processadas foram pareadas com os dados de geração FV correspondentes ao mesmo instante.

2) *Particionamento dos Dados*: O conjunto de dados foi particionado em um conjunto de desenvolvimento (para treino e validação) e um conjunto de teste. O conjunto de teste foi criado a partir da seleção manual de 10 dias ensolarados e 10 dias nublados, distribuídos entre os anos de 2017 e 2019. O

restante dos dados compõe o conjunto de desenvolvimento, resultando em 349.372 amostras para desenvolvimento e 14.003 para teste, uma proporção aproximada de 96% para 4%.

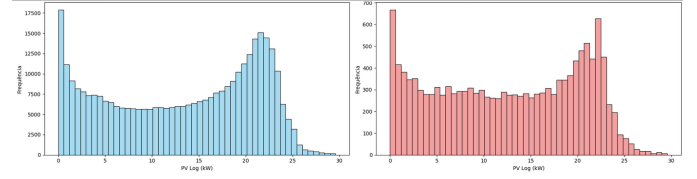


Figura 2: Histogramas da geração FV dos conjuntos de desenvolvimento (azul) e teste (vermelho).

3) *Arquivos*: O dataset de benchmark se encontra disponível publicamente no repositório da Universidade de Stanford [4]. Na Tabela I são apresentados os arquivos disponíveis para download.

Tabela I: Descrição dos Arquivos do Dataset de Benchmark

Arquivo	Tamanho	Descrição
2017_2019_images_pv_processed.hdf5	4,47 GB	Arquivo com uma estrutura semelhante a um diretório de arquivos para armazenar os conjuntos de desenvolvimento e teste.
times_test.npy	433.7 kB	Array NumPy de timestamps correspondentes ao conjunto de teste no arquivo .hdf5.
times_trainval.npy	10.83 MB	Array NumPy de timestamps correspondentes ao conjunto de desenvolvimento no arquivo .hdf5.

Os arrays Numpy de teste e desenvolvimento possuem o mesmo tamanho dos seus respectivos conjuntos no arquivo .hdf5, onde cada time stamp dentro do array possui o mesmo índice que o seu respectivo par foto-geração dentro do dataset armazenado no .hdf5. Abaixo segue um exemplo de amostra retirada dos arquivos, correspondendo a uma foto com sua geração FV e time stamp correspondentes.

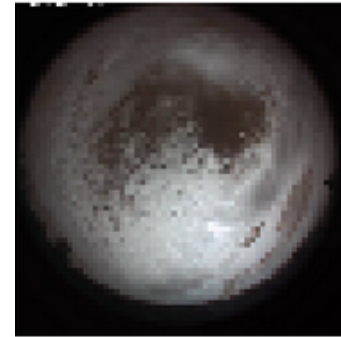


Figura 3: Amostra do SKIPP'D.

- Time Stamp: 2018-10-01 15:12:30
- Geração: 11.32 kW

B. Metodologia Experimental

O projeto compara o desempenho de dois modelos de CNN: o modelo de referência SUNSET, mostrado na Figura 1, e um modelo que usa como backbone a arquitetura da ResNet50V2. Para este último, foi aplicada a técnica de fine-tuning, carregando os pesos referentes ao pré-treinamento do modelo usando a ImageNet. Ambos os modelos são treinados usando o otimizador Adam e a função perda a ser minimizada é definida como Erro Quadrático Médio (MSE - Mean Squared Error):

$$\text{Loss} = \frac{1}{N} \sum_{i \in S} (P_i - G_i)^2, \quad (1)$$

onde S é o conjunto de amostras, $N = |S|$ é o número de amostras, P_i é a saída predita e G_i é a saída real para a amostra i .

1) *SUNSET*: Durante a etapa de desenvolvimento, adotou-se o método de validação cruzada GroupKFold. Nesta abordagem, as amostras são agrupadas por dia, e a divisão dos k folds é realizada sobre esses grupos. Essa estratégia é crucial para garantir que todas as amostras de um mesmo dia permaneçam exclusivamente no conjunto de treino ou de validação, prevenindo o vazamento de dados e fornecendo uma estimativa de desempenho mais robusta.

Para o modelo SUNSET, diversas combinações de taxa de aprendizado e batch size foram inicialmente testadas, de forma a escolher a que melhor se adequava ao problema. Com a melhor combinação definida, foi realizado um treinamento final com validação cruzada de 10 folds, alinhando-se à metodologia dos trabalhos de referência [1], [2].

Como já foi introduzido, o clima da Califórnia, de onde os dados foram coletados, é caracterizado por longos períodos de céu claro e invernos mais curtos com o céu parcialmente nublado. Isso resulta em um conjunto de dados desbalanceado, com uma prevalência de dias ensolarados, que são mais fáceis para o modelo aprender e prever. Dessa forma, a validação cruzada k -fold se apresenta como uma estratégia robusta para obter uma estimativa de desempenho mais justa.

Por fim, a predição final para cada amostra do conjunto de teste é a média das predições dos 10 submodelos treinados em cada fold. Por se tratar de 10 folds, os submodelos são treinados com 90% dos dados de desenvolvimento, o que significa que a predição final é composta por 10 submodelos menos enviesados, pois são treinados com quase todos os dados disponíveis.

2) *ResNet50V2*: A segunda arquitetura avaliada utiliza o modelo ResNet50V2 como base, com pesos pré-treinados na base de dados ImageNet. Para adaptá-lo à tarefa de regressão, sua base convolucional foi instanciada com as seguintes configurações: as camadas de classificação foram removidas, o formato de entrada foi definido para $64 \times 64 \times 3$, e uma camada de pooling médio global foi aplicada para vetorizar os mapas de features. Adicionalmente, o modelo incorpora a função `preprocess_input`, específica da família ResNetV2, para

normalizar os dados de entrada conforme o esperado pela arquitetura.

Sobre esta base, foi adicionada uma camada de Dropout com uma taxa de 0.2 para regularização, seguida por uma camada final densa com uma única unidade para produzir a saída da predição.

A estratégia de treinamento adotada foi a de fine-tuning direto. Os pesos de toda a base convolucional foram descongelados e tornados treináveis, com exceção das camadas de Batch Normalization, que foram mantidas congeladas para preservar suas estatísticas de média e variância aprendidas com a ImageNet.

Sobre a etapa de desenvolvimento, é importante ressaltar que o modelo final possui aproximadamente 23.5 milhões de parâmetros treináveis, mais ou menos 10 milhões a mais que o modelo SUNSET. Fato esse, motivou uma adaptação no número de folds da estrutura de validação cruzada GroupK-Fold empregada anteriormente: em vez de 10, utilizaram-se 5. Esta decisão se deve ao custo computacional significativamente maior da arquitetura ResNet, cujo tempo de treinamento por época foi de aproximadamente 150 segundos, mais que o dobro dos 60 segundos registrados para o modelo SUNSET. Apesar da perda de robustez estatística da avaliação, é um trade-off justo pra viabilizar o experimento. Ainda assim, 5 folds significa que cada submodelo gerado é treinado com uma porção substancial (80%) dos dados de desenvolvimento. Portanto, o viés de cada submodelo utilizado para a predição final no conjunto de teste continua baixo.

C. Métricas

As métricas de avaliação utilizadas foram a Raiz do Erro Quadrático Médio (Root Mean Squared Error - RMSE) e o Erro Absoluto Médio (Mean Absolute Error - MAE):

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i \in S} (P_i - G_i)^2}, \quad (2)$$

onde S é o conjunto de amostras, $N = |S|$ é o número de amostras, P_i é a saída predita e G_i é a saída real para a amostra i .

Ao final, também foi utilizado o Erro Médio Absoluto

$$\text{MAE} = \frac{1}{N} \sum_{i \in S} |P_i - G_i|, \quad (3)$$

onde S é o conjunto de amostras, $N = |S|$ é o número de amostras, P_i é a saída predita e G_i é a saída real para a amostra i .

III. RESULTADOS

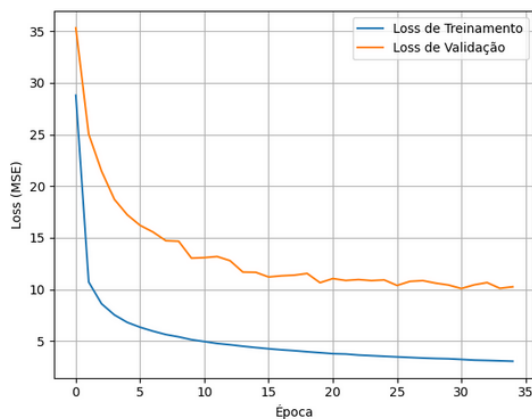
A seguir serão apresentados os resultados dos métodos de treinamento, validação e teste introduzidos na seção anterior.

A. SUNSET

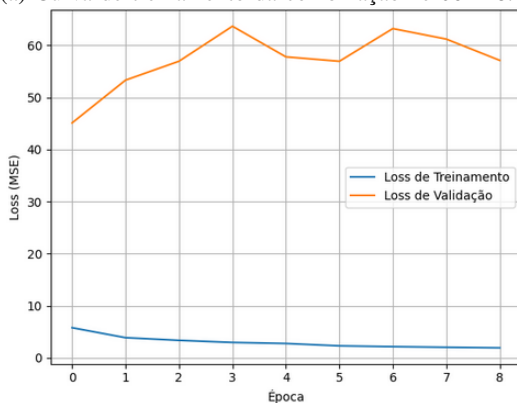
1) *Otimização dos Hiperparâmetros*: Para otimização da taxa de aprendizado e batch size, foi empregado um grid search utilizando validação k-fold com 2 folds apenas. Apesar da baixa robustez dessa abordagem, o objetivo era descartar combinações que claramente não se adequavam ao problema. Segue abaixo uma tabela resumindo o valor da loss de validação média entre os dois folds para cada combinação, bem como as curvas de treinamento e validação ao longo das épocas para a melhor combinação e a pior combinação.

Tabela II: Resultados do Grid Search para a otimização dos hiperparâmetros do modelo SUNSET.

Taxa de Aprendizado	Batch Size	Perda Média (MSE)
1e-06	128	9.9525
1e-06	256	10.8420
1e-05	256	14.1517
1e-05	128	16.2486
0.001	32	25.9229
0.001	64	30.3609
0.0001	32	34.1772
0.0001	64	43.4011



(a) Curva de treinamento da combinação 1e-06-128.



(b) Curva de treinamento da combinação 1e-04-64.

Figura 4: Comparativo das curvas de perda para a melhor e pior combinação de hiperparâmetros.

A análise dos resultados da Tabela II indica que o desempenho do modelo é altamente sensível à taxa de aprendizado, mas demonstra baixa sensibilidade à variação do tamanho do lote (batch size) na faixa testada. Taxas de aprendizado mais altas resultaram em um erro significativamente maior, cumprindo o objetivo primário do grid search de eliminar as piores configurações do espaço de busca. Especificamente para a taxa de aprendizado de 1e-6, a variação no desempenho entre os batch sizes de 128 e 256 foi mínima. Portanto, diante deste cenário, foi selecionada a combinação que apresentou empiricamente o melhor resultado: uma taxa de aprendizado de 1e-6 e um batch size de 128.

2) *Treinamento Final*: Uma limitação metodológica na etapa de otimização de hiperparâmetros foi a estratégia de criação dos datasets para a validação cruzada. A abordagem adotada apresenta uma ineficiência significativa de memória ao exigir o carregamento prévio de todas as imagens em um array NumPy. Esse array é então filtrado com os índices correspondentes às datas de treinamento/validação para o fold em questão e serve como entrada de uma função específica para criar o `tf.data.Dataset`. Para o caso do treinamento, a prática é embaralhar os dados antes de separá-los em batches. Essa abordagem carrega o seguinte problema: ao tentar fazer um embaralhamento completo do conjunto, o tamanho do buffer alocado para a operação atingia o limite de memória RAM do Kaggle, o que forçou a limitação do tamanho do buffer em 5000 amostras, um número significativamente menor que o número total de amostras do conjunto de treinamento. Como é possível observar nas curvas apresentadas na Figura 4, mesmo a melhor combinação de hiperparâmetros resultou em um certo nível de overfitting.

Para corrigir as deficiências da estratégia inicial, a pipeline de dados foi reestruturada para um carregamento sob demanda. O novo processo consiste em criar um dataset a partir dos índices das amostras de cada fold, que são então embaralhados e agrupados em batches. Uma função de mapeamento subsequente utiliza cada batch de índices para carregar dinamicamente os dados correspondentes do arquivo HDF5. Esta estratégia resultou em um consumo de memória drasticamente menor e, ao permitir um embaralhamento completo do dataset a cada época, foi crucial para corrigir o problema de overfitting observado anteriormente.

Feito essa correção, o modelo foi submetido a uma validação cruzada de 10 folds, cujos resultados são apresentados na Tabela III. Para cada fold, o treinamento foi configurado para um máximo de 50 épocas e gerenciado por um conjunto de callbacks para otimização do processo. Foi empregado um mecanismo de EarlyStopping com uma paciência de 5 épocas para encerrar o treinamento caso a perda de validação não apresentasse melhora, bem como um callback ReduceLROnPlateau com paciência de 3 épocas para ajustar dinamicamente a taxa de aprendizado. Em todos os folds ambos os callbacks foram acionados.

Podemos observar na Figura 5 o efeito do embaralhamento completo para o controle do overfitting. Ao embaralhar

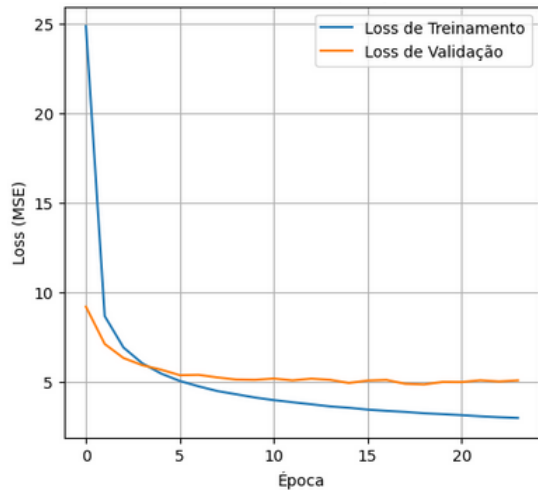


Figura 5: Curva de treinamento para o fold 4.

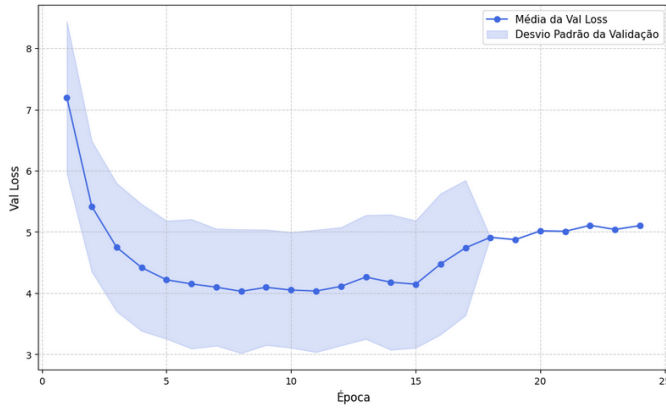


Figura 6: Curva de desempenho médio da validação cruzada.

completamente o conjunto de treinamento antes de agrupá-lo em batches, o modelo é forçado a aprender padrões que são verdadeiros independentemente da ordem em que são apresentados. Cada lote de dados que o modelo vê é agora uma amostra aleatória e representativa de todo o conjunto de treino. Um único batch pode conter imagens de um dia ensolarado de verão, um dia nublado de inverno e um dia de primavera

Tabela III: Resumo do desempenho por fold, considerando o ponto de menor perda de validação para cada modelo.

Fold	RMSE Treino	RMSE Validação	Melhor Época
1	1.9345	2.3971	12
2	2.0387	1.7304	11
3	2.1941	1.7375	8
4	1.8073	2.2078	19
5	2.0419	1.7398	10
6	2.0169	1.7462	12
7	1.9797	2.0771	12
8	2.0948	2.1477	8
9	2.1608	2.0737	8
10	2.2682	1.7404	6
Média	2.0537	1.9598	—

com nuvens variáveis, o que beneficia gradientes mais estáveis e generalistas: o modelo não faz ajustes extremos para uma única condição, mas sim pequenos ajustes que beneficiam a previsão de todas as condições.

3) *Resultados de Teste*: Como já apresentado, a previsão final para cada amostra do conjunto de teste é a média das previsões dos 10 submodelos treinados em cada fold da validação cruzada. Em cima dessas previsões finais, é calculado o RMSE geral. Contudo, para uma análise mais detalhada, essa avaliação foi estratificada por condição climática: tanto os dados de teste quanto as previsões foram divididos entre dias ensolarados e nublados, e o RMSE foi calculado individualmente para cada cenário a fim de aferir a robustez do modelo em diferentes condições de céu. Nessa etapa também é apresentado o MAE das previsões.

Tabela IV: Desempenho final do modelo no conjunto de teste, estratificado por condição do céu.

Condição do Céu	RMSE	MAE
Dias Ensolarados	0.865	0.718
Dias Nublados	3.356	2.361
Geral	2.453	1.541

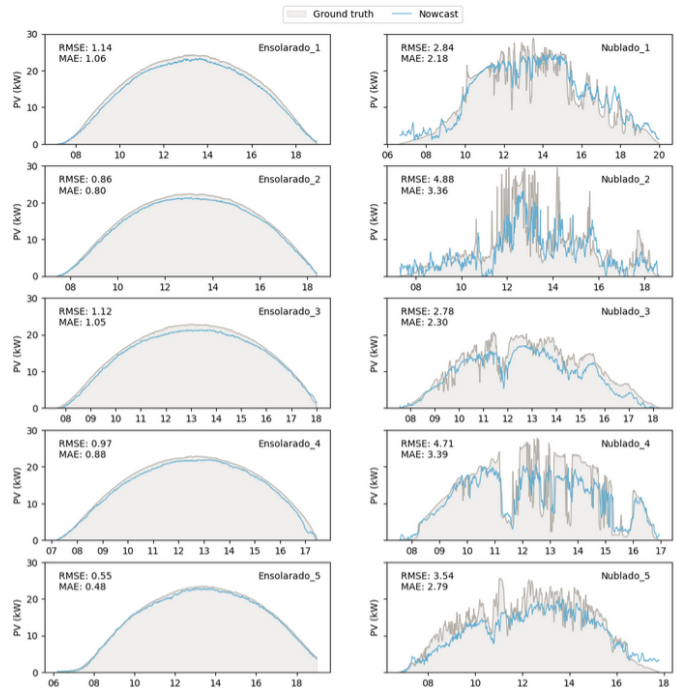


Figura 7: Desempenho do modelo em diferentes dias e condições climáticas.

Como é possível observar na 7, o modelo apresenta um desempenho excelente em dias ensolarados, capturando com sucesso o padrão parabólico e previsível da irradiância solar em condições de céu limpo. O RMSE para dias ensolarados foi de apenas 0.865, com um MAE de 0.718. Esses valores baixos indicam que, em média, o erro de previsão do modelo é

inferior a 1 kW nessas condições, o que é um resultado muito robusto.

Contudo, para dias nublados, o desempenho cai bastante. A geração real nesses dias é extremamente volátil, com picos e vales acentuados causados pela passagem de nuvens sobre o sol. Embora o modelo capture a tendência geral da curva diária (o nascer e o pôr do sol), falha em prever essas flutuações de alta frequência.

No geral, a predição do modelo tende a ser uma versão "suavizada" da realidade. Ele parece "hesitar" em prever os picos e vales extremos.

B. ResNet50V2

A técnica de fine-tuning tipicamente exige taxas de aprendizado conservadoras para não corromper os pesos pré-treinados. Visto que a melhor taxa de aprendizado identificada nos experimentos anteriores já era muito baixa ($1e-6$), esta foi mantida como ponto de partida para um teste inicial, juntamente com o batch size de 128. Para cada fold, o treinamento foi configurado para um máximo de 15 épocas e gerenciado por um conjunto de callbacks para otimização do processo. Foi empregado um mecanismo de EarlyStopping com uma paciência de 5 épocas para encerrar o treinamento caso a perda de validação não apresentasse melhora, bem como um callback ReduceLROnPlateau com paciência de 3 épocas para ajustar dinamicamente a taxa de aprendizado.

Tabela V: Desempenho por fold, avaliado no ponto de menor perda de validação, e a média final dos resultados.

Fold	RMSE Treino	RMSE Validação	Melhor Época
1	1.7500	2.0064	15
2	1.7745	2.0414	14
3	1.8762	1.8940	12
4	1.7377	2.1477	15
5	1.8940	1.7740	12
Média Final	1.8065	1.9727	–

Como é possível observar na Tabela V, os resultados foram bem promissores. Vale observar que em todos os folds o ReduceLROnPlateau foi acionado, ao passo que o EarlyStopping não foi acionado nenhuma vez. Juntando isso ao fato de a melhor época de todos os submodelos ter sido muito próxima da época final de treinamento, tudo indica que o modelo precisaria ser treinado por mais épocas. Além disso, como é possível observar na Figura 8, a perda de validação apresentou um comportamento mais ruidoso que o modelo SUNSET para os mesmos hiperparâmetros.

Considerando o elevado custo computacional do novo modelo, não foi realizada uma nova busca extensiva por hiperparâmetros. Em vez disso, foi adotada uma taxa de aprendizado menor de $3e-7$ e aumentado o limite de épocas por fold para 20, visando promover uma convergência mais estável e com menor variância.

Como é possível ver na Figura 9, o desempenho continua oscilando, porém parece ser mais estável. Contudo, na Tabela VI podemos ver que o desempenho foi, na verdade, ligeiramente pior que a primeira taxa.

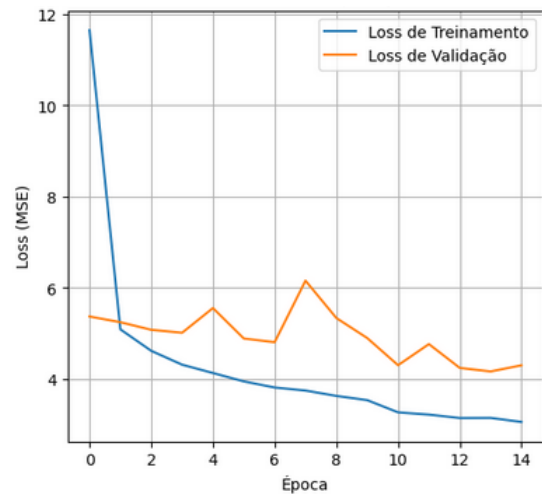


Figura 8: Curva de treinamento para o fine-tuning. Taxa: $1e-6$

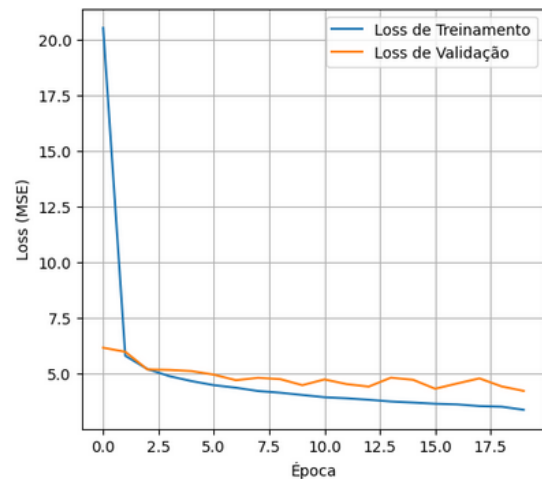


Figura 9: Curva de treinamento para o fine-tuning. Taxa: $3e-7$

1) *Resultados de Teste:* Apesar da instabilidade observada durante o treinamento com a taxa de aprendizado inicial de $1e-6$, os cinco submodelos gerados pela validação cruzada foram utilizados para realizar a predição no conjunto de teste. Notavelmente, mesmo sem fazer a devida otimização de hiperparâmetros, a média das predições de todos os submodelos atingiu um RMSE de 2.7182. Isso indica que, mesmo com

Tabela VI: Resumo do desempenho por fold e média final.

Fold	RMSE Treino	RMSE Validação	Melhor Época
1,00	1,8763	2,0442	20
2,00	1,8661	2,0737	20
3,00	1,9798	1,9383	13
4,00	1,8091	2,1808	20
5,00	1,9528	1,8101	17
Média Final	1,8968	2,0094	–

uma otimização ainda não ideal, a arquitetura ResNet já se aproxima do desempenho do baseline, sugerindo um alto potencial de melhoria.

Tabela VII: Resultado final no conjunto de teste para diferentes taxas de aprendizado.

Taxa de Aprendizado	RMSE de Teste
1e-06	2.7182
3e-07	2.7570

Infelizmente, o histórico de treinamento e submodelos da primeira combinação foram perdidos durante o desenvolvimento do projeto. Contudo, de forma a comparar o desempenho da ResNet com o SUNSET em diferentes condições climáticas, foram utilizados os submodelos gerados pelo treinamento com a taxa de aprendizado 3e-07 para gerar a seguinte tabela e imagens.

Tabela VIII: Desempenho final do modelo ResNet50V2 no conjunto de teste, estratificado por condição do céu.

Condição do Céu	RMSE	MAE
Dias Ensolarados	0.917	0.631
Dias Nublados	3.786	2.642
Geral	2.757	1.639

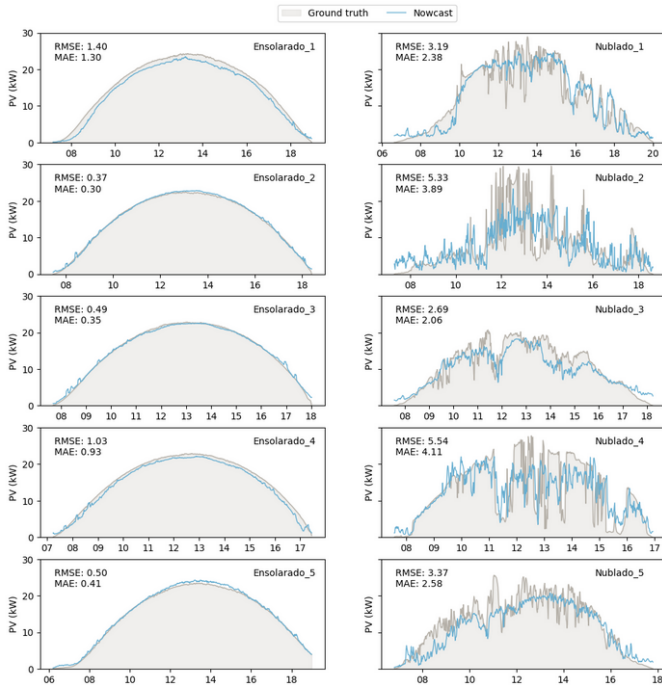


Figura 10: Desempenho da ResNet em diferentes condições climáticas.

IV. CONCLUSÃO

Com base nos resultados obtidos, este projeto demonstrou com sucesso a viabilidade da utilização de redes neurais convolucionais (CNNs) para a tarefa de "nowcasting" da geração

de energia solar fotovoltaica a partir de imagens do céu. A replicação dos resultados do modelo de referência, SUNSET, foi alcançada, obtendo-se um RMSE geral de 2.453 e um MAE de 1.541 no conjunto de teste. A análise detalhada revelou um desempenho excelente do modelo em dias ensolarados, com um RMSE de apenas 0.865. Contudo, em dias nublados, caracterizados pela alta volatilidade da geração, o desempenho foi inferior, com o modelo apresentando dificuldade em prever as flutuações de alta frequência, resultando em uma predição "suavizada" da realidade.

A investigação de uma arquitetura mais complexa, a ResNet50V2 com fine-tuning, também apresentou resultados promissores. Mesmo sem uma otimização extensiva de hiperparâmetros, devido ao seu maior custo computacional, a ResNet50V2 alcançou um desempenho próximo ao do SUNSET, com um RMSE geral de 2.757. Similarmente ao SUNSET, a ResNet50V2 exibiu um ótimo desempenho em dias de céu claro (RMSE de 0.917), mas enfrentou dificuldades em dias nublados (RMSE de 3.786). O desempenho próximo ao da baseline, mesmo com uma otimização incipiente, sugere um alto potencial de melhoria para a arquitetura ResNet.

O desenvolvimento do projeto também destacou a importância crítica da metodologia experimental. A implementação de uma pipeline de dados eficiente, com carregamento sob demanda, foi crucial para superar limitações de memória e mitigar o overfitting por meio de um embaralhamento completo do conjunto de dados a cada época. A utilização da validação cruzada Group K-Fold mostrou-se uma estratégia robusta para obter uma estimativa de desempenho justa em um dataset com desbalanceamento de condições climáticas.

Em suma, o estudo confirma que modelos de CNN, como o SUNSET e a ResNet50V2, podem extrair informações valiosas de imagens do céu para estimar a geração de energia solar instantânea. Enquanto o modelo SUNSET se mostrou eficaz e mais leve computacionalmente, a performance inicial da ResNet50V2 indica que arquiteturas mais profundas e pré-treinadas são uma via promissora para avanços futuros, especialmente se for realizada uma otimização mais aprofundada de seus hiperparâmetros. Futuros trabalhos poderiam focar nessa otimização e em técnicas para melhorar a previsão durante condições de nebulosidade variável.

REFERÊNCIAS

- [1] Y. Sun, G. Szűcs, and A. R. Brandt, "Solar pv output prediction from video streams using convolutional neural networks," *Energy Environ. Sci.*, vol. 11, pp. 1811–1818, 2018. [Online]. Available: <http://dx.doi.org/10.1039/C7EE03420B>
- [2] Y. Nie, X. Li, A. Scott, Y. Sun, V. Venugopal, and A. Brandt, "Skipp'd: A sky images and photovoltaic power generation dataset for short-term solar forecasting," *Solar Energy*, vol. 255, pp. 171–179, 2023.
- [3] Y. Sun, V. Venugopal, and A. Brandt, "Short-term solar power forecast with deep learning: Exploring optimal input and output configuration," *Solar Energy*, vol. 188, pp. 730–741, 08 2019.
- [4] Y. Nie, A. Scott, V. Venugopal, X. Li, Y. Sun, and A. Brandt, "SKIPP'D: A SKY Images and Photovoltaic Power Generation Dataset (Benchmark Data)," <https://purl.stanford.edu/dj417rh1007>, 2022, acessado em: 7 de julho de 2025.