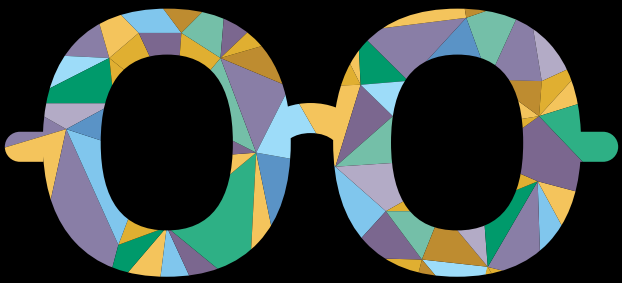


OBSERVATÓR!O2016

Experiências em *Deep Learning*

UM PROJETO DO VISGRAF LAB

Para mais informações sobre o projeto visite nossas páginas:

OBSERVATÓR!O2016: oo.impa.br

***Deep Learning* e Rio-2016:** lvelho.impa.br/dl_rio2016/index.html

Documentação do projeto: visgraf.impa.br/rio2016/project.html

SOBRE AS EXPERIÊNCIAS

As experiências apresentadas neste folheto correspondem a desdobramentos do OBSERVATÓR!O2016, um projeto de monitoramento e visualização dos Jogos Olímpicos Rio 2016 a partir de dados de redes sociais. A concepção e o desenvolvimento do trabalho foram realizados por uma equipe interdisciplinar que atua no VISGRAF, o Laboratório de Visão e Computação Gráfica do IMPA.

Em decorrência do projeto, foram coletadas aproximadamente 180 mil imagens únicas sobre a Rio-2016 compartilhadas no Twitter antes e durante os Jogos. No atual momento, inauguramos uma nova fase de pesquisa na qual a equipe está interessada em explorar formas automatizadas de reagrupar as imagens e remixá-las em novos produtos audiovisuais a partir de princípios e metodologias de *Deep Learning*, uma abordagem que vem ganhando destaque dentro de pesquisas empreendidas no campo da visão computacional.

Ao apresentar os resultados audiovisuais das imagens da Rio-2016 queremos, dentre outros objetivos, problematizar a relação entre tecnologia e arte, mais precisamente no que diz respeito a aspectos de autoria e criatividade.

MOSAICO DA TOCHA

Coleção de imagens do *tour da tocha* visualizadas na forma de um mosaico interativo

Essa experiência recorre à abordagem de *Deep Learning* - mais especificamente à prática de treinamento supervisionado utilizando redes neurais convolucionais - para reconhecer visualmente a presença da tocha Olímpica nas imagens do nosso banco de dados. Em virtude dos esforços da Google, o problema de reconhecimento de objetos em imagens tornou-se relativamente mais acessível e replicável. Isso porque a empresa liberou, no último ano, o uso *opensource* da rede neural *Inception-v3* - capaz de classificar imagens a partir de 1000 categorias fornecidas pela *ImageNet* - e da biblioteca de códigos *TensorFlow* - que auxilia o desenvolvimento de um sistema de classificação de imagem.

Além disso, a rede *Inception-v3* pode ser retreinada para detectar uma nova classe de imagem. Para retreinar a rede a fim de reconhecer a tocha Olímpica nas imagens, precisamos, além de uma infraestrutura computacional adequada para otimizar o processo, de uma amostra de cem imagens da tocha Olímpica manualmente rotuladas.

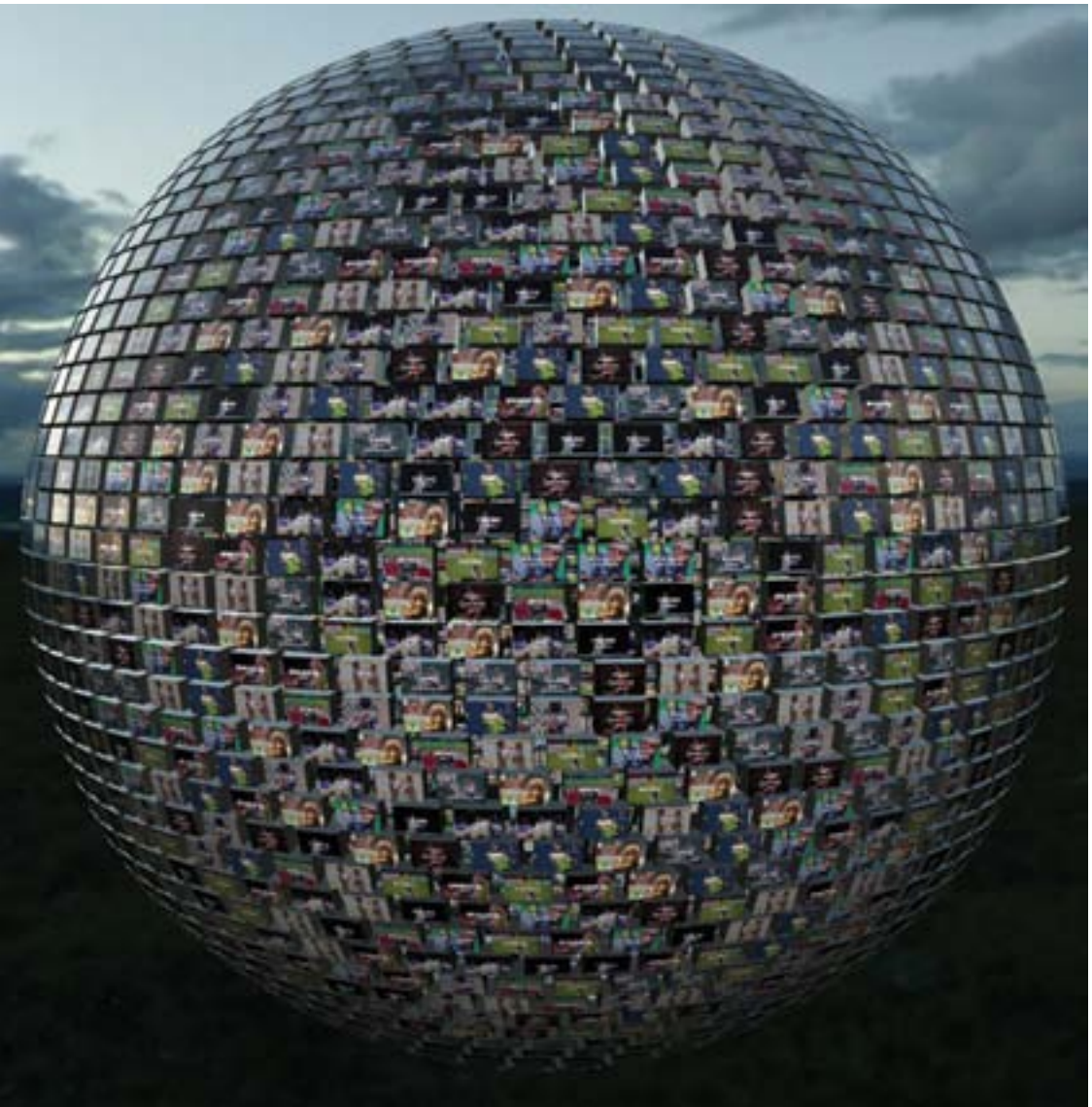
Acima, detalhe do mosaico. À esquerda, 12 das 2091 imagens classificadas na categoria tocha Olímpica com mais de 83% de probabilidade.

Após ser retreinada, a rede neural reconheceu a presença da tocha Olímpica nas imagens do banco de dados segundo uma lógica de probabilidade. Até o dia 25 de junho, 1500 imagens com mais de 85% probabilidade de conter a tocha Olímpica tinham sido classificadas pela rede. A partir delas, e aplicando operações de espelhamento e rotação, criamos o mosaico de imagens que é apresentado no site.

Veja mais em **http://lvelho.impa.br/dl_rio2016/mosaico.html**

GLOBO DE VÍDEOS

Montagens audiovisuais reunem imagens de esportes olímpicos e atletas



Assim como fizemos para a categoria de imagens tocha Olímpica, retreinamos a rede neural *Inception-v3* para reconhecer distintos esportes e a premiação de medalhas. No entanto, propomos uma nova solução de síntese para as imagens. Dessa vez, exploramos formas automatizadas de remixá-las imagens em resultados audiovisuais, mais precisamente *slideshows* musicais que apresentamos no formato de um globo de vídeos em nosso site. Para a síntese, recorremos a recursos automatizados, oferecidos pela Apple, para produção de *slideshows*.

O novo sistema operacional da Apple (iOS 10) traz uma forte presença de inteligência artificial em seus aplicativos nativos. O aplicativo *Photos* utiliza *Deep Learning* para reconhecer rostos, objetos e cenas nas fotografias e vídeos armazenados no album do usuário. *Photos* oferece a recente funcionalidade *Memories* que cria automaticamente *slideshows* musicais com base no reconhecimento de pessoas, lugares, cenas e datas. Depois de pronto, o aplicativo permite personalizar algumas propriedades do *slideshow* através de seu ambiente de autoria.

Para que *Photos* criasse automaticamente *slideshows* musicais com as imagens da Rio-2016, importamos, progressivamente, coleções de images no aplicativo com base nas classificações previamente feitas pela rede neural. A sequência de telas nessa página apresenta parte do processo de criação dos *slideshows* através do aplicativo *Photos* no *iPad*. Conheça mais detalhes sobre a metodologia desse experimento em nosso site.

Veja mais em **http://lvelho.impa.br/dl_rio2016/videos.php**

IMPRESSÕES SONORAS

Declarações de atletas e condutores da tocha combinados com áudios que marcaram a Rio-2016

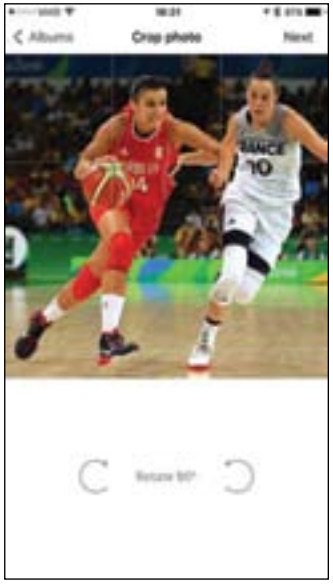
Essa experiência surgiu do desejo de explorar uma mídia até então pouco trabalhada no projeto: o áudio. Durante o revezamento da tocha Olímpica e a realização dos Jogos, coletamos registros sonoros que marcaram o megaevento - declarações de atletas e condutores, reportagens, locuções de rádio, canções, paródias que viralizaram nas redes, etc. Esses registros sonoros foram extraídos de vídeos e armazenados em um conta no *SoundCloud*.

Através de uma curadoria, combinamos alguns desses áudios com textos de declarações públicas de atletas e condutores da tocha. A mescla de sons e textos foi realizada, no nível da interface, através do uso do *plugin SoundCite*.

Embora a hibridização entre sons e textos não empregue *Deep Learning*, as imagens que compõem o fundo da página do site foram sintetizadas a partir de um modelo de aprendizagem. Selecionamos uma única fotografia da base de dados do OBSERVATÓR!O2016 e a submetemos diversas vezes ao aplicativo de edição de imagem *Prisma*, disponível para os sistemas iOS e Android. Esse aplicativo utiliza redes neurais profundas para resolver um problema específico da visão computacional: a transferência do estilo artístico de uma imagem para o conteúdo de outra imagem.

Prisma foi lançado em junho de 2016 por Alexey Moiseenkov e seu desenvolvimento tem o artigo *A Neural Algorithm of Artistic Style* (2015) como principal base teórica.

As imagens abaixo são reproduções de tela do aplicativo *Prisma* no *iPhone*:



Da esquerda para a direita, a imagem original importada no app. Ao lado, o resultado após a imagem ser fundida com o estilo Roy (referência ao artista Roy Lichtenstein).

Veja mais em **http://lvelho.impa.br/dl_rio2016/som.html**

EXIBIÇÃO ARTÍSTICA

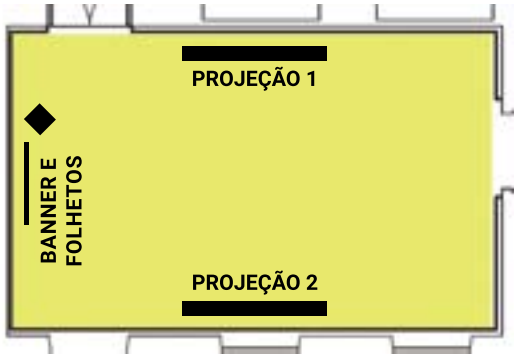
Proposta de trabalho para o encontro *Indisciplinas: a arte frente ao urgente*

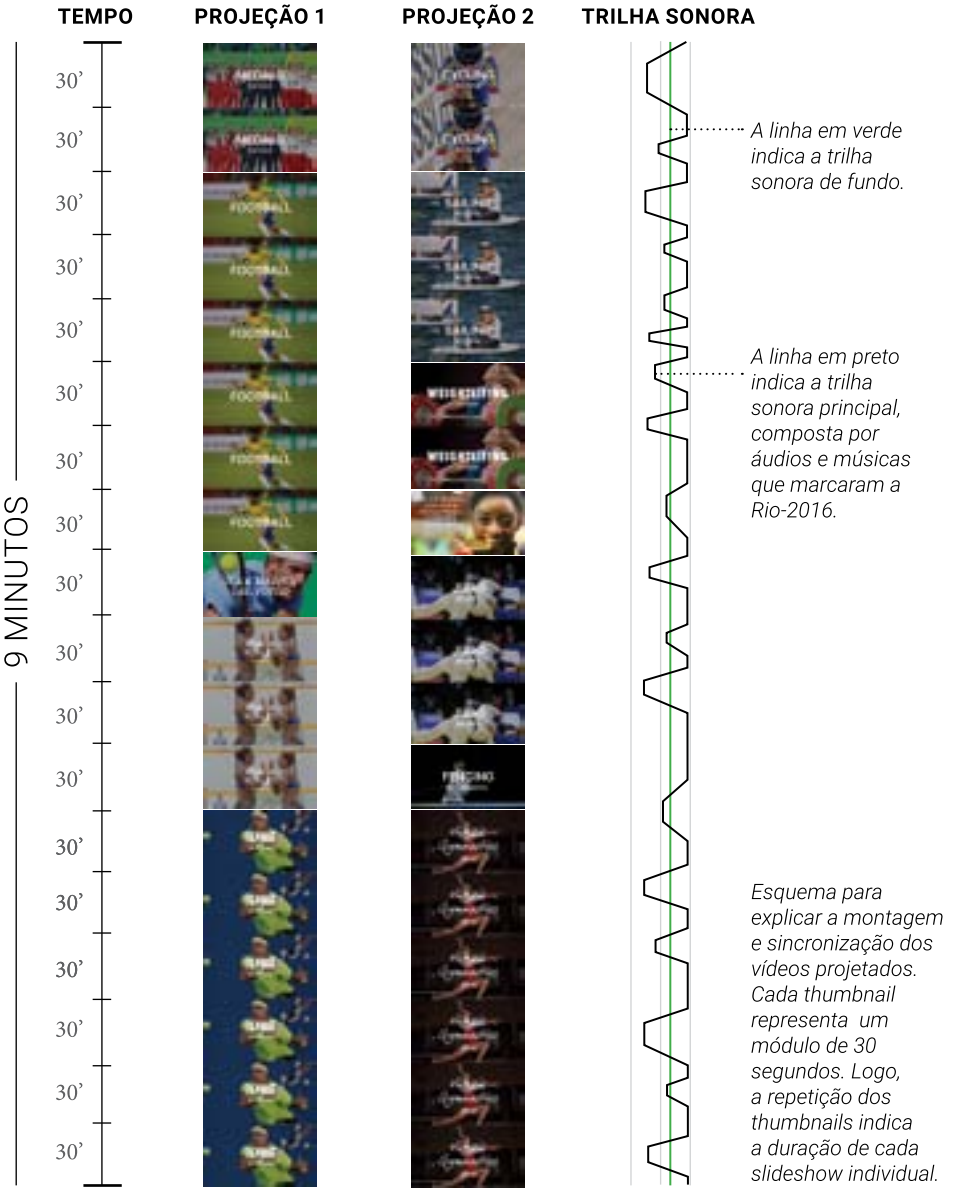
Em razão da convocatória de trabalhos para o *Indisciplinas* - 4º encontro dos Programas de Pós-Graduações em Artes Visuais do Estado do Rio de Janeiro -, subtemos as experiências envolvendo *Deep Learning* e as imagens da Rio-2016 para serem apresentadas no formato de uma exibição artística. A proposta foi aceita e durante os dias 22 e 25 de novembro exporemos vídeo-projeções que mesclam as três experiências antes descritas. Como resultado, nosso trabalho busca problematizar questões já discutidas na área de Artes: podemos ensinar a máquina a ser criativa? De quem é a autoria do trabalho: do artista, do computador, de ambos?

Para a exibição artística, reunimos os *slideshows* musicais que compõem o experimento *Globo de Vídeos* em dois únicos vídeos para serem projetados simultaneamente. Os vídeos receberam uma nova trilha sonora, editada a partir dos áudios e músicas do experimento *Impressões Sonoras*.

SALA DE PROJETORES
56,53m²

Possível disposição das projeções a partir da planta baixa da Casa França Brasil disponibilizada no edital de convocatória.





Veja mais em **http://lvelho.impa.br/dl_rio2016/indisciplinas.html**

DEEP LEARNING E ARTE

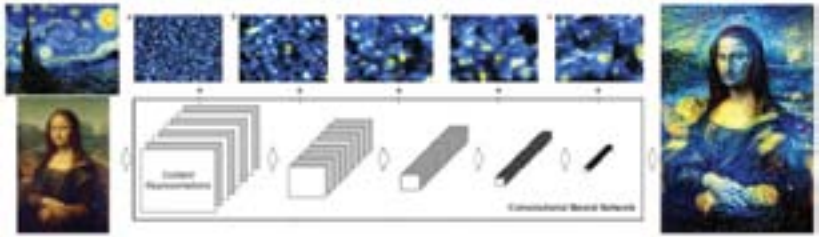
Projetos e discussões emergentes sobre a relação entre arte e *Deep Learning*

INTERSEÇÃO ENTRE ARTE E INTELIGÊNCIA ARTIFICIAL

O grupo de pesquisa da Google faz um histórico e comenta alguns trabalhos recentes gerados da convergência entre Arte e *Machine Learning*.

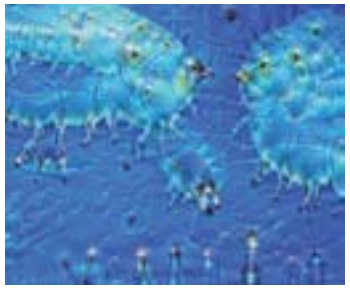
REDES NEURAIS COM TALENTO ARTÍSTICO

O texto de Christopher Burger explica como programas computacionais são capazes de gerar automaticamente imagens artísticas com alta qualidade perceptual.



MACHINE LEARNING PARA ARTISTAS

Gene Kogan destaca pontos interessantes do seu curso *Machine Learning* para artistas, oferecido na Universidade de Nova Iorque (NYU).



REDES NEURAIS PROFUNDAS SÃO CRIATIVAS?

O texto de Zachary Chase Lipton discute se os *outputs* de redes neurais profundas - mais precisamente redes adversariais generativas - podem ser considerados resultados criativos.

MAGENTA

Magenta é um braço de pesquisa da Google voltado para lançar perguntas e buscar respostas para problemas como: “podemos usar *Deep Learning* para criar produtos artísticos e musicais atraentes? Se sim, como? Se não, por quê?”.

Veja mais em **http://lvelho.impa.br/dl_rio2016/dlearte.html**

CONCEPÇÃO

Somos um time interdisciplinar que desenvolve projetos e pesquisas no VISGRAF

JULIA GIANNELLA
Assistente de pesquisa no VISGRAF e doutoranda no PPDESDI-UERJ
autoria e design

LUIZ VELHO
Professor titular do IMPA e líder do VISGRAF
autoria e supervisão do projeto

JULIANO KESTENBERG
Assistente de pesquisa no VISGRAF e doutorando no PPGAV-UFRJ
edição de vídeo

VITOR GUERRA
Pós-Doc no VISGRAF
edição de som

DJALMA LUCIO
Desenvolvedor de sistemas no IMPA
programação e suporte de sistemas

O VISGRAF é o Laboratório de Visão e Computação Gráfica do IMPA (Instituto Nacional de Matemática Pura a Aplicada)
Conheça nossa página: **visgraf.impa.br**

VISGRAF LAB
visgraf.impa.br