

Nome: Pedro Henrique Rodrigues Marques Dos Santos.

Matrícula: 11611ECP017.

Matéria: Machine Learning.

Docente: keiji Yamanaka.

Melhorando o Realismo de Imagens Sintéticas

(por Apple Computer, Inc.)

As redes neurais mais bem-sucedidas são treinadas com supervisão, porém para tal treinamento é necessário um amplo conjunto de treinamentos. Uma alternativa é usar imagens sintéticas a partir de um simulador, mas as imagens sintéticas podem não ser suficientemente realistas, resultando em imagens fracas de teste reais. Então foi desenvolvido um método para refinar imagens sintéticas para que se pareçam mais realistas. Os modelos de treinamento nessas imagens refinadas levam a melhorias significativas na precisão em várias tarefas de aprendizagem de máquinas.

A formação de modelos de aprendizagem de máquinas em imagens sintéticas padrão é problemática, pois as imagens podem não ser suficientemente realistas, levando o modelo a aprender detalhes presentes apenas em imagens sintéticas e não generalizando bem em imagens reais. Para este fim, foi desenvolvido um método para refinar imagens sintéticas para torná-las mais realistas. A tarefa é aprender um modelo que melhore o realismo de imagens sintéticas a partir de um simulador usando dados reais não marcados, preservando as informações de anotação.

Com objetivo de melhorar o realismo foi criada uma "Rede de refinadores", que processa imagens sintéticas. Para aprender uma rede de refinadores, precisamos de algumas imagens reais. Uma opção seria exigir pares de imagens reais e sintéticas com correspondência em pixels, ou imagens reais. Para criar uma correspondência em pixel, é preciso renderizar uma imagem sintética que corresponde a uma determinada imagem real ou capturar uma imagem real que corresponda a uma imagem sintética renderizada. Então aprenderam um mapeamento sem correspondência em pixels ou qualquer rótulo para as imagens reais. Para a rede de refinadores aprenda de forma não supervisionada, utilizou-se uma rede auxiliar de

discriminação que classifica as imagens reais e refinadas em duas classes. A rede de refinadores tenta enganar essa rede discriminadora para pensar que as imagens refinadas são as reais. As duas redes se alternam, e o treinamento para quando o discriminador não consegue distinguir as imagens reais das falsas.

Além de gerar imagens realistas, a rede deve preservar as informações de anotação do simulador como a estimativa do olhar. A transformação aprendida não deve mudar a direção do olhar. Para preservar as anotações de imagens sintéticas, complementamos a perda contraditória com uma perda de auto regularização L1 que penaliza grandes mudanças entre as imagens sintéticas e refinadas. Outro requisito importante para a rede de refinadores é que ele deve aprender a modelar as características reais da imagem sem apresentar nenhum artefato.

Treinaram a rede de refinadores com apenas perda de auto regularização e apresentamos a perda adversa depois que a rede de refinadores começa a produzir versões embaçadas das imagens sintéticas de entrada. No início, ele produz uma imagem embaçada que se torna cada vez mais realista à medida que o treinamento avança. A perda do discriminador é baixa no começo - o que significa que pode facilmente dizer a diferença entre real e refinado. Lentamente, a perda do discriminador aumenta e a perda do gerador diminui à medida que o treinamento avança, gerando mais imagens reais.