

APLICAÇÃO DE VISÃO COMPUTACIONAL PARA IDENTIFICAÇÃO DE EQUIPES EM IMAGENS DE FÓRMULA 1

Sérgio Medeiros Fonte¹

Centro de Educação Aberta e à Distância (CEAD)

¹Universidade Federal do Piauí (UFPI) – Teresina, PI – Brazil

sergio.fonte@ufpi.edu.br

Resumo: *Este artigo apresenta o desenvolvimento de um sistema de classificação de imagens para identificar carros de Fórmula 1 por equipe e piloto. A base de dados foi construída manualmente por meio de webscraping e captura de imagens diversas, totalizando 2.019 registros. Utilizando técnicas de visão computacional e aprendizado de máquina, o modelo alcançou uma acurácia de aproximadamente 68%. O trabalho destaca os desafios de coleta, pré-processamento e treinamento com imagens reais e heterogêneas. Resultados e limitações são discutidos para futuras melhorias.*

Palavras-chave: *Visão computacional; Fórmula 1; Classificação de imagens; Aprendizado profundo; Dataset personalizado.*

1. Introdução

A Fórmula 1 é um dos esportes mais assistidos e tecnologicamente avançados do mundo, com equipes que investem milhões de dólares anualmente em inovação, estratégia e performance. O reconhecimento visual dos carros durante as corridas é um desafio tanto para espectadores quanto para profissionais envolvidos na cobertura, como narradores, fotógrafos e analistas. A semelhança entre os veículos, especialmente no que diz respeito a cores e design aerodinâmico, frequentemente causa confusões, dificultando a identificação rápida e precisa das equipes na pista. Apesar de a tecnologia televisiva atual empregar sistemas gráficos inspirados em videogames para exibir informações dos pilotos, esses dados ainda são alimentados manualmente, o que limita a escalabilidade e a precisão em tempo real.

Neste cenário, este trabalho propõe o desenvolvimento de um sistema automatizado de classificação de imagens de carros de Fórmula 1 por equipe, com o uso de técnicas de visão computacional e aprendizado profundo. A motivação do projeto

surge tanto do interesse pessoal pelo esporte quanto da constatação de uma demanda técnica real: a ausência de soluções automatizadas para identificação visual de veículos em registros fotográficos e audiovisuais. A proposta visa, além do entretenimento e da organização de acervos históricos, auxiliar profissionais da área — como fotógrafos esportivos — a identificar corretamente os carros e organizar seus portfólios de forma automatizada, otimizando o tempo e reduzindo erros.

O problema central da pesquisa reside na inexistência de bases públicas de dados para esse domínio e na dificuldade de se distinguir visualmente os carros de diferentes equipes, especialmente aqueles com esquemas de cores semelhantes, como Aston Martin, Mercedes e Williams. Diante disso, foi necessário construir uma base personalizada com 2.019 imagens, obtidas por meio de *webscraping*, capturas de tela de corridas e acervo pessoal. A partir dessa base, foi desenvolvido um classificador de imagens capaz de reconhecer a equipe de cada carro com acurácia aproximada de 68%, organizando automaticamente as imagens em pastas nomeadas conforme o resultado da predição.

A metodologia empregada envolve o uso de modelos de *deep learning*, técnicas de *data augmentation* e ferramentas como *Python*, *Keras* e *Google Colab* para treinamento e validação do classificador. Esta é uma pesquisa aplicada e empírica, com foco em resolver um problema prático, utilizando dados reais e estratégias modernas de ciência de dados.

No campo acadêmico, a relevância do projeto está na aplicação concreta de conceitos de visão computacional e inteligência artificial, demonstrando como tais técnicas podem ser aplicadas a nichos específicos com impacto direto em práticas profissionais e comerciais. Além disso, o trabalho contribui para a discussão sobre acessibilidade de dados esportivos e desenvolvimento de soluções personalizadas onde o mercado ainda não oferece alternativas eficazes.

O artigo está estruturado da seguinte forma: na seção seguinte, apresenta-se o referencial teórico com base nos conceitos de redes neurais convolucionais e classificação de imagens. Em seguida, detalha-se a metodologia, abordando as etapas de coleta, organização, modelagem e treinamento. A seção de resultados discute o desempenho do classificador, seus desafios e perspectivas de melhoria. Por fim, a conclusão resume os principais achados e propõe caminhos futuros para a evolução da proposta.

2. Referencial Teórico

O avanço da inteligência artificial (IA) e do aprendizado de máquina (*machine learning*) tem transformado significativamente a forma como sistemas computacionais interagem com dados visuais. No campo da visão computacional, o objetivo principal é permitir que as máquinas interpretem, analisem e compreendam imagens e vídeos, automatizando tarefas que, até então, dependiam exclusivamente da percepção humana (SZEWCZYK et al., 2019). Essa área é particularmente relevante em contextos como o esporte, a medicina, a indústria automotiva e a segurança, onde a identificação visual rápida e precisa é essencial.

Entre as abordagens mais eficazes para tarefas de reconhecimento de padrões visuais estão as redes neurais convolucionais (CNNs), uma classe especializada de redes neurais artificiais projetadas para processar dados com estrutura de grade, como imagens (LECUN; BENGIO; HINTON, 2015). As CNNs se destacam por sua capacidade de extrair automaticamente características relevantes das imagens, utilizando camadas convolucionais, funções de ativação, pooling e camadas densas, permitindo a construção de modelos altamente precisos e escaláveis.

A utilização de CNNs em sistemas de classificação de imagens já demonstrou resultados expressivos em competições como o ImageNet Large Scale Visual Recognition Challenge (ILSVRC), onde arquiteturas como *AlexNet*, *VGG*, *ResNet* e *EfficientNet* alcançaram desempenhos notáveis (KRIZHEVSKY et al., 2012; HE et al., 2016; TAN; LE, 2019). Tais modelos, quando combinados com técnicas de transfer learning, tornam-se ainda mais eficientes para aplicações específicas, permitindo que redes pré-treinadas em grandes bases de dados sejam adaptadas para domínios menores e personalizados, como o caso deste projeto.

Outro aspecto fundamental para o sucesso de sistemas de classificação é a qualidade e diversidade do conjunto de dados utilizado. Em contextos onde não existem bases públicas disponíveis, como é o caso da Fórmula 1, a construção de *datasets* personalizados se torna uma etapa crítica. A coleta de imagens reais, a padronização de formatos, o balanceamento entre classes e o uso de técnicas de data augmentation são estratégias fundamentais para mitigar o risco de overfitting e melhorar a generalização do modelo (SHORTEN; KHO, 2019).

Além disso, a literatura destaca que a aplicabilidade de sistemas de visão computacional transcende o reconhecimento de objetos, abrangendo também a organização automática de imagens, análise de desempenho, identificação de padrões visuais temporais e até apoio à tomada de decisão em ambientes complexos (GONZALEZ; WOODS, 2018). No caso da Fórmula 1, onde detalhes sutis como mudanças aerodinâmicas ou alterações nos patrocínios são comuns ao longo da temporada, um classificador robusto pode contribuir significativamente para a análise técnica e comercial das equipes envolvidas.

Portanto, o desenvolvimento de classificadores personalizados baseados em CNNs representa uma aplicação concreta e inovadora da inteligência artificial em nichos específicos, como o automobilismo de elite, e demonstra o potencial transformador da IA mesmo em contextos com acesso limitado a dados estruturados e recursos públicos.

3. Metodologia

Este trabalho utilizou uma abordagem empírica e aplicada para o desenvolvimento de um classificador de imagens capaz de identificar a equipe correspondente a um carro de Fórmula 1 com base em características visuais como cor, design, logotipos e aerodinâmica. A metodologia foi dividida em quatro grandes etapas: coleta e preparação dos dados, processamento e aumento dos dados (data augmentation), definição do modelo de aprendizado profundo e avaliação dos resultados obtidos.

3.1. Coleta e organização dos dados

Dada a inexistência de *datasets* públicos específicos para a Fórmula 1, foi necessário construir uma base de dados autoral, composta por 2.019 imagens. As imagens foram obtidas por meio de *webscraping* de sites de fabricantes de carros, acervos de fotógrafos profissionais, revistas especializadas e capturas de tela de corridas reais. Também foram utilizadas imagens pessoais do autor. As imagens foram organizadas em pastas nomeadas de acordo com a equipe correspondente, com uma separação entre conjuntos de treino e validação.

3.2. Pré-processamento e aumento dos dados

Com o intuito de melhorar a robustez do modelo e reduzir o risco de overfitting, foram aplicadas técnicas de data augmentation, como rotação, inversão horizontal, alteração de brilho e zoom. Esse processo ampliou artificialmente a diversidade do conjunto de dados, permitindo que o modelo aprendesse a reconhecer os carros sob diferentes ângulos e condições de iluminação. As imagens também foram redimensionadas para um tamanho padrão e normalizadas para que pudessem ser processadas corretamente pela rede neural.

3.3. Definição e treinamento do modelo

Foi adotado o uso de redes neurais convolucionais (CNNs) como modelo de aprendizado profundo para a classificação das imagens. O modelo foi construído em Python, utilizando as bibliotecas *TensorFlow*, *Keras*, *NumPy* e *Matplotlib*, entre outras. O desenvolvimento e treinamento foram realizados na plataforma Google Colab, aproveitando o suporte a GPUs para acelerar o processamento.

O pipeline do modelo envolveu:

1. Camadas convolucionais para extração de características.
2. Camadas de pooling para redução de dimensionalidade.
3. Camadas totalmente conectadas para classificação.
4. Função de ativação ReLU e softmax.
5. Otimizador Adam e função de perda categórica cruzada.

A divisão dos dados seguiu uma proporção aproximada de 80% para treino e 20% para validação, e o modelo foi treinado por múltiplas épocas até atingir um desempenho estável.

3.4. Avaliação e organização automática

O desempenho do modelo foi avaliado com base na acurácia obtida no conjunto de validação, que alcançou aproximadamente 68% de precisão. Após o treinamento, o modelo foi utilizado para organizar automaticamente novas imagens em pastas de acordo com a equipe classificada. A estrutura automatizada ainda requer validação humana em casos de baixa confiança, mas representa um avanço em direção à automação completa da triagem de imagens.

4. Resultados e Discussões

Após a construção de uma base de dados personalizada com 2.019 imagens distribuídas entre as dez equipes da temporada de Fórmula 1, foi possível treinar um classificador de imagens baseado em redes neurais convolucionais (CNNs). O modelo atingiu uma acurácia de validação de aproximadamente 68%, o que, embora não represente um desempenho perfeito, revela uma capacidade significativa de generalização, sobretudo considerando o desafio imposto pela semelhança visual entre os carros das diferentes equipes.

4.1. Desempenho do modelo

O modelo desenvolvido foi treinado ao longo de nove épocas, utilizando técnicas de data augmentation para aumentar a diversidade do conjunto de treinamento e reduzir o risco de overfitting. A Figura 1 apresenta as curvas de perda e acurácia para os conjuntos de treino e validação, revelando o comportamento do modelo ao longo do processo de aprendizado.

Nas primeiras épocas, observa-se uma evolução rápida da acurácia, tanto para os dados de treino quanto para validação. O modelo partiu de uma acurácia inicial inferior a 30%

no treino e 45% na validação, e em poucas iterações ultrapassou a marca dos 60% em ambos os conjuntos. Esse crescimento inicial evidencia que o modelo conseguiu capturar rapidamente padrões visuais básicos presentes nas imagens, beneficiando-se da aplicação das transformações de data augmentation, que forneceram exemplos variados de ângulos, luminosidades e rotações.

Conforme o treinamento avançou, as curvas começaram a se distanciar, especialmente a partir da quinta época, quando a acurácia no treino continuou subindo — chegando a 79% ao final da nona época — enquanto a acurácia de validação estagnou em torno de 68%. Esse descolamento entre as curvas é um indicativo claro de overfitting moderado, sugerindo que o modelo começou a memorizar características específicas do conjunto de treino, em vez de continuar generalizando bem para novas imagens. A estabilidade da acurácia de validação nas últimas épocas reforça esse comportamento, indicando que, apesar do ganho contínuo no desempenho sobre os dados vistos, não houve melhoria significativa sobre os dados não vistos.

A curva de perda oferece outro ponto de análise relevante. A perda no conjunto de treino caiu consistentemente até atingir 0.64, enquanto a perda de validação se estabilizou em 1.03, já na quarta época. Esse comportamento é típico em cenários onde o modelo apresenta boa capacidade de ajuste aos dados conhecidos, mas limitada capacidade de generalização, possivelmente por causa da complexidade visual dos dados e da quantidade relativamente pequena de imagens por classe.

Mesmo diante desse cenário, é importante destacar que o uso de técnicas como dropout e data augmentation ajudou a conter o agravamento do overfitting. A diferença entre as curvas não foi drástica, e o modelo foi capaz de manter uma acurácia consistente de validação em torno de 68%, o que representa um resultado expressivo diante da dificuldade da tarefa. A presença de classes com identidade visual muito semelhante — como Mercedes, Aston Martin e Williams — somada à diversidade de fontes de imagens, provavelmente impôs um limite natural à capacidade de generalização com o modelo atual.

De modo geral, o modelo demonstrou boa aprendizagem nos dados de treino, com sinais controlados de overfitting, e um nível de desempenho estável na validação, suficiente para justificar seu uso em aplicações práticas, como a organização automática

de acervos de imagens. Contudo, os resultados também apontam caminhos claros para evolução, como o uso de arquiteturas pré-treinadas, aumento do volume e da variedade do dataset, e estratégias de regularização mais refinadas.

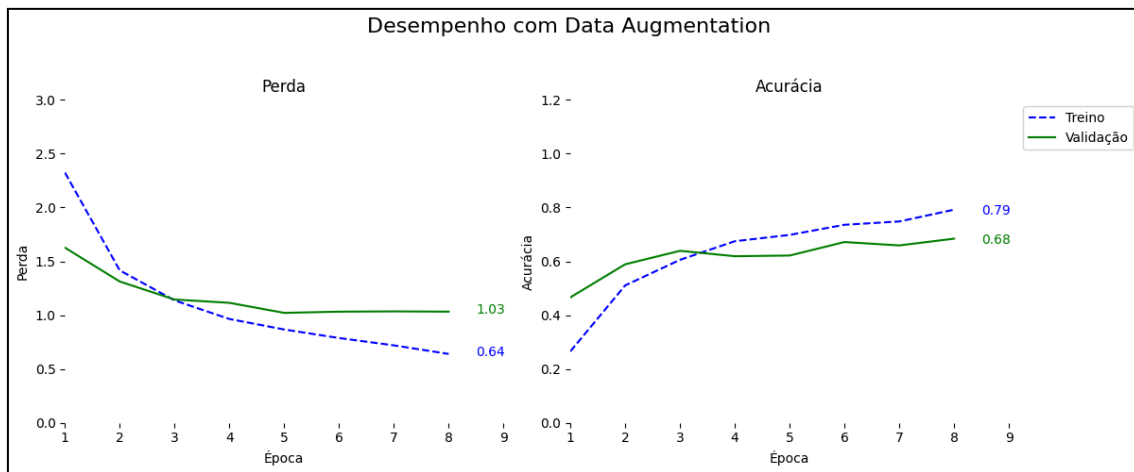


Figura 1 - Gráfico de acurácia e perda

4.2. Desafios enfrentados

O desenvolvimento do classificador enfrentou diversos desafios, sendo o principal a inexistência de bases públicas específicas para imagens de Fórmula 1 rotuladas por equipe. Isso exigiu a construção de um dataset manual com 2.019 imagens, obtidas a partir de múltiplas fontes e contextos visuais, o que gerou uma base com grande variação de qualidade, resolução, ângulo, iluminação e fidelidade visual. Essa heterogeneidade aumentou a complexidade da tarefa de classificação e contribuiu para ruídos no treinamento do modelo.

Além disso, a semelhança visual entre os carros de algumas equipes impactou diretamente o desempenho do classificador. A matriz de confusão (Figura 3) evidencia que Aston Martin, Mercedes e Williams foram frequentemente confundidas entre si, devido ao uso de esquemas cromáticos similares (tons escuros de verde, azul e preto), à presença de patrocinadores visuais semelhantes e ao design aerodinâmico quase idêntico.

A Mercedes, apesar de ter atingido uma precisão de 69%, apresentou recall de apenas 49%, com erros notáveis para Aston Martin (8 casos), Sauber (3) e Williams (2). A Aston Martin, embora tenha um recall de 72%, teve imagens confundidas principalmente com Alpine (12 casos) e Mercedes (2 casos). A Williams foi corretamente classificada em apenas 53% dos casos, com dispersão de erros principalmente para Aston Martin, Mercedes e Alpine.

	precision	recall	f1-score	support
alpine	0.53	0.48	0.51	48
aston_martin	0.51	0.72	0.60	50
ferrari	0.68	0.68	0.68	44
haas	0.48	0.47	0.47	30
mclaren	0.71	0.78	0.74	45
mercedes	0.69	0.49	0.57	41
racing_bulls	0.75	0.73	0.74	37
redbull	0.62	0.68	0.65	37
sauber	0.67	0.67	0.67	33
williams	0.69	0.53	0.60	34
accuracy			0.63	399
macro avg	0.64	0.62	0.62	399
weighted avg	0.63	0.63	0.62	399

Figura 2. Relatório por Equipe

Esses padrões da figura 2, indicam que o modelo consegue capturar características gerais das classes, mas tem dificuldade em identificar nuances visuais que diferenciam equipes com paletas de cores e posicionamento de logos muito semelhantes. Além das classes mencionadas, outras equipes como Haas (F1-score: 0.47) e Alpine (F1-score: 0.51) também apresentaram desempenho modesto. Os erros para Haas, por exemplo, ocorreram em diversas direções (Ferrari, Racing Bulls, Mercedes), sugerindo que a identidade visual da equipe talvez não esteja suficientemente distinta no dataset atual.

Apesar das limitações, o classificador apresentou desempenho robusto em algumas classes com maior identidade visual ou quantidade de imagens mais homogênea. É o caso da McLaren, com F1-score de 0.74 e recall de 78%, e da Racing Bulls, com F1-score também de 0.74, ambas com poucos erros de classificação cruzada.

Esses desafios reforçam a necessidade de estratégias complementares, como: Expansão do dataset com novas imagens sob diferentes condições climáticas, circuitos e ângulos; Rebalanceamento entre as classes, para mitigar viés de aprendizado; Transfer learning com redes pré-treinadas mais profundas, como EfficientNet ou ResNet; E, futuramente,

o uso de segmentação de logotipos ou detecção de regiões de interesse (ROI) para reforçar sinais visuais determinantes.

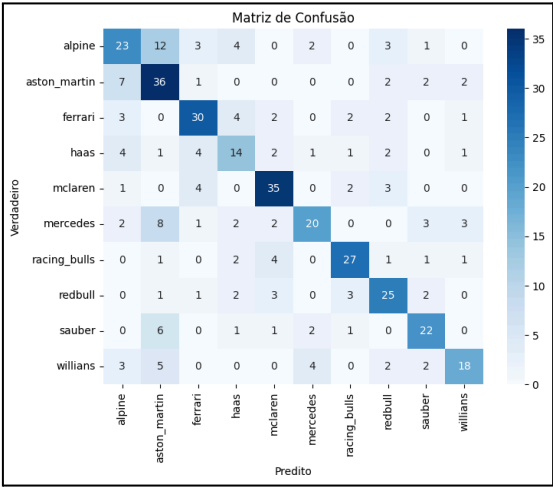


Figura 3 - Matriz de confusão

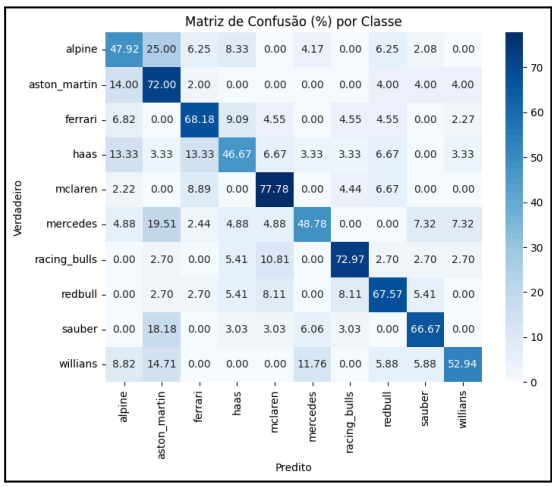


Figura 4. Matriz de confusão (%)

4.3. Aplicações práticas

O classificador desenvolvido foi integrado a um sistema de organização automática de imagens, no qual imagens novas são submetidas ao modelo e, com base na predição, movidas para a pasta correspondente à equipe identificada. Esse recurso pode ser útil tanto para fotógrafos e entusiastas do automobilismo quanto para fins históricos e estatísticos, promovendo o arquivamento automatizado por equipe.

Além disso, o modelo pode servir como base para futuras aplicações mais complexas, como o reconhecimento em vídeo em tempo real, por exemplo, durante transmissões ao vivo, esse sistema poderia identificar automaticamente os carros que passam por uma câmera onboard ou aérea, sem depender de sensores embarcado que podem se degradar ou exigir manutenção, além disso pode apoiar o suporte a análises táticas de corrida e até o monitoramento da evolução visual dos veículos ao longo da temporada, auxiliando equipes de engenharia e marketing.

4.4. Possibilidades de melhoria

Para melhorar os resultados obtidos, algumas estratégias podem ser exploradas podem ser o aumento do dataset, especialmente com imagens em diferentes condições

climáticas e de diferentes circuitos, também o uso de modelos pré-treinados (transfer learning) como ResNet, MobileNet ou EfficientNet, que podem aumentar a acurácia com menos dados, como alternativa também temos os ajustes nos hiperparâmetros, como taxa de aprendizado, número de camadas e funções de ativação e por fim a implementação de um sistema de feedback humano, para corrigir classificações erradas e gerar dados rotulados com mais precisão.

5. Conclusão

Este trabalho apresentou o desenvolvimento de um sistema de classificação de imagens voltado para a identificação de equipes da Fórmula 1 a partir de fotos reais dos veículos. A proposta surgiu da combinação entre uma motivação pessoal — o apreço pela categoria — e a observação de uma lacuna prática: a dificuldade de reconhecimento preciso e automatizado dos carros durante transmissões, acervos visuais e análises técnicas.

Diante da ausência de bases de dados públicas e estruturadas, foi necessário construir um *dataset* autoral com mais de 2.000 imagens obtidas por meio de *web scraping*, capturas de tela e registros pessoais. Essa base serviu de insumo para o treinamento de um modelo de *deep learning* baseado em redes neurais convolucionais (CNN), com uso de técnicas como *data augmentation* e validação cruzada. O classificador final alcançou cerca de 68% de acurácia, valor considerado promissor dado o alto grau de similaridade visual entre carros de diferentes equipes.

Além da criação do modelo em si, foi implementado um mecanismo de organização automática de imagens, em que novas imagens são classificadas e movidas para pastas correspondentes à equipe prevista. Tal funcionalidade pode ter aplicações reais para fotógrafos, jornalistas, equipes técnicas e plataformas de mídia esportiva, promovendo ganhos em agilidade e precisão na curadoria de conteúdo.

Entre as principais dificuldades enfrentadas, destacam-se a obtenção de imagens rotuladas, as variações de iluminação e ângulo, e a semelhança estética entre

determinadas equipes. Esses fatores impactam diretamente a performance do modelo e demonstram a importância de continuar expandindo e refinando a base de dados.

Como perspectivas futuras, sugere-se o uso de arquiteturas mais avançadas via transfer learning a introdução de classificação hierárquica (primeiro por equipe, depois por piloto), a adaptação do modelo para reconhecimento em vídeo ou em tempo real; a integração com sistemas de arquivamento digital, para fins históricos e analíticos. Em síntese, este projeto não apenas demonstrou a viabilidade técnica de um classificador visual de Fórmula 1, mas também apontou caminhos para soluções mais completas e aplicáveis em contextos reais do esporte e da tecnologia.

6. Referências

GONZALEZ, R. C.; WOODS, R. E. **Digital Image Processing**. 4. ed. Pearson, 2018.

LECUN, Y.; BENGIO, Y.; HINTON, G. **Deep learning**. *Nature*, v. 521, p. 436–444, 2015.

HE, K. et al. **Deep Residual Learning for Image Recognition**. In: CVPR, 2016.

KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. **ImageNet classification with deep convolutional neural networks**. *Advances in Neural Information Processing Systems*, 2012.

SHORTEN, C.; KHO, Z. **A survey on image data augmentation for deep learning**. *Journal of Big Data*, 2019.

SZEWCZYK, M. M. et al. **Computer Vision in Sports: A Review**. *Machine Vision and Applications*, 2019.

TAN, M.; LE, Q. **EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks**. In: ICML, 2019.