UNIVERSIDADE FEDERAL DE PELOTAS
Centro de Desenvolvimento Tecnológico
Programa de Pós-Graduação em Computação

Dissertação

Uma abordagem de extração de grafite com multiagente e identificação por CNNs

Glauco Roberto Munsberg dos Santos

Pelotas, 2019

Glauco Roberto Munsberg dos Santos

Uma abordagem de extração de grafite com multiagente e identificação por CNNs

Dissertação apresentada ao Programa de Pós-Graduação em Computação da Universidade Federal de Pelotas, como requisito parcial à obtenção do título de Mestre em Ciência da Computação

Orientador: Prof. Dr. Ricardo Matsumura Araújo

Pelotas, 2019

Universidade Federal de Pelotas / Sistema de Bibliotecas
Catalogação na Publicação

S237a Santos, Glauco Roberto Munsberg dos

Uma abordagem de extração de grafite com multiagente e identificação por CNNs / Glauco Roberto Munsberg dos Santos ; Ricardo Matsumura Araújo, orientador. — Pelotas, 2019.

66 f.

Dissertação (Mestrado) — Programa de Pós-Graduação em Computação, Centro de Desenvolvimento Tecnológico, Universidade Federal de Pelotas, 2019.

1. Grafite. 2. Arte urbana. 3. Rede neural convolucional. 4. Aprendizado profundo. 5. Sistemas multiagente. I. Araújo, Ricardo Matsumura, orient. II. Título.

CDD : 005

Elaborada por Maria Inez Figueiredo Figas Machado CRB: 10/1612

Dedico este trabalho a comunidade do grafite que, mesmo recebendo duras críticas e olhares, luta para o que o fenômeno seja compreendido como forma de arte e comunicação.

AGRADECIMENTOS

Agradeço ao Prof. Ricardo Matsumura por sempre me desafiar e suas inúmeras contribuições no meu crescimento, suas orientações foram fundamentais para a minha formação pessoal e profissional. Obrigado por acreditar em mim e no meu trabalho.

Ao programa *Google Research Awards for Latin America* pelo financiamento e reconhecimento da importância que este trabalho tem para a comunidade.

Agradeço aos meus pais, minha irmã, ao meu sobrinho Rajeh e amigos por perceberem que este trabalho exigiu de mim grande esforço para melhorar a forma como me comunico. Obrigado pelas palavras e gestos de encorajamento que permearam toda a trajetória deste trabalho.

Agradeço mais uma vez a minha *alma mater* UFPel por me permitir desbravar mais esta área de conhecimento, ao PPG da Computação por me ensinar a técnica e a curso de antropologia, através da figura da Prof. Cláudia Turra, por me ensinar a arte do grafite.

Não pergunte quem eu sou e não me peça para seguir sendo o mesmo: deixe que os nossos burocratas e a nossa polícia vejam que os nossos documentos estão em ordem. Pelo menos, evitamos a sua moralidade quando ~~escrevemos~~ pichamos.

— MICHEL FOUCAULT

RESUMO

SANTOS, Glauco Roberto Munsberg dos. **Uma abordagem de extração de grafite com multiagente e identificação por CNNs**. 2019. 66 f. Dissertação (Mestrado em Ciência da Computação) – Programa de Pós-Graduação em Computação, Centro de Desenvolvimento Tecnológico, Universidade Federal de Pelotas, Pelotas, 2019.

O grafite é um intervenção urbana que utiliza muros, paredes e postes como suporte e geralmente está ligado a uma mensagem social ou política, estas representações urbanas são, muitas vezes, um importante indicador social. Mapeá-los e rastreá-los permite compreender como essas intervenções interagem com os demais elementos do meio urbano. Este trabalho tem por objetivo avaliar o uso de Redes Neurais Convolucionais e Sistemas Multiagente para localizar e mapear grafites em cidades, a partir de imagens em nível de rua provenientes do Google Street View. O método utilizado foi a elaboração de quatro experimentos com as redes neurais pré-treinadas e reuso dos seus classificadores para o novo contexto de identificação de grafite. Utilizamos para isso a técnica de *fine-tuning* com imagens extraídas do Flickr e do Google Street View. Através da análise dos modelos será mostrado que o reuso dos classificadores é promissor, diminuindo o tempo de treinamento das redes e obtendo modelos com resultados de 76,9% para a taxa de verdadeiros positivos quando testado o *dataset* do Flickr e sensibilidade de 71,43% em imagens do ambiente urbano. Destaca-se ainda neste trabalho o sistema multiagente capaz de percorrer o ambiente urbano do Google Street View e analisar em média 61 imagens por minuto para cada agente.

Palavras-Chave: grafite; arte urbana; rede neural convolucional; aprendizado profundo; sistemas multiagente

ABSTRACT

SANTOS, Glauco Roberto Munsberg dos. **A graffiti extraction approach with multi-agent and identification CNNs**. 2019. 66 f. Dissertação (Mestrado em Ciência da Computação) – Programa de Pós-Graduação em Computação, Centro de Desenvolvimento Tecnológico, Universidade Federal de Pelotas, Pelotas, 2019.

Graffiti is an urban intervention that uses walls and posts as support generally linked to a social or political message. These urban representations are often an important social indicator. Mapping and tracking them allows us to understand how these interventions interact with other elements of the urban environment.

This work aims to evaluate the use of Convolutional Neural Networks and Multiagent Systems to locate and map graffiti in cities from street level images from Google Street View. The method used was the elaboration of four experiments with the pretrained neural networks and reuse of their classifiers for the new context of identification of Graffiti. We used the fine-tuning technique with images extracted from Flickr and Google Street View. Through the analysis of models, it will be shown that the reutilization of the classifiers is promising, reducing the network training time and getting models with results of 76.9% for the true positive rate on Flickr's dataset and sensitivity 71.43% images in the urban environment. Also, it is worth noting that the multi-agent system can navigate the urban environment of Google Street View and analyze an average of 61 images per minute for each agent.

Keywords: graffiti; urban art; convolutional neural networks; deep learning; multiagents system

LISTA DE FIGURAS

LISTA DE TABELAS

LISTA DE ABREVIATURAS E SIGLAS

CNN	Redes Neurais convolucionais
MAS	Sistemas multiagente
GSV	Google Street View
GWR	Geographically weighted Regression
SPMD	Single Program Multiple Data
SIFT	Scale Invariant Feature Transform
ML	Machine Learning
MLP	Multilayer Perceptron
FC	Fully Connected
GSV	Google Street View
OSM	Open Street Map

SUMÁRIO

1 INTRODUÇÃO

O grafite é um fenômeno urbano que utiliza muros, paredes e postes como suporte a sua manifestação. Sempre visando passar uma mensagem, que nem sempre está direcionada a quem os observa no dia-a-dia, mas as autoridades públicas ou até mesmo a outros grupos de grafite. Os grafiteiros usam letras, desenhos, muitas vezes gramática e referências próprias a seu meio com o objetivo de expressar sua ideia (CAMPOS, 2007, p. 268).

Muitas vezes o grafite é confundido com pichação, porém há distinções entre ambos: A técnica de criação do grafite exige uma maior organização na sua elaboração e execução, apresenta uma diversidade de cores maior e objetos complexos (Figura 1). Já as pichações prevalecem características como o tracejado simples, letras ou marcações (*tags*) monocromáticas e quase sempre efêmeras (LASSALA, 2010, p. 46)(CAMPOS, 2010). Porém tanto para a expressão do grafite como a pichação são utilizados como materiais o *spray* de tinta, tinta convencional com rolinho, canetões e até mesmo de matrizes para a sua rápida replicação nos suportes.

Dado que o grafite, em geral, está ligado a uma mensagem social ou política, estas representações urbanas são, muitas vezes, um importante indicador social. Mapeá-los e rastreá-los permite compreender como estas intervenções interagem com os demais elementos do meio urbano. Com isso uma série de trabalhos no campo da sociologia, antropologia e arquitetura já tiveram como objeto de estudo o mapeamento (BÁRBARA HYPOLITO, 2015) e classificação dos grafites (CAMPOS, 2007, 2010) no ambiente urbano através de técnicas manuais de extração, classificação e armazenamento através de GIS[1]. Na Ciência da Computação também já houve abordagens para a extração (YANG et al., 2012; HANNUN et al., 2014) tendo o grafite como parte de indicadores sociodemográficos (PARRA; BOUTIN; DELP, 2012; FERRELL, 1993; PARRA et al., 2013), mapeamento (JAIN; LEE; JIN, 2009) e recuperação de pichações por similaridade de assinatura (*tags*) em pichações (LECUN et al., 1998; TONG et al., 2011).

[1]GIS acrônimo para *Geographic information system* são sistemas para armazenar, manipular e visualizar informações geográficas.

Figura 1 – A imagem contém o grafite dos artistas Os Gêmeos com pichações ao redor. Fonte: (FEIXA; OLIART, 2016)

Apesar da aparente facilidade na identificação e classificação de grafite por humanos, muitos autores divergem sobre o que se considera grafite e pichação (LASSALA, 2010) e até mesmo sobre a sua classificação (BÁRBARA HYPOLITO, 2015; CAMPOS, 2007). Para a computação visual a tarefa de identificação também não é fácil, trabalhos recentes neste campo, ainda carecem da intervenção humana na coleta e classificação (PARRA; BOUTIN; DELP, 2012, p. 7) (JAIN; LEE; JIN, 2009, p. 3) visto a quantidade de ruídos que podem ocorrer na detecção (Figura 2) do grafite pelos modelos.

Na área de computação visual – área da computação responsável por permitir aos computadores a ver através de imagens e vídeos – as técnicas de aprendizado de máquina vêm contribuindo para a identificação de objetos em imagens. Hoje o estado da arte para a identificação em imagens é através da técnica das redes neurais convolucionais (Convolutional Neural Networks - CNNs), estas redes têm arquiteturas muito mais profundas que suas antecessoras e são o resultado de novas técnicas e o avanço do poder computacional destes últimos anos. A técnica das CNNs será abordada com mais detalhe na seção 2.2 no Capítulo 2 de Referencial Teórico.

Recentes aplicações de CNNs, para análise complexa de imagens, vem mostrando o quão promissor é a técnica para aplicar em problemas de reconhecimento e classificação de imagens. Em SONG; XIAO (2016) o autor introduz uma arquitetura de CNN capaz da estimar o volume tridimensional dos objetos presentes em uma imagem. Com a análise de uma imagem RGB, além da detecção e segmentação de objetos na imagem, a arquitetura da CNN, permite compreender questões volumétricas do objeto

(i) (ii) (iii) (iv) (v)

Figura 2 – A imagem mostra a variedade de suportes e ruídos em que o grafite pode existir. Em (i) um grafite de um gato cobrindo parte de um muro e parcialmente coberto pela árvore, (ii) um grafite usando canetões sobre uma folha, (iii) um grafite no estilo *bomb* cobrindo parte de um vagão de metrô, (iv) um *stencil* sobre uma parede e (v) um grafite que usa elementos da fachada do prédio na sua própria composição. Fonte: Flickr

presente na cena (SONG; XIAO, 2016, p. 815).

No trabalho de RUSSAKOVSKY et al. (2015) os autores demonstram nos resultados (RUSSAKOVSKY et al., 2015, p. 235) como diversas CNNs já possuem precisões significativas no reconhecimento e distinguem não somente classes genéricas de objetos como “Cão”e “Gato”, mas a partir da introdução de classificação fina de imagens, também reconhece espécies como “gato persa”, “gato egípcio” e “gato siamês” e “husky siberiano”, “dálmata” e “pastor alemão policial” para cães entre outras classes destes e outros animais. Dado o desenvolvimento atual destas redes em reconhecimento de elementos específicos, vislumbramos o emprego destas redes para a identificação de grafites.

Trabalhos como o de JAIN; LEE; JIN (2009) obtiveram resultados importante para a identificação de *crews*[2] e casamento de imagens do mesmo grafite (PARRA; BOUTIN; DELP, 2012; PARRA et al., 2013) – estes trabalhos serão descritos na seção de Trabalhos Relacionados. Porém, os recentes trabalhos, não apresentam solução à identificação de grafites no meio urbano ou que não sejam similares aos presentes do *dataset*.

Como os trabalhos anteriores não solucionam a identificação do grafite dentro de uma área urbana qualquer — que pode ser por exemplo uma cidade, bairro ou área predeterminada —, sem que haja a intervenção humana para a detecção e classificação do grafite, nós propomos neste trabalho a investigação da identificação de grafite a partir das redes neurais convolucionais.

Como resultado obtivemos classificador para o grafite a partir do emprego de uma arquitetura VGG-16. O modelo PRC obteve o melhor resultado quando observado os verdadeiros positivos 76,9% e falsos negativos 19.9% aplicado ao Flickr. Entretanto quando os mesmos modelos são usados no ambiente urbano, com imagens do GSV,

[2] *Crew* é um grupo de pichadores que assinam seus grafites com a *tag* que os identificam como parte do grupo

o modelo PRC+GSV obteve uma melhora de 9% de sensibilidade a pesar de uma diminuição de 8% no valor preditivo positivo em relação ao modelo PRC.

No capítulo 1 realizamos uma introdução ao problema, o objetivo e os resultados alcançados. O capítulo 2 abordamos referenciais teóricos para este trabalho, na seção 2.6 abordaremos os trabalhos relacionados ao problema. O capítulo 3 descrevemos os objetivos e metodologias utilizadas na coleta de imagens, a rotulagem das imagens, definição da classificação, metodologia de treino e teste e criação de um multiagente.

O capítulo 4 expomos os resultados dos treinamentos das redes CNNs. A seção 4.3 é dedicado à descrição dos resultados da coleta das imagens com MAS (multi-agent system). Posteriormente no capítulo 5 realizamos a conclusão do trabalho realizado.

2 REFERENCIAL TEÓRICO

Abaixo apresentamos o referencial teórico sobre o grafite, redes neurais profundas, Google Street View, Open Street Map, ao sistema Multiagente e por fim apresentamos trabalhos relacionados.

2.1 Grafite

No âmbito da antropologia e urbanismo, em trabalhos como o de LASSALA (2010, p. 20) e de RICARDO CAMPOS (2017, p. 3) o grafite é um fenômeno presente em diversos contextos históricos como também geográficos, este nasce de uma propensão do humano de manifesta sua opinião, se fazer ouvido pelo outro e até mesmo ocupar e apropriar-se do espaço público. O grafite, muitas vezes informa e tantas vezes transgressor, não deve ser visto apenas como uma poluição visual, mas um fenômeno de comunicação (LASSALA, 2010, p. 27).

No universo do grafite, que não cabe a este trabalho esmiuçar, a autora BÁRBARA HYPOLITO (2015, p. 29-31) trás o grafite como uma expressão gráfica resultante da combinação do grafite, estêncil Figura 2(iv), lambe[1] e pixação. Já LASSALA (2010, p. 34-78) trás, além dos elementos apresentados pela autora anterior, o bomb, o grapixo e também os *stickers*[2] como elementos do grafite. Já no extenso trabalho realizado em CAMPOS (2007, p. 291-308) o autor apenas trás a classificação de *tags*, *throw ups* que compõe o *bombing* Figura 2(iii), — Chamado de *bomb* pelos dois autores anteriores — e *"hall of fame"* o grafite artístico Figura 2(v).

Apesar da divergência dos elementos que compõe o fenômeno para todos os autores, é evidente a existência do grafite como resistência e protesto (LASSALA, 2010, p. 20), mas também como instrumento de ação (BÁRBARA HYPOLITO, 2015, p. 175-179) que impacta o indivíduo que observa fazendo o grafite um elemento do entendimento dos conflitos sociais e do uso do espaço do público (BÁRBARA HYPOLITO, 2015, p. 202-203), (CAMPOS, 2007, p. 6) e (LASSALA, 2010, p. 38).

[1] Papéis com mensagens fixado a paredes e postes com cola.
[2] Adesivos semelhantes ao stencil.

2.2 Redes Neurais Profundas - Deep Learning

Nos últimos anos o aprendizado de máquina (*Machine Learning* - ML) vem realizando progressos significativos com auxílio de redes neurais artificiais profundas – também conhecidas como arquiteturas de *Deep Learning* (LECUN; BENGIO; HINTON, 2015, p. 438). Tais redes neurais profundas são criadas a partir de neurônios artificiais e são utilizadas para diversos objetivos como o de processar dados, reconhecer objetos e classificação de elementos a partir de atributos como cor, forma etc. Esta extração de atributos é o aspecto central do aprendizado profundo, áreas como de reconhecimento de voz (HANNUN et al., 2014), classificação de texto (ZHANG; ZHAO; LECUN, 2015) e reconhecimento facial (TAIGMAN et al., 2014; SCHROFF; KALENICHENKO; PHILBIN, 2015) vêm obtendo resultados competitivos ou até mesmo superiores à de humanos na realização destas tarefas. A arquitetura típica destas redes – que pode ser vista na Figura 3 – conta com uma camada de entrada (*Input Layer*), onde os dados são inseridos e propagados para os neurônios seguintes, a última camada de neurônios é chamada de camada de saída (*Output Layer*). Já as camadas intermediárias são conhecidas como camadas ocultas (*Hidden Layer*) (SRIVASTAVA et al., 2014, p. 1929)(CAI et al., 2017, p. 2137)(LECUN; BENGIO; HINTON, 2015, p. 437).

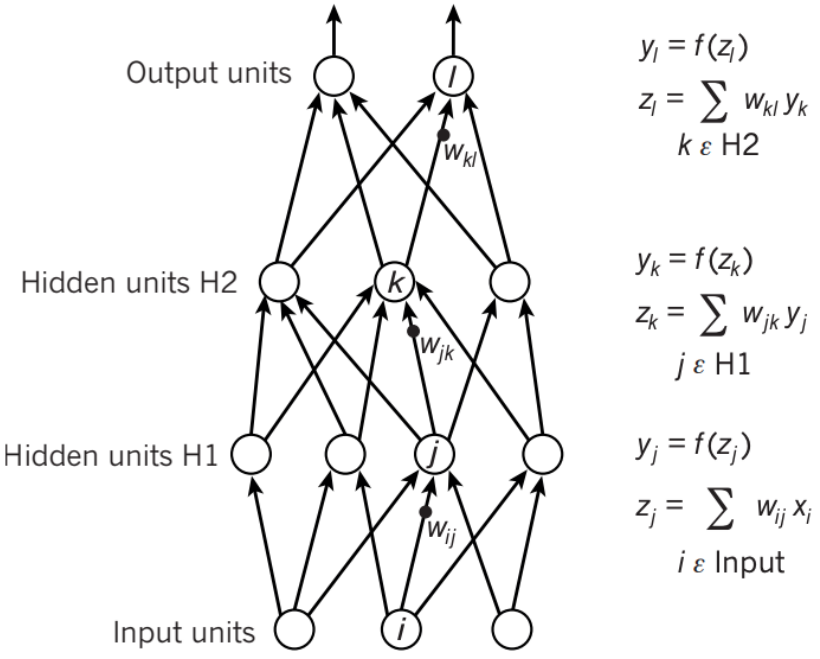


Figura 3 – Representação de uma rede neural e equações para calcular o *forward* para as camadas seguintes. A rede neural conta com duas camadas ocultas e uma camada de saída. Autor: (LECUN; BENGIO; HINTON, 2015).

Na área da computação visual — campo do aprendizado de máquina responsável

por ensinar máquinas a enxergar e reconhecer elementos em imagens(SONG; XIAO, 2016) — as Redes Neurais Convolucionais (CNNs), são consideradas o estado da arte no desafio de classificação de imagens (RUSSAKOVSKY et al., 2015, p. 235)(LECUN; BENGIO; HINTON, 2015, p. 439).

2.2.1 Redes Neurais Convolucionais

As CNNs tiveram a sua primeira aparição em FUKUSHIMA (1975), nela o autor concebeu a primeira rede neural com múltiplas camadas e convoluções. A partir dos trabalhos LECUN et al. (1989, 1998) foi combinado a técnica de *backpropagation* com as CNNs, permitindo que as redes profundas obtivessem resultados promissores. Unindo a utilização de novas técnicas como o ReLU, que ajudam em treinamentos mais profundos, e *dropout* (SRIVASTAVA et al., 2014), que aumenta a generalização dos modelos, somado ao avanço nos últimos anos da tecnologia de GPUs permitiram a diminuição drástica do tempo de treinamento das redes (LECUN; BENGIO; HINTON, 2015, p 439).

As CNNs possuem estruturas semelhante a uma *multilayer perceptron* (MLP) — também é chamado de *fully connected - FC* por ter todas as unidades da camada conectadas a todas unidades da camada seguinte — entretanto uma CNN conta com camadas especiais de convolução (DUMOULIN; VISIN, 2016, p. 12) responsáveis pela extração de padrões, *pooling* (DUMOULIN; VISIN, 2016, p. 10) responsável por simplificar e diminuir as dimensões da saída da convolution anterior e *dropout* (SRIVASTAVA et al., 2014) que evitam ajustes do modelo ao *dataset* de treinamento – também conhecido como *overfitting*.

As unidades convoluções — também chamados de *kernels* — em CNNs criam *feature maps* de tamanho AxA que ao deslizar sobre uma imagem de tamanho BxB (B > A) extraem padrões locais. Assim as primeiras camadas muitas vezes detectam bordas (YOSINSKI et al., 2015, 2014, p. 12), porém as próximas convoluções tendem a expressar conceitos mais sofisticados de padrões como demonstramos na Figura 4.

Esta característica de gerar padrões locais faz com que as primeiras camadas de convolução, após a entrada, não usem todos os atributos da imagem ao mesmo tempo como faria uma *fully connected*, mas ter o comportamento de explorar a localidade espacial na imagem, onde *pixels* próximos tendem a estar relacionados.

No processamento de imagens, em uma CNN, a ideia de conexão e formulação de atributos se dá implicitamente pela distância espacial entre os *pixels*. Por fim as CNNs usam camadas totalmente conectadas MLPs ou FCs ao final da arquitetura para realizar as classificações. Por fim é usada uma *softmax* que também é uma MLP totalmente conectada, acoplada ao classificador, esta estrutura usa uma função de ativação entre zero e um para cada classe de saída e também divide pela soma das saídas, esta operação dá a probabilidade de a imagem de entrada estar em uma deter-

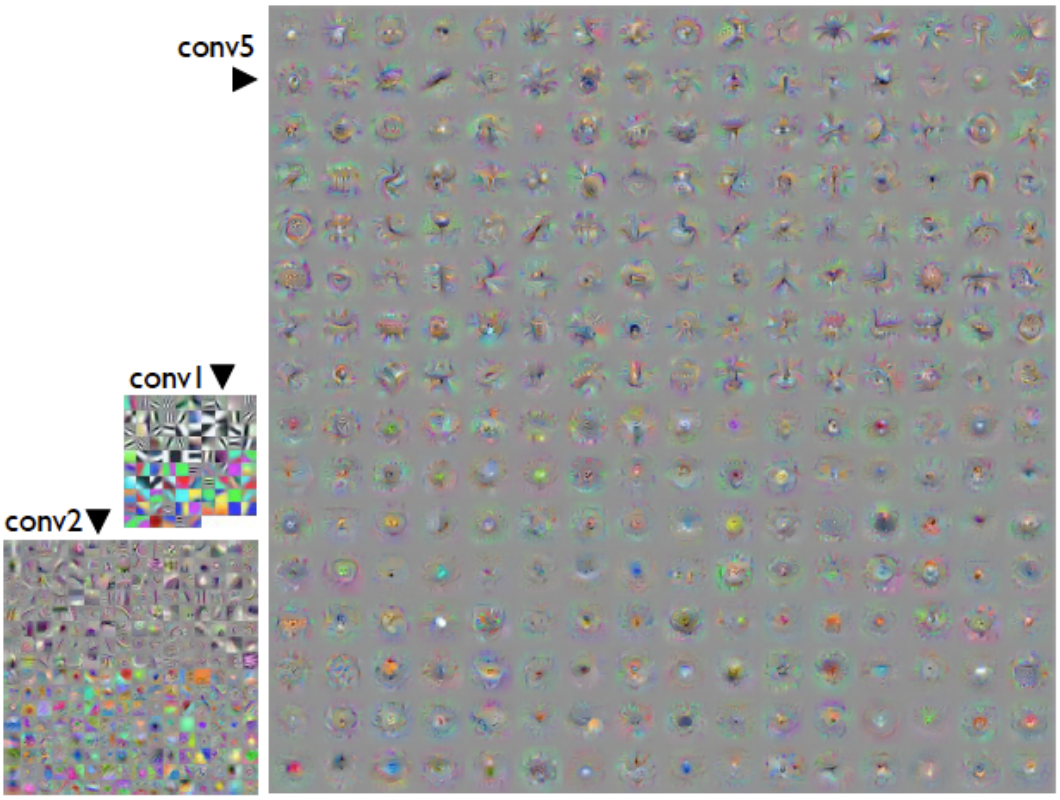


Figura 4 – Em **conv1** podemos ver a detecção de padrões e bordas simples enquanto que em **conv5** padrões mais complexos são observados. Fonte: (YOSINSKI et al., 2015).

minada classe. A combinação das camadas de *convolution*, *pooling*, *fully connected* e *softmax* são chamadas de arquiteturas. A arquitetura da VGG-16 é apresentada na Figura 5.

Ainda no campo de identificação e detecção de imagens diversas arquiteturas como a VGG16 (Figura 5) (SIMONYAN; ZISSERMAN, 2014), e Inception V3 (GoogLeNet) vem obtendo resultados de classificação semelhante à de humanos, na mesma tarefa, em competições como a ILSVRC (RUSSAKOVSKY et al., 2015, p. 242-244) que utilizam grandes *datasets* como a ImageNet (DENG et al., 2009).

2.2.2 ImageNet

A ImageNet é um dos *datasets* mais usados como *benchmarks* para a classificação de imagens e detecção de objetos(RUSSAKOVSKY et al., 2015), o *dataset* conta com 1000 classes diferentes organizados através da arquitetura da WordNet (MILLER, 1998). Cada uma das mil classes contém em média 1,200 imagens. Por sua riqueza de representação esta é muitas vezes escolhidas para o pré-treinamento de CNNs antes de serem aplicadas a novas tarefas.

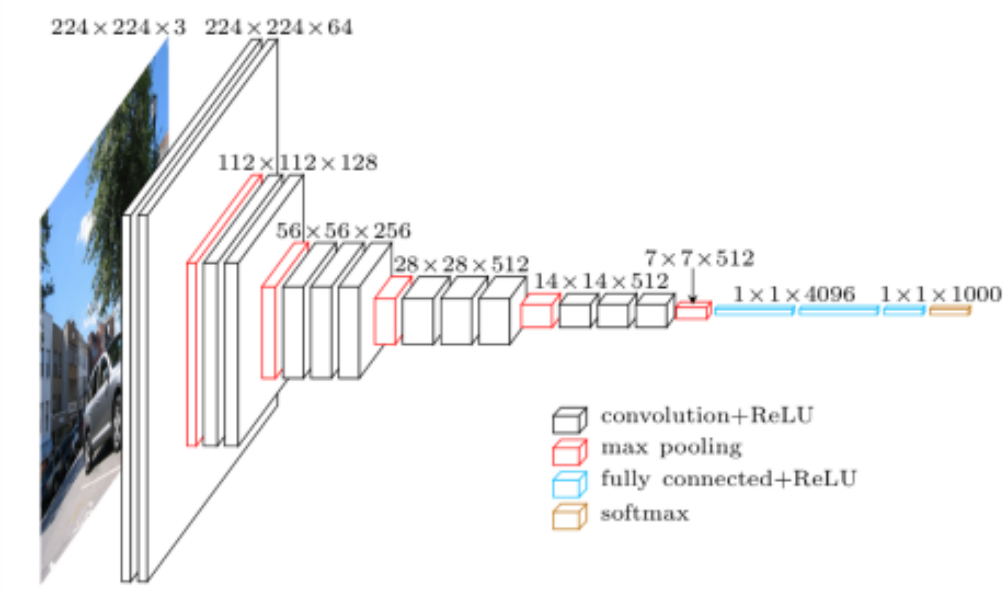


Figura 5 – A arquitetura da CNN VGG16. Fonte: (BLIER, 2016, p. 7)

2.2.3 Transfer Learning

O *transfer learning* (aprendizado por transferência) é um método de *machine learning* em que um modelo desenvolvido para uma tarefa é reutilizado como ponto de partida para um novo modelo para uma segunda tarefa. Hoje uma arquitetura CNN possui aproximadamente 60 milhões de parâmetros, aprender diretamente cada um desses parâmetros com um *dataset* com alguns milhares de imagens é complexo (OQUAB et al., 2014, p. 1718).

Em um modelo novo todos os parâmetros da CNN são inicializados com *Gaussian* aleatório (SHIN et al., 2016, p. 1291), o benefício do *Transfer Learning* está no aceleramento do tempo necessário para criar e treinar um modelo, assim reutilizando a arquitetura e conhecimentos adquiridos no modelo. Isso ajuda a acelerar o processo de treinamento do novo modelo e acelera os resultados, como demonstramos na Figura 6.

Entretanto a técnica de transferência exigirá que o domínio do problema seja semelhante, como por exemplo em que ambos os domínios sejam de identificação de objetos em imagens e também precisará que o domínio dos dados seja semelhante, como por exemplo a identificação de objetos de imagens que compartilhem características.

Com base nas características das CNNs, que são demonstradas ao longo das seções anteriores, e o sucesso que as mesmas vêm demonstrando em diversos trabalhos no campo do reconhecimento de imagens (SIMONYAN; ZISSERMAN, 2014; RUSSAKOVSKY et al., 2015; SONG; XIAO, 2016; SZEGEDY et al., 2015), nós propomos a utilização destas redes para a detecção de grafite.

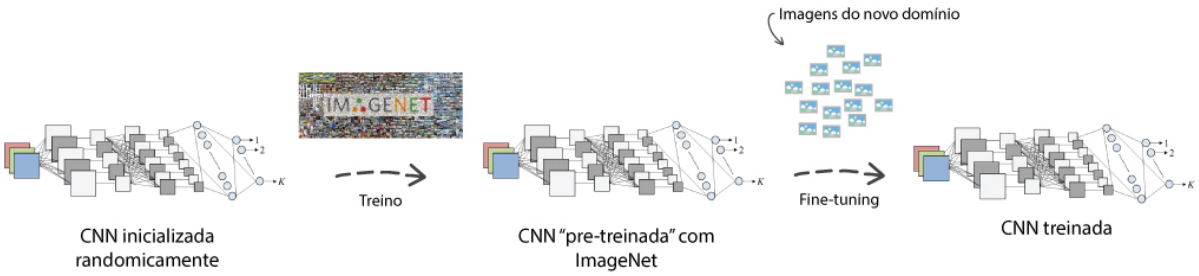


Figura 6 – A imagem demonstra como é realizado o processo de "pré-treinamento" de uma CNN com um *dataset* chamado de ImageNet e posteriormente o modelo é treinado para um novo domínio.

Com o propósito de acelerar o processo, de treinamento da rede, neste trabalho utilizaremos uma CNN pré-treinada com o *dataset* ImageNet (seção 2.2.2). A abordagem tradicional para este processo, de *fine-tuning*, nos modelos pré treinados com ImageNet é por meio da substituição da camada FC1000, que representa as 1000 classes, por uma nova de tamanho igual ao número de classificadores desejado – esta metodologia será utilizada na seção 3.3.2.1. Neste trabalho também apresentaremos uma nova abordagem com a reutilização não apenas dos pesos dos parâmetros da CNN, através do *fine-tuning*, mas também a reutilização da FC1000 para o novo domínio reutilizando uma classe pré-disposta a ser a classe classificadora do grafite. A técnica é nova e a metodologia utilizar será apresentada através dos modelos PRC e PIRC na seção 3.3.2.

2.3 Google Street View

O *Google Street View* - (GSV) é uma ferramenta de vista panorâmica disponibilizada pela Google através de seus softwares online Google Map e Google Earth(ANGUELOV et al., 2010, p. 34). Através desta ferramenta o usuário pode visualizar imagens em 360° na horizontal e 180° na vertical (SUPPORT, 2018). Como pode ser visto na Figura 7 a partir de uma coordenada de latitude e longitude o GSV permite ao usuário uma vista esférica de 360° do local, a cada 12 metros é possível gerar uma nova esfera de visualização (ZAMIR; SHAH, 2010, p. 257).

O GSV tem sido usado em diversos projetos para extração automática de informações, trabalhos como o de densidade demográfica de carros (GEBRU et al., 2017), pesquisas de criminologia (VANDEVIVER, 2014), estimativa da densidade infantil em áreas carentes (ODGERS et al., 2012) etc. No trabalho de PROCTOR (2011, p. 2018) encontramos a utilização do GSV como ferramenta para a extração de conteúdo e contexto de obras de artes através do projeto Google Art Project, entretanto o projeto foca no mapeamento interno de museus e suas respectivas obras de arte, deixando de forma, por exemplo, os grafites que usam como suporte as muros e paredes, objeto

Figura 7 – Visualização espérica das imagens do Google Street View.

do estudo desta dissertação.

Apesar dos trabalhos apresentados (GEBRU et al., 2017; VANDEVIVER, 2014; ODGERS et al., 2012; PROCTOR, 2011) terem como objetivo a extração de informação e contexto geográfico através do GSV, nenhuma deles disponibiliza uma forma automática para a coleta de imagens. A API[3] disponibilizada pelo Google Street View permite a coleta das imagens a partir da localização (latitude e longitude) e o angulo da câmera, porém não disponibiliza uma forma automática de extrair um conjunto de imagens que representa uma determinada área.

2.4 Open Street Map

A iniciativa OpenStreetMap - OSM[4] tem como finalidade permitir a colaboração e compartilhamento de informações geográficas(BENNETT, 2010, p. 26), trabalhos como HAKLAY (2010); GOODHUE; MCNAIR; REITSMA (2015) demonstram a sua boa precisão na localização de ruas (HAKLAY, 2010, p. 689) (ZHANG; MALCZEWSKI, 2018, p. 8). Dada a sua precisão, trabalhos como HENTSCHEL; WAGNER (2010, p 1649) apresentam resultados promissores para a navegação de robôs autônomos em tempo real, a partir de informações detalhadas sobre ruas, trilhas, ferrovias, cursos de água, pontos de interesse, uso da terra e informações sobre a construção presentes no local.

A iniciativa possui a ferramenta HOT Export Tool[5] que é capaz de delinear uma área de extração dos bancos de dados geográficos. Os dados são obtidos no formato PBF (*Protocolbuffer Binary Format*), o padrão de extração da OSM contém algumas entidades primitivas para representar as ruas e seus elementos: A entidade *way* representam ruas e estradas (Figura 8(ii)), cada instância da entidade *way* conta de dois e dois mil nodos, o nodo é um entidade *node*, e cada um dos nodos representa um fragmento da rua ou elemento como um semáforo como é demonstrado na Figura 8(i). Já a entidade *relations* representa a relação entre *ways*, por exemplo, um conjunto

[3] Acessível em https://developers.google.com/maps/documentation/streetview
[4] Acessível em https://www.openstreetmap.org
[5] Acessível em http://export.hotosm.org

de ruas que forma uma rótula. Todas as instâncias de entidades possuem *tags* que descrevem o significado do elemento específico ao qual estão relacionados.

```
<node id="298884269" lat="54.0901746" lon="12.2482632" user="SvenHRO" uid="46882" visible="true"
version="1" changeset="676636" timestamp="2008-09-21T21:37:45Z"/>
<node id="261728686" lat="54.0906309" lon="12.2441924" user="PikoWinter" uid="36744" visible="true"
version="1" changeset="323878" timestamp="2008-05-03T13:39:23Z"/>
<node id="1831881213" version="1" changeset="12370172" lat="54.0900666" lon="12.2539381"
user="lafkor" uid="75625" visible="true" timestamp="2012-07-20T09:43:19Z">
  <tag k="name" v="Neu Broderstorf"/>
  <tag k="traffic_sign" v="city_limit"/>
</node>
```

(I)

```
<node id="298884272" lat="54.0901447" lon="12.2516513" user="SvenHRO" uid="46882" visible="true"
version="1" changeset="676636" timestamp="2008-09-21T21:37:45Z"/>
<way id="26659127" user="Masch" uid="55988" visible="true" version="5" changeset="4142606"
timestamp="2010-03-16T11:47:08Z">
  <nd ref="292403538"/>
  <nd ref="298884289"/>
  ...
  <nd ref="261728686"/>
  <tag k="highway" v="unclassified"/>
  <tag k="name" v="Pastower Straße"/>
</way>
```

(II)

Figura 8 – Na imagem é possível visualizar a estrutura OSM no formato PBF. O item (i) contém três instâncias de *Node*, que compõem a instância rua (entidade *Way* em (ii).

2.5 Sistemas Multiagente

Sistemas Multiagente (MAS) é uma área da computação que é caracterizado por um conjunto de agente autônomos e heterogêneos que podem ou não interagir em um ambiente compartilhado (Figura 9). De forma geral estes agentes devem cooperar, coordenar e negociar para atingir um propósito (WOOLDRIDGE, 2009, p. 19). Ainda segundo o autor WOOLDRIDGE (2009, p. 15, 27) a autonomia é parte fundamental que o distingue MAS de outros sistemas e objetos, ao permitir os agentes decidam por si próprios se devem ou não realizar uma ação.

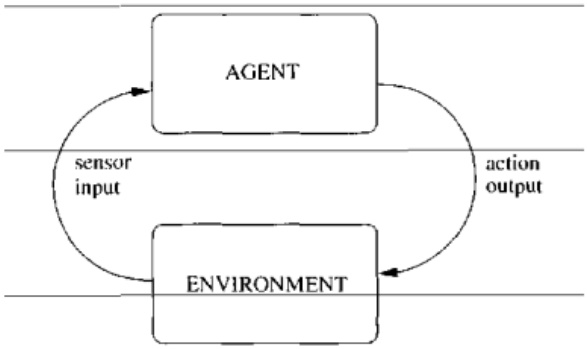


Figura 9 – Um agente em seu ambiente. O agente recebe informações sensoriais do ambiente e produz como ações de saída que o afetam. A interação geralmente é contínua e não finalizadora. Fonte: (WOOLDRIDGE, 2009, p. 16)

Dentro do universo do MAS há o problema de patrulhamento de um ambiente

conhecido como *Multi-Agent Patrolling Problem - MAPP* (HU; ZHAO, 2010), nele objetiva-se que uma equipe de agentes deve otimizar uma métrica coletiva movendo-se através de um determinado modelo de ambiente.

No trabalho de pesquisa sobre algoritmos de patrulhamento de PORTUGAL; ROCHA (2011, p. 137) podemos identificar como o problema pode ser utilizado para o encontro de grafites de forma autônoma na cidade. O autor lembra que o MAPP também tem como objetivo monitorar e supervisionar ambientes, obter informações, procurar objetos e detectar anomalias, a fim de proteger os terrenos da invasão, o que envolve frequentes visitas a todos os pontos da infraestrutura.

No trabalho de SAMPAIO; SOUSA; ROCHA (2016) os autores propõem uma solução alternativa às estratégias heurísticas que em sua maioria das vezes são baseadas em conhecimento prévio do ambiente e uma coordenação central. Nas abordagens chamadas de ZR-STRATEGIES (SAMPAIO; SOUSA; ROCHA, 2016, p. 159)(Figura 10) cada agente escreve uma informação no nodo que passa e o agente decide seu próximo nó de destino sentindo a informação armazenada neste nodo. O autor conclui que a nova estratégia de *0-range* têm desempenho competitivo quando comparadas com as estratégias de *1-range* comparadas no artigo. A estratégia proposta demonstra-se mais descentralizadas, dando maior autonomia e robustez aos fracassos dos agentes (SAMPAIO; SOUSA; ROCHA, 2016, p. 162).

ZrAgent (*S* – a 1-range strategy; *m* – memory policy)

```
1   x_decdata = read decdata stored in node x
2   x_decdata[x] = apply the update rule of strategy S
3   x_decdata[x].time = current time
4   for each node y for which x_decdata has an entry (which
        depends on parameter m)
5       ag_decdata[y] = most recent among
                x_decdata[y] and ag_decdata[y]
6       x_decdata[y] = ag_decdata[y]
7   write back x_decdata on node x
8   x = chooses the next node, using the decision rule of S ap-
        plied to the neighbor's decision data in ag_decdata
9   move to node x
10  upon arrival at node x, go to step 1
```

Figura 10 – Pseudo código da estratégia de *0-range*. Fonte: (SAMPAIO; SOUSA; ROCHA, 2016, p. 159)

2.6 Trabalhos Relacionados

Como vimos na seção 2.1, sobre o Grafite, há uma vasta bibliografia de trabalhos que têm como objetivo a definição do grafite através dos artefatos geográficos, sociológicos e antropológicos. Nesta seção são apresentados estes trabalhos e outros 5 trabalhos relacionamento com o desenvolvimento de ferramentas para a extração do grafite de imagens. No campo da sociologia e geografia destacamos trabalhos como de FERRELL (1993) descreve sobre o grafite e a correlação da criminalização do mesmo, demonstrando que a proibição da manifestação política subversiva fez o grafite porta-voz das comunidades como o *hip-hop*.

O trabalho de HAWORTH; BRUCE; IVESON (2013) descreve a ocorrência de grafites na cidade de Sidney a partir das ocorrências de remoção solicitadas a cidade. O mesmo analisa zonas de maior propensão da área e período, porém o autor conclui que a imprecisão na coleta de dados não permite que um modelo seja criado e sugere zonas livres – onde é permitido a aplicação de grafite – através de uma tolerância por parte da cidade com o grafite e reconhecendo esta como expressão artística.

Contudo o trabalho de MEGLER; BANIS; CHANG (2014) aplica a técnica de *Geographically weighted Regression* (GWR) para a predição de dispersão de grafites de São Francisco. O modelo decorrente do trabalho descreve a incidência de dois-terços dos grafites presentes na cidade (MEGLER; BANIS; CHANG, 2014, p. 70), entretanto o modelo baseia-se em informações obtidas a partir de solicitações de remoção pelo departamento responsável pela política de remoção de grafites da cidade. Ainda neste trabalho os autores descrevem a correlação encontrada entre a concentração de jovens que residem na área com a concentração de grafites. Entretanto as áreas comerciais possuem uma menor incidência de pessoas jovens por região, porém as áreas mantêm-se com alta incidência de grafites, sugerindo então que há uma migração da população para a prática (Megler, 2014, p.71).

No campo da computação temos artigos como de JAIN; LEE; JIN (2009); TONG et al. (2011) que apresentam pesquisas de similaridade entre grafite. A técnica apresentada nos artigos permite que dando uma imagem os sistemas retornam N imagens similares que estão no banco de dados, contribuindo com a justiça na identificação na origem dos grafites. Os artigos apresentam acurácia de 49.1% (TONG et al., 2011, p. 4) até 85.9% (JAIN; LEE; JIN, 2009, p. 4) para o processo de recuperação de *tags* presentes no *dataset*.

Os sistemas dos artigos acima citados fazem o uso da técnica *Scale Invariant Feature Transform* (SIFT) para cada imagem no banco de dados (BROWN; LOWE, 2007). A técnica transforma a imagem numa coleção de vetores de características locais invariantes às transformações e o algoritmo divide-se em 4 etapas: O primeiro é a construção de um espaço que consiste de um conjunto de versões simplificadas da

imagem original, conhecido como pontos chaves, já o segundo passo foca na detecção das extremidades deste espaço-escala e o terceiro passo gera a orientação destes pontos chaves. Por fim no quarto passo são criado os descritores dos pontos chaves: definido por um vetor contendo os valores de todas as entradas do histograma. Nos sistemas a partir da imagem submetida, pressupondo que a mesma é um grafite, é construído seus descritores. Logo em seguida cada descritor da imagem é avaliado e a sua similaridade é mensurada com os descritores locais de cada imagem do banco de dados. A técnica retornando assim imagens que possuem maior similaridade entre seus descritores. Além das imagens, os sistemas retornam informações relacionadas a estas imagens como as gangues associadas (LOWE et al., 1999, p. 3).

Já nos artigos PARRA; BOUTIN; DELP (2012); PARRA et al. (2013) os autores usam a detecção de *tags* que ligam grafite com possíveis gangues, como técnica os autores usam a segmentação de cores para a obtenção de grafites textuais presentes nas imagens através de descritores SIFT, obtendo acurácia entre 33.6% e 66.7% (PARRA et al., 2013, p. 181). Entretanto a técnica foca na recuperação de *tags* similares presentes em um *dataset* e não na detecção de grafite com base em sua complexidade de cores e formas, como é abordado nesse trabalho.

3 OBJETIVOS E METODOLOGIA

3.1 Objetivos Geral e Específicos

O objetivo deste trabalho é avaliar o uso de Redes Neurais Convolucionais e Sistemas Multiagente para localizar e mapear grafites em cidades, a partir de imagens em nível de rua provenientes do *Google Street View*. Este trabalho tem aos seguintes objetivos específicos:

1. Identificar a presença de grafites em imagens urbanas utilizando redes neurais convolucionais;
2. Avaliar técnicas de treinamentos de redes neurais convolucionais aplicadas a identificação de grafites;
3. Criar e disponibilizar um *dataset* rotulado com imagens com e sem a presença de grafites;
4. Propor métodos para extrair automaticamente imagens urbanas em nível de rua, adequadas para a identificação de grafites, a partir de mapas de cidades.

3.2 Criação dos Datasets

A criação do *dataset* é um importante passo para a definição dos conceitos que serão aprendidos posteriormente pelos modelos CNNs, hoje bases como o Flickr(JOULIN et al., 2016) e o próprio GSV são fontes ricas de imagens, permitindo que se crie *datasets* que representem diversos conceitos. Esta metodologia tem como objetivo a coleta de imagens para a composição de duas classes, a do grafite e a classe que representa o *não-grafite*.

3.2.1 Escolha das bases

Para este trabalho escolhemos as bases de imagens do Flickr e GSV como fontes dos *datasets* pela natureza das suas imagens e como representam o mundo: De

forma genérica as imagens que estão presentes no Flickr[1] representam um recorte do ambiente, contendo predominantemente imagens focada em um objeto e são descritas razoavelmente pelos rótulos vinculados a esta imagem (HUISKES; THOMEE; LEW, 2010, p. 531). Já o GSV apresenta imagens do ambiente de rua, coletadas e remontadas para gerar uma visão 360° de um determinado ponto geográfico (ANGUELOV et al., 2010, p. 34) conforme Figura 11 e demonstramos mais detalhadamente no Anexo D.

A tarefa de criação dos dois *datasets* será realizada em duas etapas: A primeira etapa constitui o *dataset* a partir das imagens obtidas do Flickr, já a segunda etapa constitui o *dataset*, que serão imagens extraídas a partir do GSV. Ambas coletas serão utilizadas para realizar *fine-tuning*, o primeiro *dataset* usado nos ajustes dos modelos da seção 3.3.2 e o segundo *dataset* para o ajuste proposto na seção 3.3.2.4.

3.2.2 Flickr

Para a metodologia de extração de imagens do Flickr utilizamos a extração a partir da API disponível[2] pela própria empresa. Propomos a extração de imagens a partir de *tags* que rotulam as imagens na base, para este trabalho propomos a extração de imagens rotuladas com a *tag "graffiti"* para compor o *set* de imagens que definem o grafite e imagens com a *tag "street"* para compor o *set* que representa o *não-grafite*, conforme Figura 11.

O período de busca proposta é de imagens que forma inseridas pelos usuários entre 2004 e 2017[3] para ambas as *tags*. O *dataset* será construído de forma que imagens que sejam rotuladas tanto como *graffiti* e *street* serão descartadas evitando assim imagens duplicadas.

Como pode ser visto no trabalho de THOMEE et al. (2016, p. 64–73) e HUISKES; THOMEE; LEW (2010, p. 528) as imagens são fracamente rotuladas pela comunidade do Flickr, ou seja, com isso os rótulos apresentam distorções na sua classificação por utilizar a colaboração na marcação. Embora essas imagens sejam fracamente classificadas, uma visão prévia das imagens coletadas mostra a predominância de rótulos.

3.2.3 Google Street View

Atualmente a API do GSV permite que seja extraída as imagens a partir de um ponto geográfico definido pelo usuário, porém a API não tem suporte a navegação, não permitindo uma sequência encadeada de imagens que representam uma rua seja extraída ou a exportação de imagens a partir de uma área pré-definida.

[1] O Flickr é um serviço de armazenamento de imagens e vídeos, que permite a comunidade organizar o conteúdo por *tags*(THOMEE et al., 2016).

[2] Acessível em https://www.flickr.com/services/api/

[3] A data inicial é a data inicial do Flickr e a data final o início deste trabalho.

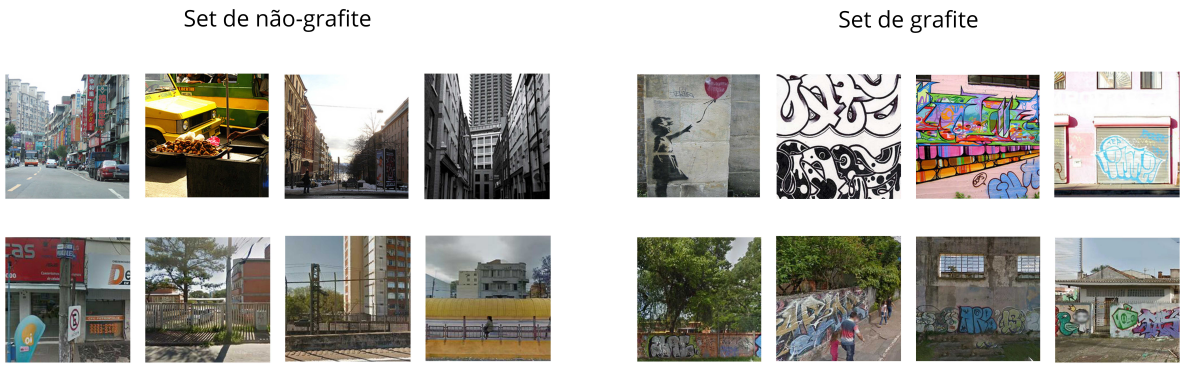


Figura 11 – A Figura demonstra imagens coletadas para compor os *sets* dos *datasets*. A primeira linha de imagens correspondem a exemplos do Flickr e a segunda linha com imagens extraídas do GSV.

Dado o problema de navegação e extração automatizada de imagens do GSV, propomos a coleta de dados através de um MAS – proposto na seção 3.4 –, estes sistemas são indicados para a cobertura e patrulhamento de ambiente, aonde agentes independentes, que se comunicam ou não, cooperem para atingir um objetivo (WOOLDRIDGE, 2009) que no nosso caso é a cobertura de uma área do Google Street View. Logo propomos que os agentes tenham a capacidade de navegar, através dos dados OSM, até as coordenadas corretas de forma arranjada para realizar a extração das imagens.

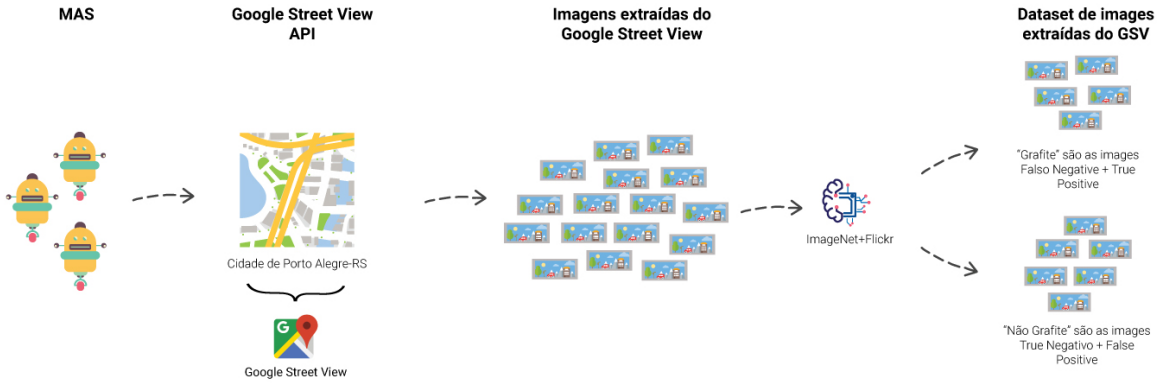


Figura 12 – A Figura descreve o MAS atuando com o modelo ImageNet+Flickr para a extração das imagens do GSV.

Como mencionado acima, a capacidade de extração das imagens está intimamente ligada ao agente, por este motivo a metodologia de utilização para extração de imagens pela API do GSV está descrita junto a seção 3.4.2 que trata do comportamento esperado pelos agentes.

As imagens coletadas, através do MAS, serão utilizadas para dois propósitos: A análise preditiva dos modelos na identificação do grafite e compor um *dataset* que será utilizado para o *fine-tuning* (Figura 12) dos modelos – propostos da seção 3.3.2.

3.3 Metodologia de Seleção de Arquitetura de CNN e Técnicas de Treinamento

O primeiro passo da nossa metodologia será a avaliação das classes do modelo VGG16 treinado com ImageNet para a identificação de grafite, assim almejamos extrair as classes, deste modelo, que estão mais propícias a identificação do grafite. O segundo passo desta metodologia são os experimentos com *fine-tuning* no modelo.

3.3.1 Avalição do modelo VGG16+ImageNet

A primeira parte da metodologia será a aplicação do nossos *dataset* de grafite na rede VGG16 treinada apenas com a ImageNet para observar qual classe no domínio da ImageNet terá mais ativações para o *dataset* do Flickr, conforme Figura 13 (i). propomos a utilização do *1-top* (classe mais ativa) como forma de avaliação do modelo.

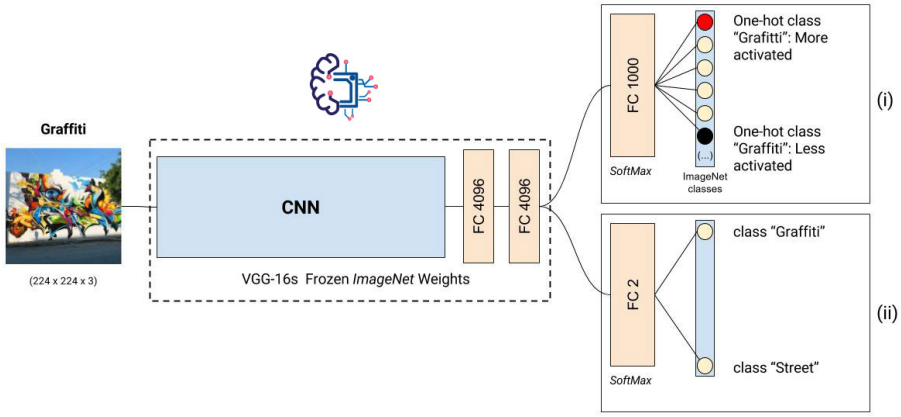


Figura 13 – A arquitetura proposta para nossos experimentos, (i) demonstra nossa abordagem proposta para o *fine-tuning* da rede e (ii) o *fine-tuning* tradicional.

Ao fazer isso, pretendemos verificar quais classes do ImageNet estão mais relacionadas às classes de destino, servindo como base para o processo de *fine-tuning*. Portanto, as classes obtidas a partir desta observação serão utilizadas em três modelos de classificadores descritos na seção seguinte que trata dos experimentos.

3.3.2 Experimentos com o modelo VGG16+ImageNet

A metodologia de treinamento e teste divide-se em dois momentos, o primeiro conjunto possui 3 experimentos com o objetivo verificar a capacidade de adaptação dos modelos para o contexto de identificação de grafite a partir de um *fine-tuning* com o *subset* do Flickr. Já o segundo momento será destinado a um *fine-tuning* no modelo que obtiver o melhor resultado no primeiro momento. As imagens utilizadas neste último treinamento serão extraídas do GSV através do MAS (Figura 14).

Figura 14 – A imagem demonstra o *subset* do Flickr utilizado para o *fine-tuning* em três modelos. O melhor modelo é utilizado para novo *fine-tuning* com imagens de grafite do GSV.

Para todos os quatro experimentos, nós propomos a utilização da arquitetura CNN VGG-16 (SIMONYAN; ZISSERMAN, 2014), a CNN em questão alcançou resultados de estado da arte para o desafio da ImageNet em 2016. A VGG-16 é uma rede neural mais profunda do que seus predecessores, contendo múltiplas camadas convolucionais (SIMONYAN; ZISSERMAN, 2014, p. 8). A CNN em questão não é excessivamente profunda como suas sucessoras VGG-19 (SIMONYAN; ZISSERMAN, 2014), GoogLeNet(SZEGEDY et al., 2015) e ResNet (HE et al., 2016), o que exigiria um longo treinamento com o hardware disponível para a pesquisa.

Nós inicializamos a rede com os pesos pré treinados com o *dataset* ImageNet para os três experimentos iniciais e o último será a ImageNet com o Flickr mais o *dataset* do GSV. Para o treinamento das nossas redes usamos os mesmos parâmetros para todas: utilizando *Stochastic Gradient Descent* com taxa de aprendizagem $\alpha = 1e - 4$, com *batch size* de 128 e execução máxima de 150 épocas. Os *datasets* divididos em 70-15-15 para treino, validação e teste. Utilizaremos a acurácia na avaliação dos quatro treinamentos.

Neste momento serão realizados três experimentos para verificar a adaptabilidade da CNN VGG16 para o novo contexto de identificação do grafite:

O primeiro experimento aborda a método clássica do *fine-tuning* com a substituição da MLP (FC1000) por uma nova MLP (FC2) com duas classes para representar o grafite e o *não-grafite* – como pode ser visto na Figura 13(ii). O segundo e o terceiro experimento será demonstrado uma nova técnica de reutilização da MLP para o novo contexto, no *fine-tuning* a MLP original (FC1000) será mantida e será analisada a sua reutilizados para a detecção do grafite e *não-grafite* – como pode ser visto na Figura 13(i).

Tanto o treinamento como a evolução do *fine-tuning* dos três experimentos serão executados de forma randômica, com um *subset* (A.2.2) construído a partir do *dataset* de imagens extraídas do Flickr. O *subset* em questão contém as classes de grafite e não-grafite, ambas as classes contam com um conjunto de 25,000 imagens de treino

e 7,500 para o conjunto de teste (Figura 15).

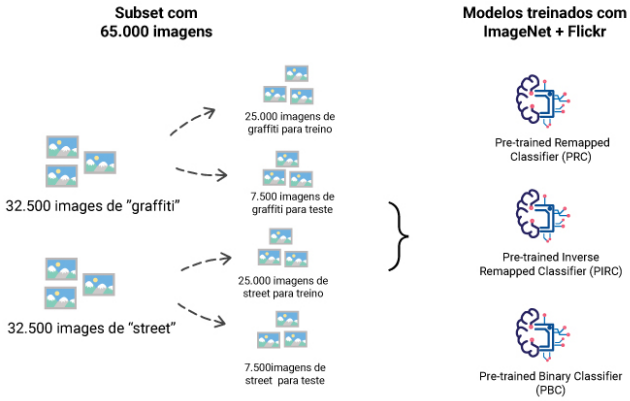


Figura 15 – A imagem demonstra o *subset* do Flickr utilizado para o *fine-tuning* na ImageNet nos modelos PRC, PIRC e PBC.

3.3.2.1 Pre-trained Binary Classifier (PBC)

Esta é a abordagem tradicional do *fine-tuning* para o reuso de uma rede para um novo domínio. Nesta configuração, a camada totalmente conectada FC1000 da VGG-16 é substituída por um classificador binário, como demonstrado na Figura 13 (ii).

3.3.2.2 Pre-trained Remapped Classifier (PRC)

Para esta abordagem, nós realizaremos o *fine-tuning* da rede mantendo a camada FC1000, remapeando as classes *"Comic Book"* e *"Jacamar"* para *"Graffiti"* e *"Street"*, respectivamente. Para as demais 998 classes não serão apresentadas imagens no treinamento. Nosso objetivo é fazer uso de conhecimentos previamente obtidos para facilitar o aprendizado das novas classes.

3.3.2.3 Pre-trained Inverse Remapped Classifier (PIRC).

Nesta abordagem nós invertemos as classes *graffiti* e *street* do modelo PRC, mapeando grafite para a classe menos ativada. Esta configuração é proposta de forma a possibilitar a comparação com a configuração do modelo PRC. A razão é verificar a associação do conhecimento a classe que estava menos pré-disposto a identificar, verificando assim se o processo de *fine-tuning* será menos eficiente no aprendizado do novo conceito.

Acima descrevemos a abordagem utilizada para o *fine-tuning* nos três modelos com a arquitetura VGG-16, utilizaremos a rotulação proposta na seção 3.3. A partir dos resultados obtidos utilizaremos o melhor modelo para ser o identificador do grafite do MAS e coletar imagens que darão subsídio para o novo *fine-tuning* proposto na seção seguinte.

3.3.2.4 Best Pre-trained Classifier (BPC).

O quarto e último experimento será a investigação de uso de imagens do GSV com o intuito de melhorar a classificação do grafite pelo melhor modelo dos experimentos anteriores PBC, PRC e PIRC. Tanto o treinamento como a evolução do *fine-tuning* do experimento será executado de forma randômica, com um *subset* construído a partir do *dataset* de imagens do GSV (Anexo A.2.3) como pode ser observado na Figura 16.

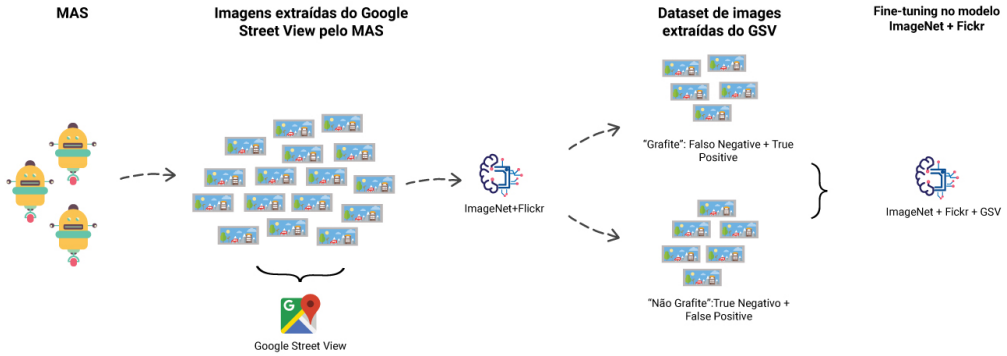


Figura 16 – O *dataset* de imagens extraídas do GSV pelos agentes e usada no *fine-tuning* do modelo.

3.4 Metodologia do MAS

Propomos neste trabalho a criação de um MAS capaz de percorrer um ambiente pré-definido – como por exemplo um centro urbano – e, através de seu comportamento, extrair as imagens do GSV e submetê-las a análise do modelo CNN proposto na seção 3.3.2. O MAS, chamado de Graphium[4], receberá um ambiente de execução, a definição do número de agentes e o modelo de identificação de grafite que será usado pelos agentes do MAS.

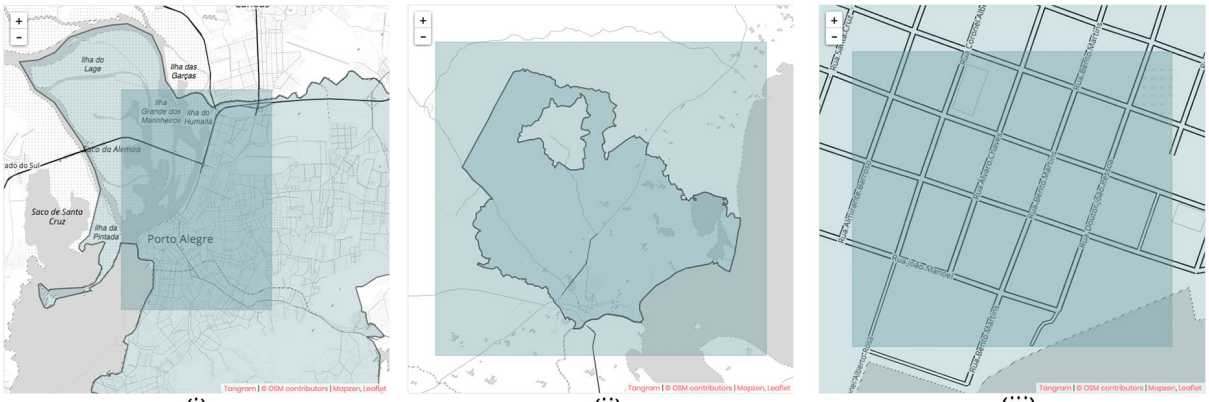
3.4.1 Ambiente

Para a criação do ambiente do MAS é fundamental a definição da geolocalização e metadados das ruas de uma determinada área, visto que as informações de latitude e longitude são essenciais para que os agentes possam realizar a extração de imagens do *Google Street View*. Para a definição do ambiente será usada a ferramenta HOT Export Tool descrita na seção 2.4.

Para este trabalho serão criados três ambientes que representam as áreas interesse para a extração de imagens: A região Metropolitana de Pelotas, a região portuária de Pelotas e o centro histórico de Porto Alegre (Figura 17). A escolha dos 3 ambientes se dá a partir do conhecimento do elevado número de grafites destas

[4]Acessível em https://github.com/glaucomunsberg/graphium e sua documentação acessível em https://glaucomunsberg.github.io/graphium/

regiões. Em adição a escolha de regiões centrais e portuária segue os padrões de regiões mencionadas em trabalhos MEGLER; BANIS; CHANG (2014); FERRELL (1993, p. 64), onde os autores descrevem estes ambientes como os mais prósperos para a manifestação do grafite.



(i) (ii) (iii)

Figura 17 – A imagem mostra as três regiões que compõem os ambientes. A região (i) representa o centro histórico de Porto Alegre - HCP, (ii) representa a áreas metropolitana de Pelotas - PMe e (iii) representa a região portuária de Pelotas - PP.

Mediante a aplicação Graphium cada ambiente é criado a partir do carregamento do arquivo pdf que representa uma área e extraída a partir da ferramenta HOT Export. A aplicação armazena as informações das entidades e instâncias do OSM que posteriormente serão usados pelos agentes para percorrer o ambiente.

3.4.2 Comportamento do agente

A metodologia para o comportamento dos agentes neste trabalho difere das estratégias *0-range* e *1-range* proposto em PORTUGAL; ROCHA (2011) e SAMPAIO; SOUSA; ROCHA (2016), ambas as estratégias tem como princípio o patrulhamento da área de forma que nodos centrais – que neste caso seriam os que apresentam grafite – tenham um maior número de visitas por parte dos agente envolvidos na tarefa.

Este trabalho tem como objetivo a identificação do grafite e cobertura do ambiente e não o patrulhamento, dado que o ambiente não sofrerá mudanças em relação a presença de novos grafites. Logo o comportamento esperado para cada agente, concorrentemente, deve ser a de selecionar uma das ruas do ambiente que ainda não foi percorrida pelos demais agentes evitando percorrer diversas vezes o mesmo local.

Ao visitar a rua o agente deverá percorrer os nodos que compõem a rua, como demonstramos na Figura 13. Cada um dos nodos representa um trecho variado de distância(Figura 18), com isso o agente deverá identificar então o intervalo de espaço entre o nodo atual e o próximo nodo, para então dividir esta distância em quadrantes de 12 metros(ZAMIR; SHAH, 2010, p. 257). Cada quadrante representará uma esfera

3D de imagens do GSV (Figura 7).

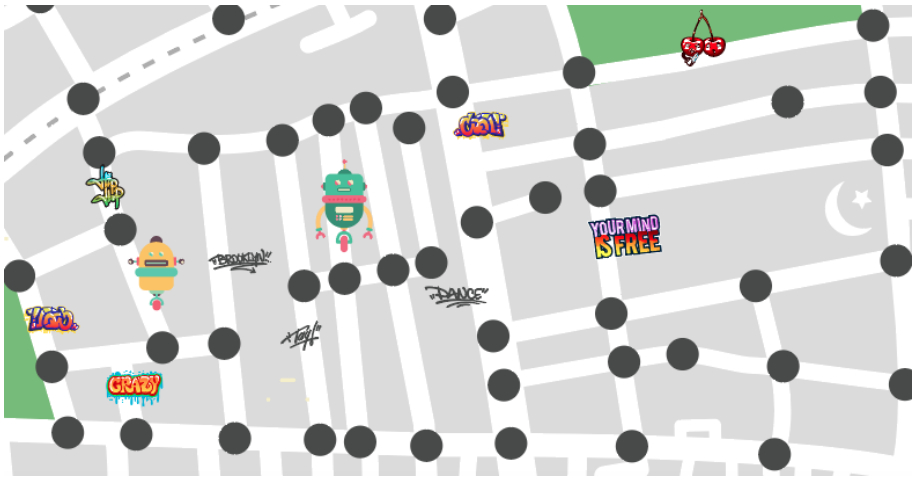


Figura 18 – A figura ilustra dois agentes percorrendo um ambiente, os pontos representam interseções entre os segmentos de ruas. A identificação do grafite é o alvo dos agentes.

O próximo passa do agente é a orientação correta da direção frontal da rua, a formula para encontrar a direção frontal – demonstrada abaixa – permite que seja identificada, em grau, a direção a partir de dois pontos (nodos) da rua. A partir do calculo, o agente é capaz de realizar a rotação de +90° e -90° posicionando a câmera do GSV para as laterais esquerda e direita. As laterais contam com as fachadas dos prédios, casas, muros etc. que são reconhecidas como suportes dos grafites. Ao posicionar a visão do agente para as laterais da rua, evitaremos a análise de imagens do horizonte frontal e traseiro da rua. O pseudocódigo que representa o comportamento do agente está presente no Anexo C.

$$\theta = atan2(sin(\Delta long) \cdot cos(lat2), cos(lat1) \cdot sin(lat2) - sin(lat1) \cdot cos(lat2) \cdot cos(\Delta long)) \quad (1)$$

A formula retorna o ângulo em graus entre o ponto A(lat1,lng1) e B(lat2,lng2). atan2 é a função que retorna um número entre $-\pi$ e π.

3.4.3 Módulo de identificação de grafite

O módulo de identificação de grafite é um componente de cada agente: sua função é a identificação de grafite a partir de um dos modelos propostos neste trabalho nas imagens do GSV. O modelo será carregado de acordo com a parametrização da execução do *swarm*, permitindo assim que diferentes modelos sejam usados para a identificação de grafite em um mesmo ambiente. A primeira execução utilizará o melhor modelo dos experimentos propostos na seção 3.3.2.

A Figura 19 demonstra o agente fazendo a cobertura de todos os nodos de uma determinada rua e coletando, através da API do GSV, as imagens que serão subme-

tidas ao modelo. A identificação do grafite na imagem se dará pelo método *Top-1*, neste método apenas a classe mais ativada – das classes presentes na FC1000 do modelo – (com o maior peso) é considerada a classe correta, neste caso apenas será classificado como grafite imagens que tiverem o maior peso de ativada para a classe "*Comic Book*".

As imagens coletadas e os metadados referente a sua localização, o identificador do modelo de identificação de grafite e a classificação – se é um grafite ou não, assim como o identificador do agente e do *swarm* são armazenados na aplicação para a análise posterior da cobertura e classificação.

Figura 19 – A imagem descreve o agente em uma, cada nodo é dividido em quadrantes de 12 metros, para que o agente analise as imagens de fachadas de prédios e casas, muros etc.

3.4.4 Comportamento do swarm

Cada execução de cobertura de uma área é chamada de *swarm*, a execução do *swarm* conta com um número máximo de agentes e a execução será encerrada após a cobertura total das ruas do ambiente selecionado.

Como foi visto na seção anterior, que define o comportamento dos agentes, os mesmos são modelados para seguirem o princípio da autônoma – cada agente é independente em relação a posição e análise realizada pelos demais agentes – e cooperam (WOOLDRIDGE, 2009, p. 3, 27) na cobertura dos ambientes até que o objetivo seja alcançado.

Com o propósito de assegurar a cobertura total do ambiente, o *swarm* é capaz de gerar novos agentes caso detecte que um ou mais agentes tenham falhado, na sua missão de cobertura, respeitando o numero máximo de agente configurado para aquela execução de cobertura.

3.4.5 Avaliação das execuções do MAS

Dado que a avaliação é pautada pelo sucesso na identificação do grafite, para a avaliação dos modelos empregados na cobertura dos ambientes, propostos na seção 3.4.1, utilizaremos a sensibilidade (KAWAMURA, 2002) dos modelos, para mensurar o quanto o modelo é eficaz em identificar corretamente o grafite, a prevalência, com o intuito de analisar a porção de imagens com grafite em relação ao total de imagens dos ambientes, assim como valor preditivo positivo e negativo, que mensuram a probabilidade do preditor estar correto sobre sua predição positiva e negativa respectivamente.

4 RESULTADOS

Este capítulo dedicamos a demonstrar os resultados da criação dos *datasets* criados, os resultados que obtivemos para os três modelos treinados com o *dataset* Flickr bem como o modelo treinado com o *dataset* GSV. Também apresentamos os resultados dos modelos PRC e PRC+GSV junto as execuções do MAS.

4.1 Resultados dos datasets

Esta seção é dedicada aos resultados obtidos na extração das imagens do Fickr e do Google Street View, a construção dos *datasets* e também descrevemos como foram compilados os *subsets* utilizados nos treinamentos dos modelos.

4.1.1 Coleta no Flickr

Na coleta foram utilizadas as *tags*(rótulos) de busca “graffiti” e “street”, a primeira representar a classe grafite e a segunda para a classe *não-grafite*. Com a API coletamos 719.352 imagens do Flickr, para compor as duas classes, sendo 351.338 imagens para a *tag* “graffiti” e 368.040 imagens para a *tag* “street”. Deste *dataset* criamos um *subset* de 65 mil imagens, 25 mil imagens para o treino e 7.500 para teste em cada uma das classes. Veja a representação da criação do *dataset*[1] e o *subset*[2] de treinamento.

A coleta pela API acontece pelos parâmetros *tag* e o período desejado de extração das imagens, o sistema retorna um JSON como os identificadores de forma paginada, das imagens vinculadas a *tag*. Entretanto, durante a coleta, a API demonstrou instabilidade para retornar identificadores para um período longo, como por exemplo a lista de imagens vinculados a uma *tag* para o período de um ano, para solucionar a limitação, imposta pela API, foram extraídas as imagens pelas *tags* mês a mês durante o período indicado na metodologia.

[1] O acesso as imagens está disponível no Anexo A.2.1
[2] O acesso as imagens está disponível no Anexo A.2.2

4.1.2 Coleta no Google Street View

Ao total foram coletadas 23.741 imagens do GSV através do MAS (Figura 12) na área que corresponde ao centro histórico de Porto Alegre como é demonstrado na Figura 17(i). No processo de coleta, que compõem este *dataset*, as imagens foram classificadas pelo modelo PRC acoplado ao MAS – apresentaremos os resultados do modelo na seção 4.2.2.

Um *subset* foi construído a partir deste *dataset*, neste conjunto de imagens, que foram coletas e classificas pelos agentes no ambiente, foram submetidas a análise humana para catalogar as imagens de acordo com os valores preditivos. A categorização das imagens foram realizadas manual pelo autor através da verificação manual das imagens coletas pelo MAS. Com base nesta classificação foi construído o *subset* com as classes grafite e *não-grafite*, contendo 181 imagens com grafite e 23,519 imagens sem grafite.

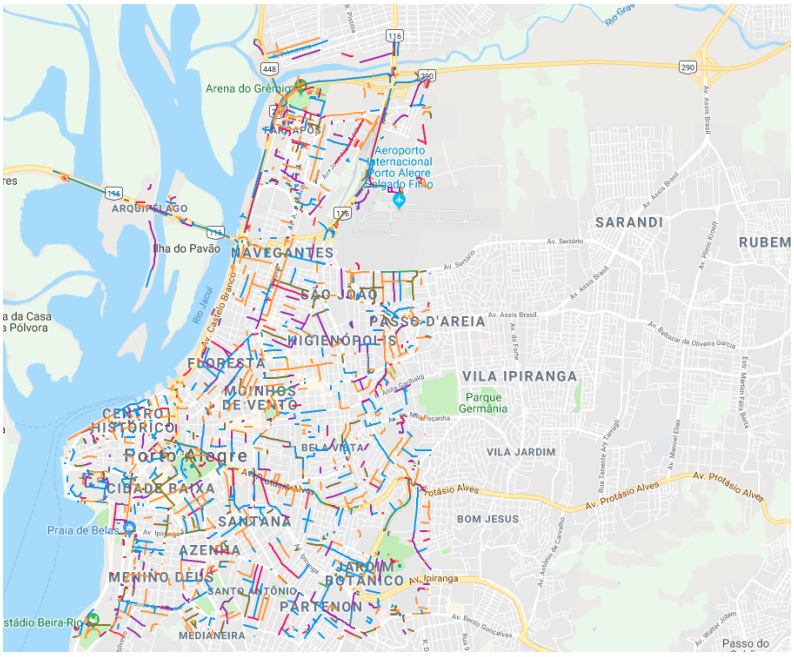


Figura 20 – O centro histórico de Porto Alegre a as áreas que foram cobertas para a coleta de imagens pelo MAS.

4.2 Classificação com CNN

Nesta seção descrevemos os resultados sobre a escolha dos rótulos de ativação na VGG16 + ImageNet e os resultados do *fine-tuning* nos quatro modelos propostos na seção de Metodologia de Treinamento e Teste.

4.2.1 Resultados da escolha dos rótulos

Os histogramas na Figura 21 e Figura 22 demonstram o número de vezes que cada classe da ImageNet foi ativada para cada uma das imagens do nosso *dataset* do Flickr com *graffiti* e *street*. Para este experimento nós apresentamos as imagens, do *dataset* do Flickr, a rede VGG16 treinada apenas com a ImageNet.

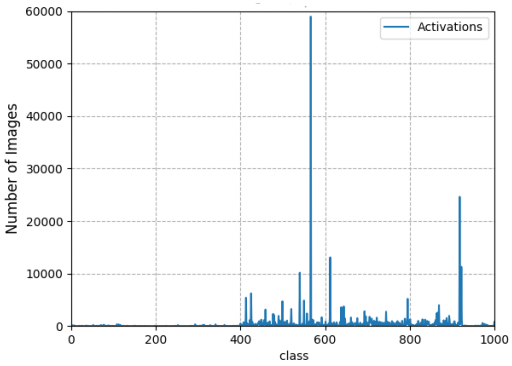


Figura 21 – Histograma da ativação Top-1 para “graffiti”.

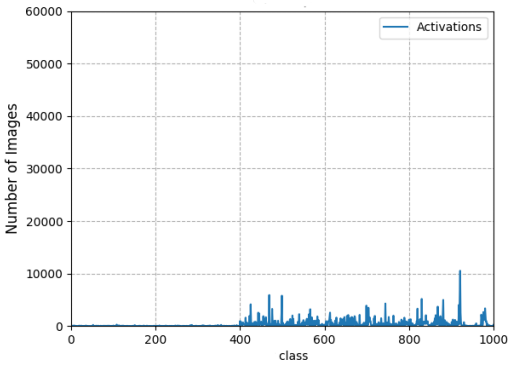


Figura 22 – Histograma de ativação Top-1 para “street”.

A Tabela 1 resume os resultados obtidos, demonstrando as cinco principais ativações para a classe de imagens *graffiti* e a menos ativada. Enquanto a classe “freight car” foi a classe que teve mais ativações para *graffiti*, a análise do *dataset* revelou que muitas imagens de grafite contêm vagões de trem como suporte. Um exemplo é demonstrado na Figura 33 (iv). As características dos vagões de trem também são muitos similares ao de caminhões, o que poderia levar a rede a classificar as laterais dos caminhões como grafites dado que o suporte (lateral dos vagões) possuem os mesmos frisos horizontais e metálicos. Com isso escolhemos a classe subsequente “*Comic book*”, já que a mesma demonstra similaridade com o grafite, possuindo forte saturação de cores e bordas sempre bem definidas. A classe “Jacamar”, um pássaro específico, foi a classe menos ativada para as imagens com grafite.

Figura 23 – A esquerda um exemplo de imagem da classe “*Comic Book*“ (i), demonstando saturação na paleta de cores e bordas definidas do grafite (ii)(iii). A direita, um exemplo do dataset que inclui um vagão de trem com aplicação de um grafite (iv). Fonte: Flickr

Já a Tabela 2 demonstra os resultados das cinco classes mais ativadas para o *set* de *street*. Como pode ser observado as classes que representam elementos presentes nos cenários de rua são os mais ativados, dado as suas especificidades classes como “Japanese spaniel”, um cachorro de porte pequeno, teve o menor número de ativações.

Com os resultados observados na Tabela 1 e 2, nós escolhemos mapear na classe “Jacamar“, a menos sensível ao grafite para a classe *street*. Esta escolha da classe *street* foi feita, pois a classe já está sendo usado como *proxy* para a classificação do *não-grafite*. Quando formos usar a rede treinada para realizar as predições com o MAS, nós consideraremos qualquer classe que não seja *Comic Book* como um resultado negativo ao grafite.

Para a identificação das classes mais ativadas para grafite e *não-grafite*, demonstradas nas Tabelas 1 e 2, foram utilizados os métodos Top-5 e Top-1 respectivamente. No método Top-1 apenas a classe mais ativada - das classes presentes na FC1000 do modelo – (com o maior peso) é considerada a classe correta, já o Top-5 utiliza as cinco classes mais ativadas como corretas.

Classe	**Synset**	**Top-5 Ocorrência**
freight car	n03393912	58,955
comic book	n06596364	24,612
jigsaw puzzle	n03598930	13,079
book jacket, dust cover, dust jacket, dust wrapper	n07248320	11,282
doormat, welcome mat	n03223299	10,155
jacamar	n01843065	0

Tabela 1 – As cinco classes mais atividades e a menos ativada para o *set* de imagens *Graffiti*.

Classe	**Synset**	**Top-1 Ocorrência**
traffic light, traffic signal, stoplight	n06874185	10,533
cab, hack, taxi, taxicab	n02930766	5,909
street sign	n06794110	5,860
cinema, movie theater, movie theatre, movie house...	n03032252	5,786
unicycle, monocycle	n04509417	4,971
Japanese spaniel	n02085782	0

Tabela 2 – As cinco classes mais atividades e a menos ativada para o *set* de imagens *Street*.

4.2.2 Resultados dos modelos pré treinados

Nas Figuras 24 e 25 é demonstrado o *loss* e acurácia durante o treinamento para os três modelos propostos na seção 3.3.2. Como é possível observar tanto o modelo

Experimentos Flickr	TPR	FNR
PBC	43,9%	43.1%
PRC	76,9%	14.9%
PIRC	72,1%	12,7%

Tabela 3 – Resultado Top-1 obtido nos testes realizados com as três redes propostas. Os valores demonstrados são a taxa dos verdadeiros positivos (TPR) e taxa dos falsos negativos (FNR).

PRC como PIRC convergem rapidamente, enquanto que o modelo PBC falha continuamente mesmo com o longo número de épocas – os resultados semelhantes para o modelo PRC e PIRC fazem que as linhas se sobreponham nos resultados nas Figuras 24 e 25. A rápida convergência dos modelos é uma evidência do benefício de reutilização da camada FC pré-treinada mesmo que o número de classe não corresponda com o número de classes do novo domínio.

Na Tabela 3 é possível visualizar o resultado final das redes. A rede sem o *fine-tuning* tem uma taxa de acerto aproximado de 6.33% para o grafite. Este é um resultado melhor que o randômico, pois seria de menos de 0.1% sobre as 1000 classes da ImageNet, entretanto esta ainda é um resultado ruim. Ao substituir a ultima camada FC por uma binária geramos um resultado igualmente fraco, já que a rede não consegue convergir, gerando resultados comparáveis à escolha aleatória.

O modelo PRC e PIRC demonstram resultados similares: PRC com uma taxa de verdadeiro positivo de 76.9% e taxa de falso positivo de 14.9%, já a PRIC obteve taxas de 72.1% e 12.7% respectivamente. Estas taxas são similares aos observados na ImageNet como pode ser visto em SIMONYAN; ZISSERMAN (2014).

Como esperado para a inversão da aplicação das classes no modelo PIRC reduziu os falsos positivos e aumento os falsos negativos, porém em ambas as classes os resultados são semelhantes.

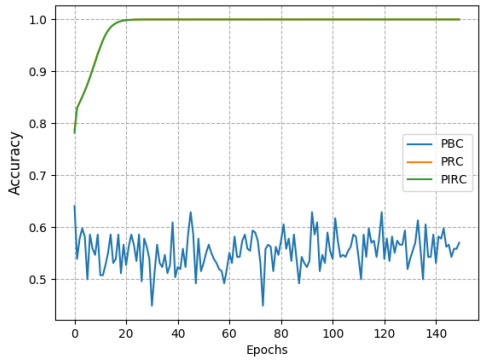


Figura 24 – Acurácia durante o treinamento para cada modelo.

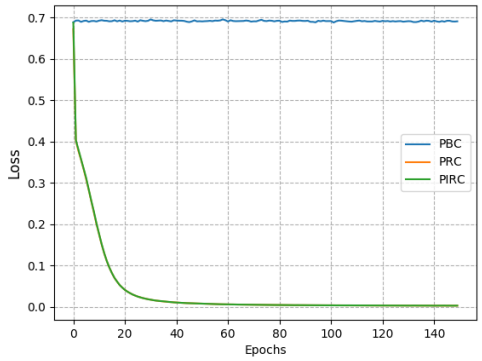


Figura 25 – *Loss* durante o treinamento para cada modelo.

Ao ajustarmos a camada, pudemos convergir rapidamente o treinamento e obter

resultados razoáveis em um conjunto de teste. Mostramos que a escolha das classes a serem remapeadas pode produzir resultados diferentes, mas essa escolha não parece ser extremamente crítica. Entretanto a camada FC binária provavelmente convergiria com uma melhor customização dos parâmetros como o tamanho de *batch* e número de épocas. Evidenciou-se assim que, usando parâmetros padrão, nos três modelos propostos até aqui, há benefícios em usar a abordagem PRC. Sendo assim a PRC é o melhor modelo e base para o modelo *Best Pre-trained Classifier* ou BPC.

4.2.3 Resultados do *fine-tuning* usando MAS

Dados os resultados obtidos com os modelos pré treinados, o *fine-tuning* foi realizado no modelo PRC com a época 20, antes do ajuste completo do modelo, para este treino utilizamos o *subset* resultante da coleta do GSV descrito na seção 4.1.2 de forma randômica e *batch size* padrão de 128. Nas Figuras 26 e 27 é demonstrada a acurácia e o o *loss* durante o treinamento para o modelo BPC – chamado de PRC+GSV por ser o modelo PRC + as imagens do GSV – proposto na seção 3.3.2 com treino no conjunto de treino e resultados das figuras com o conjunto de teste do *subset* GSV. Como é possível observar o modelo PRC+GSV converge rapidamente a partir do *subset* GSV e com o conhecimento do domínio de grafite adquirido previamente pelo modelo PRC.

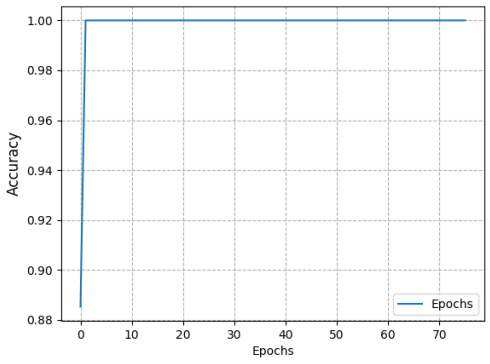


Figura 26 – Acurácia durante o treinamento para o modelo PRC+GSV.

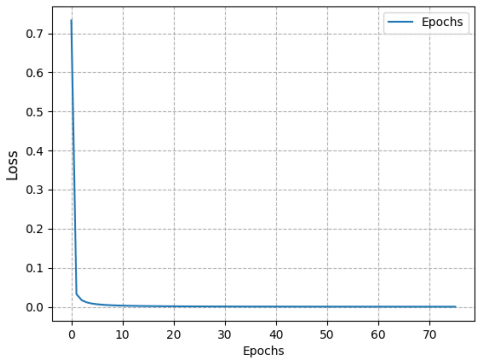


Figura 27 – *Loss* durante o treinamento para o modelo PRC+GSV.

Na Tabela 4 apresentamos os resultados comparativo entre a modelo PRC – que obteve o melhor resultado nos três primeiros testes – com o modelo PRC+GSV. A rede PRC possui uma taxa de verdadeiros positivos de 76,9% enquanto que a PRC+GSV apresenta a taxa de 71,43%, já a taxa de falsos negativos é similar para ambas as redes.

Os resultados da Tabela 3 demonstram uma queda de 5% na taxa para falsos positivos após o *fine-tuning* realizado com as imagens do GSV, já a taxa de falso

negative teve uma alteração menos significativa de 0.02%. Entretanto os resultados poderão ser validados a partir do uso destes modelos no ambiente do GSV, apesar do modelo PRC obter melhor resultados para os falsos positivos no *dataset* de teste do Flickr, os modelos serão utilizados na localização de grafite em imagens do GSV, domínio de imagens usada o *fine-tuning* do modelo PRC+GSV.

Experimento Flickr+GSV	TPR	FNR
PRC	76,9%	14.9%
PRC+GSV	71,43%	14.7%

Tabela 4 – Resultado Top-1 comparativo entre o modelo PRC e PRC+GSV. Os valores demonstrados são a taxa dos verdadeiros positivos (TPR) e taxa dos falsos negativos (FNR).

Os resultados apresentados nas Tabelas 3 e 4 correspondendo aos TPRs (verdadeiros positivos) e FNRs (falsos negativos) para os quatro modelos (PBC, PRC, PIRC, PRC+GSV) e todos os modelos e resultados foram realizados com o mesmo *subset* de teste[3] do Flickr.

4.3 Coleta de dados e identificação do grafite com MAS

Esta seção dedicamos a demonstrar o trabalho realizado com o multiagente para cobertura dos ambientes e o resultado da utilização dos modelos PRC e PRC+GSV para a identificação do grafite nas imagens encontradas pelos agentes.

4.3.1 Execuções do MAS

Foram realizadas quatro execuções com o MAS (Tabela 5) para experimentar os modelos PRC e PRC+GSV na identificação de grafites no contexto do Google Street View. Descrevemos abaixo os resultados obtidos com cada uma das execuções.

Swarm	Ambiente	Modelo	Agentes	Tempo	Imagens
PCH	PoA Centro Histórico	PRC	5	1h17m11s	23.751
PMe	Pelotas Metropolitana	PRC	5	38m47s	11.977
PP1	Pelotas Porto	PRC	5	7m20s	424
PP2	Pelotas Porto	PRC+GSV	5	7m20s	424

Tabela 5 – Resultados das execuções do MAS para os ambientes, bem como os modelos acoplados, número de agentes, tempo de execução e número de imagens analisadas.

Na Tabela 6 demonstramos os detalhes sobre o tempo de execução para cada uma dos *swarms* (PCH, PMe, PP1 e PP2), como é possível ver o tempo médio de análise de imagem por minutos(IM) é constante para as execuções PCH e PMe, bem como

[3] Veja o *subset* no Anexo A.2.2.

para PP1 e PP2. Como é possível ver a média de imagens por minutos que cada agente (IMA) realizou demonstra que execuções mais longas há uma análise de mais de uma imagem por agente para cada minuto, enquanto que nas execuções menores o tempo médio foi de 11 imagens por minuto. Aparentemente o tempo de resposta da API do GSV torna-se mais rápida a partir de um determinado número de solicitações.

Ainda na Tabela 6 podemos ver o número de agentes encerrados (AE) demonstra quantas vezes um dos agentes foi finalizado antes do termino do seu objetivo, este encerramento é ligado a oscilações de rede, erros na leitura de informações provenientes da API do GSV e erros de localizações provenientes do OSM. Para casos de encerramentos como estes, um novo agente é gerado.

Swarm	Imagens	Tempo	IM	Agentes	Erro Agente	IMA
PCH	23.751	1h17m11s	308.324	5	8	61.665
PMe	11.977	38m47s	307.102	5	12	61.420
PP1	424	7m21s	57.3050	5	3	11.461
PP2	424	7m20s	60.5714	5	1	11.616

Tabela 6 – Resultados das execuções do MAS, número de imagens, tempo de execução, bem como imagens por minuto - IM, número de agentes simultâneos, nº de agentes que tiveram erros e o número de imagens analisadas por minutos dividido pelo número de agentes - IMA.

4.3.1.1 Swarm Centro Histórico de Porto Alegre - PCH

O *swarm* PCH utilizou a região do Centro Histórico de Porto Alegre[4] como ambiente de execução e contando com 5 agentes para cobrir o ambiente. Para a identificação do grafite nas imagens, coletas pelos agentes, utilizamos o modelo PRC. O *swarm* analisou 23.751 imagens em 1 hora e 17 minutos de execução. O PCH teve como objetivo a análise das imagens do GSV com o modelo PRC, sua assertividade na identificação do grafite e subsidiar com imagens o experimento de *fine-tuning* no próprio modelo.

4.3.1.2 Swarm Pelotas Região Metropolitana - PMe

O *swarm* PMe utilizou a região do Centro Metropolitano de Pelotas[5] como ambiente de execução e contando com 5 agentes para cobrir o ambiente. Para a identificação do grafite nas imagens, coletas pelos agentes, utilizamos o modelo PRC. O *swarm* analisou 11.977 imagens em 38 minutos de execução. O PMe teve como objetivo a identificação do grafite na cidade de pelotas como um todo e também servir de paralelo na análise da execução PCH.

[4] Acessível no Anexo A.1.1.
[5] Acessível no Anexo A.1.2.

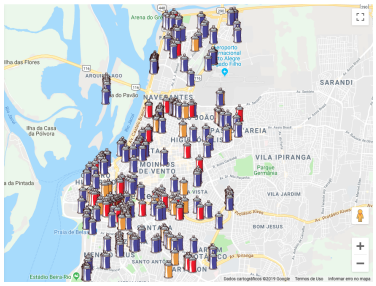
4.3.1.3 *Swarm Pelotas Região Portuária - PP1*

O *swarm* PP1 utilizou a região portuária de Pelotas[6] como ambiente de execução e contando com 5 agentes para cobrir o ambiente. Para a identificação do grafite nas imagens, coletas pelos agentes, utilizamos o modelo PRC. O *swarm* analisou 424 imagens em 7 minutos de execução. O PP1 teve como objetivo analisar a região de Pelotas mais propícia a identificação do grafite com o modelo PRC.

4.3.1.4 *Swarm Pelotas Região Portuária - PP2*

O *swarm* PP2 utilizou a região portuária de Pelotas[7] como ambiente de execução e contando com 5 agentes para cobrir o ambiente. Para a identificação do grafite nas imagens, coletas pelos agentes, utilizamos o modelo PRC+GSV. O *swarm* analisou 424 imagens em 7 minutos de execução. O PP2 teve como objetivo analisar a região de Pelotas mais propícia a identificação do grafite com o modelo PRC+GSV.

4.3.2 Representação Visual dos Resultados



HCP - Ambiente PCH com o modelo PRC

PP1 - Ambiente PP com o modelo PRC

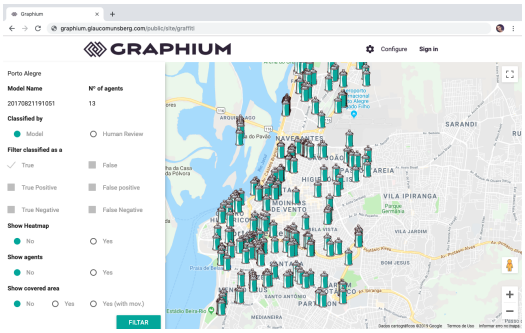
PP2 - Ambiente PP com o modelo PRC + GSV

Figura 29 – Representação da ferramenta de análise dos *swarms* e dos grafites identificados para cada uma das execuções.

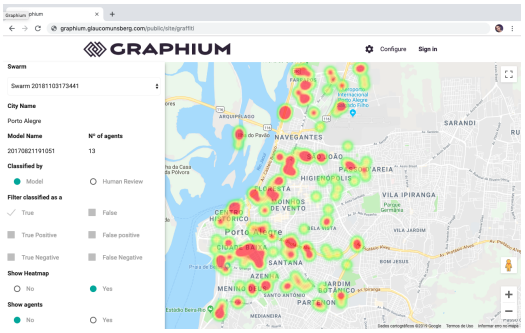
Na Figura 29 podemos visualizar representações em mapa de três das execuções, a primeira representa a região do centro histórico de Porto Alegre, a segunda e terceira imagem trás a região portuária de pelotas. Os *sprays* nas imagens são os grafites: Os vermelhos são os resultados true positive, os roxos são false positive e os amarelos para false negative identificados pelos modelos.

Já a Figura 31 demonstra o ambiente HCP após a identificação do grafite em duas formas de visualização: A primeira os grafites são identificados através de pontos (*sprays*) sobre o mapa e a segunda representação é o mapa de calor para os mesmos pontos do grafite. Com o mapa de calor é possível visualizar as zonas de mais intensidade em que os grafites se manifestam no ambiente como por exemplo no bairro Cidade Baixa e entorno de vias como a avenida Protásio Alves e avenida Farrapos.

[6] Acessível no Anexo A.1.3.
[7] Acessível no Anexo A.1.3.



HCP - Visualização por pontos



HCP - Visualização por mapa de calor

Figura 31 – Representação da ferramenta de análise da execução PCH, a primeira representa os pontos e a segunda é o mapa de calor do grafite.

4.3.3 Identificação de Grafite pelos Modelos PRC e PRC+GSV

Como é demonstrado nas Tabelas 7 e 8 a execução PCH e PMe foram as que mais abrangem territórios – Centro Histórico de Porto Alegre e região metropolitana de Pelotas, respectivamente – e para ambas utilizamos o mesmo modelo PRC como identificador de grafite. Já as execuções PP1 e PP2 são executados sobre o mesmo ambiente, região do portuária de Pelotas, porém com os modelos PRC e PRC+GSV respectivamente.

Swarm	Modelo	Nº de Imagens	TP	FP	TN	FN
PCH	PRC	23.741	92	203	23.357	89
PMe	PRC	11.977	28	192	11.739	18
PP1	PRC	424	35	19	349	21
PP2	PRC+GSV	424	40	31	337	16

Tabela 7 – Valores absolutos de imagens e a matriz de confusão true positive (TP), false positive (FP), true negative (TN) e true positive (TP) para cada execução do MAS.

Swarm	Modelo	PREV	SENS	ESPE	VPP	VPN
PCH	PRC	0.0076	0.5083	**0.9914**	0.3118	0.9962
PMe	PRC	0.0038	0.6087	0.9839	0.1273	**0.9984**
PP1	PRC	**0.1320**	0.6250	0.9484	**0.6481**	0.9432
PP2	PRC+GSV	**0.1320**	**0.7143**	0.9157	0.5634	0.9547

Tabela 8 – Tabela demonstrando a sensibilidade (SEN), especificidade (ESPE), prevalência (PREV), valor preditivo positivo (VPP) e valor preditivo negativo (VPN) para cada execução. Em negrito os melhores resultados para cada uma das categorias.

Ao analisar a Tabela de Confusão, Tabela 8, verificamos que o modelo PRC demonstrou uma menor sensibilidade[8](KAWAMURA, 2002) quando executado em am-

[8]**Sensibilidade** é o método que reflete o quanto o modelo é eficaz em identificar corretamente,

bientes maiores (PCH e PMe), porém apresentando uma especificidade[9] maior para o ambiente PCH. Porém a prevalência[10] do grafite demonstrou-se diferente para ambientes distintos, como já era esperado, Porto Alegre contém um maior número de grafites na sua região central. Ainda sobre a análise nos ambientes de PCH e PMe os valores preditivos positivos[11] se mostram significativamente discrepantes, demonstrando uma maior aptidão para a identificação de grafite no Centro Histórico de Porto Alegre em detrimento a cidade de Pelotas. Porém o valor preditivo negativo[12] são significativamente semelhantes, variando pouco menos que 0.001%.

Diferentemente das execuções PCH e PMe, que tinham ambientes distintos, as execuções PP1 e PP2 são realizadas no mesmo ambiente: Região portuária de Pelotas, neste caso os agentes, de ambas execuções, são expostos as mesmas imagens extraídas do GSV e seus resultados também são apresentados na Tabela 8. Como é possível analisar a prevalência é a mesma, dado que é o mesmo ambiente e logo são os mesmos grafites apresentados aos modelos, já a sensibilidade do modelo PRC+GSV foi maior, inclusive entre as demais execuções (HCP e PMe).

Ainda no âmbito da PP1 e PP2, a PRC obteve um resultado 3% melhor quanto a especificidade, demonstrando que a rede PRC+GSV teve esta perda em prol de 10% na sua sensibilidade. Quando comparado os VPPs e VPNs das duas execuções verificamos uma piora significativa do VPP do modelo treinado (PRC+GSV) com as imagens do GSV em relação a sua progenitora PRC. Já em relação ao valor preditivo negativo a variação é positiva a PRG+GSV, porém pouco mais de 0.01%.

Como é possível observar a sensibilidade do modelo PRC+GSV teve uma melhora após o *fine-tuning* realizado no modelo PRC, porém as melhores especificidades encontram-se nas execuções de ambientes do Centro Histórico de Porto Alegre (PCH) e Região Metropolitana de Pelotas (PMe) – ambas com o modelo PRC. Já a análise das probabilidades de uma imagem de grafite ser identificada como grafite (VPP) e uma imagem de *não-grafite* ser classificada como *não-grafite* (VPN), ambas estão relacionadas a prevalência, quando executado o modelo em ambiente que a prevalência é menor, logo espera-se que o VPN seja maior evitando assim que muitos *não-grafites* sejam classificados como grafites.

Entretanto a análise sobre a probabilidades de acerto da classificação em grafite (VPP) e não-grafite (VPN) utilizado um modelo em ambiente que a prevalência é me-

dentre todos os indivíduos avaliados, aqueles que realmente apresentam a característica de interesse.

[9]**Especifidade** é o método que reflete o quanto o modelo é eficaz em identificar corretamente os indivíduos que não apresentam a condição de interesse.

[10]**Prevalência** é a fração de imagens que apresentam as condições de interesses no *dataset* total avaliada.

[11] O **Valor Preditivo Positivo** é a probabilidade de uma imagem avaliada e com resultado positivo ser realmente a imagem com grafite.

[12]O **Valor Preditivo Negativo** é a probabilidade de uma imagem avaliada e com resultado negativo ser realmente a imagem sem grafite.

nor, um VPN maior permite evitar falsos positivos de não-grafite classificados como grafite.

5 CONCLUSÃO

As redes neurais profundas e CNNs são considerados métodos recentes e vêm acumulando diversos sucessos no campo de reconhecimento de padrões em dados, objetos em imagens, de voz etc. Este trabalho teve como objetivo avaliar o uso destas Redes Neurais Convolucionais para a identificação de grafite, abordamos também a vantagem em reuso dos classificadores de redes pré-treinadas em um novo contexto e por fim a utilização de MAS para percorrer o GSV para localizar os grafites no ambiente urbano.

Neste trabalho desenvolvemos um classificador para o grafite a partir do emprego de uma arquitetura VGG-16, obtivemos resultados promissores para alguns destes modelos (PRC e PRC+GSV). Observando a aplicação dos modelos ao *dataset* com imagens do Flickr verificamos, através das Tabelas 3 e 4 que o modelo PRC obteve o melhor resultado quando observado os verdadeiros positivos 76,9% e falsos negativos 19.9%, entretanto quando os mesmos modelos são usados no ambiente urbano, com imagens do GSV, o modelo PRC+GSV obteve um resultado superior de 9% de sensibilidade apesar de uma diminuição de 8% no valor preditivo positivo em relação ao modelo PRC.

5.1 Contribuições

A principal contribuição deste trabalho está no emprego da arquitetura VGG16 pré-treinada com ImageNet para a identificação de grafite a partir do *fine-tuning* com imagens do Flickr e GSV. Para desempenhar esta tarefa foi proposto uma nova abordagem de reuso dos classificadores da ImageNet, aproveitando o conhecimento adquirido previamente pelo modelo, para acelerar o treinamento da rede para um novo domínio.

Outra contribuição é na extração de imagens do GSV através do trabalho coordenado de agentes (MAS) capazes de cobrir um ambiente e realizar as identificação de imagens para apresentação ao modelo. Ainda sobre o MAS há a versatilidade da aplicação em configurar o ambiente de cobertura, número de agentes e definição do modelo acoplado para aquela análise.

Tanto a aplicação, os modelos e os *datasets* utilizados neste trabalho estão disponíveis para futuros trabalhos no GitHub[1].

5.2 Trabalhos futuros

Como trabalho futuro propomos a investigação de outros *datasets* que possam melhorar os resultados obtidos, nos experimentos deste trabalho, para a identificação de grafite. Ainda no campo da identificação de grafite através de modelos com CNNs, como trabalho futuro consideramos a aplicação de modelos para a classificação de grafites a partir estilos: Agrupando os grafites por suas similaridades já que o grafite conta com estilos distintos como o *bomb*, estêncil, lambe e a pixação como é observamos nas imagens 2 e 1.

No campo do MAS, como foi apresentado o sistema desenvolvido é capaz de percorrer ambientes pré-definidos, como trabalho futuro, considera-se a evolução da aplicação para a criação automática de ambientes a partir de estímulos ou áreas não cobertas. Atualmente o sistema tem como comportamento a cobertura total de uma área, logo considera-se a utilização de algoritmos de patrulhamento de ambientes, assim os agentes seriam capazes de vigiar as modificações nas imagens provenientes do GSV, submetendo essas novas versões das imagens aos modelos.

12

REFERÊNCIAS

ANGUELOV, D. et al. Google street view: Capturing the world at street level. **Computer**, [S.l.], v.43, n.6, p.32–38, 2010.

BÁRBARA HYPOLITO, B. de. **CIDADE, CORPO E ESCRITAS URBANAS**: Cartografia no espaço público contemporâneo. 2015. Dissertação (Mestrado em Ciência da Computação) — UNIVERSIDADE FEDERAL DE PELOTAS.

BENNETT, J. **OpenStreetMap**. [S.l.]: Packt Publishing Ltd, 2010.

BLIER, L. **a brief report of the heuritech deep learning meetup**. [Online; acessado em 20-Setembro-2018], `https://blog.heuritech.com/2016/02/29/a-brief-report-of-the-heuritech-deep-learning-meetup-5/`.

BROWN, M.; LOWE, D. G. Automatic panoramic image stitching using invariant features. **International journal of computer vision**, [S.l.], v.74, n.1, p.59–73, 2007.

CAI, C.-H.; KE, D.; XU, Y.; SU, K. Symbolic manipulation based on deep neural networks and its application to axiom discovery. In: NEURAL NETWORKS (IJCNN), 2017 INTERNATIONAL JOINT CONFERENCE ON, 2017. **Anais...** [S.l.: s.n.], 2017. p.2136–2143.

CAMPOS, R. **Porque pintamos a cidade?**: uma abordagem etnográfica do graffiti urbano. [S.l.]: Fim de Século, 2010.

CAMPOS, R. M. d. O. Pintando a cidade: uma abordagem antropologica ao graffiti urbano. **Universidade Aberta, Portugal**, [S.l.], 2007.

DENG, J. et al. Imagenet: A large-scale hierarchical image database. In: COMPUTER VISION AND PATTERN RECOGNITION, 2009. CVPR 2009. IEEE CONFERENCE ON, 2009. **Anais...** [S.l.: s.n.], 2009. p.248–255.

DUMOULIN, V.; VISIN, F. A guide to convolution arithmetic for deep learning. **arXiv preprint arXiv:1603.07285**, [S.l.], 2016.

FEIXA, C.; OLIART, P. **Juvenopedia**: Mapeo de las juventudes iberoamericanas. [S.l.]: NED Ediciones, 2016.

FERRELL, J. **Crimes of style**: Urban graffiti and the politics of criminality. [S.l.]: Garland New York, 1993.

FUKUSHIMA, K. Cognitron: A self-organizing multilayered neural network. **Biological cybernetics**, [S.l.], v.20, n.3-4, p.121–136, 1975.

GEBRU, T. et al. Using deep learning and Google Street View to estimate the demographic makeup of neighborhoods across the United States. **Proceedings of the National Academy of Sciences**, [S.l.], p.201700035, 2017.

GOODHUE, P.; MCNAIR, H.; REITSMA, F. TRUSTING CROWDSOURCED GEOSPATIAL SEMANTICS. **ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences**, [S.l.], v.XL-3/W3, p.25–28, 2015.

HAKLAY, M. How good is volunteered geographical information? A comparative study of OpenStreetMap and Ordnance Survey datasets. **Environment and planning B: Planning and design**, [S.l.], v.37, n.4, p.682–703, 2010.

HANNUN, A. et al. Deep speech: Scaling up end-to-end speech recognition. **arXiv preprint arXiv:1412.5567**, [S.l.], 2014.

HAWORTH, B.; BRUCE, E.; IVESON, K. Spatio-temporal analysis of graffiti occurrence in an inner-city urban environment. **Applied Geography**, [S.l.], v.38, p.53–63, 2013.

HE, K.; ZHANG, X.; REN, S.; SUN, J. Deep residual learning for image recognition. In: IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, 2016. **Proceedings...** [S.l.: s.n.], 2016. p.770–778.

HENTSCHEL, M.; WAGNER, B. Autonomous robot navigation based on OpenStreetMap geodata. In: INTELLIGENT TRANSPORTATION SYSTEMS (ITSC), 2010 13TH INTERNATIONAL IEEE CONFERENCE ON, 2010. **Anais...** [S.l.: s.n.], 2010. p.1645–1650.

HU, Z.; ZHAO, D. Reinforcement learning for multi-agent patrol policy. In: COGNITIVE INFORMATICS (ICCI), 2010 9TH IEEE INTERNATIONAL CONFERENCE ON, 2010. **Anais...** [S.l.: s.n.], 2010. p.530–535.

HUISKES, M. J.; THOMEE, B.; LEW, M. S. New trends and ideas in visual concept detection: the MIR flickr retrieval evaluation initiative. In: MULTIMEDIA INFORMATION RETRIEVAL, 2010. **Proceedings...** [S.l.: s.n.], 2010. p.527–536.

JAIN, A. K.; LEE, J.-E.; JIN, R. Graffiti-id: Matching and retrieval of graffiti images. In: FIRST ACM WORKSHOP ON MULTIMEDIA IN FORENSICS, 2009. **Proceedings...** [S.l.: s.n.], 2009. p.1–6.

JOULIN, A.; MAATEN, L. van der; JABRI, A.; VASILACHE, N. Learning visual features from large weakly supervised data. In: EUROPEAN CONFERENCE ON COMPUTER VISION, 2016. **Anais...** [S.l.: s.n.], 2016. p.67–84.

KAWAMURA, T. Interpretação de um teste sob a visão epidemiológica: eficiência de um teste. **Arquivos Brasileiros de Cardiologia**, [S.l.], v.79, n.4, p.437–441, 2002.

LASSALA, G. Pichação não é pixação: uma introdução à análise de expressões gráficas urbanas. **São Paulo: Altamira Editorial**, [S.l.], 2010.

LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. **nature**, [S.l.], v.521, n.7553, p.436, 2015.

LECUN, Y.; BOTTOU, L.; BENGIO, Y.; HAFFNER, P. Gradient-based learning applied to document recognition. **Proceedings of the IEEE**, [S.l.], v.86, n.11, p.2278–2324, 1998.

LECUN, Y. et al. Backpropagation applied to handwritten zip code recognition. **Neural computation**, [S.l.], v.1, n.4, p.541–551, 1989.

LOWE, D. G. et al. Object recognition from local scale-invariant features. In: ICCV, 1999. **Anais...** [S.l.: s.n.], 1999. v.99, n.2, p.1150–1157.

MEGLER, V.; BANIS, D.; CHANG, H. Spatial analysis of graffiti in San Francisco. **Applied Geography**, [S.l.], v.54, p.63–73, 2014.

MILLER, G. **WordNet**: An electronic lexical database. [S.l.: s.n.], 1998.

O ESPAÇO E O TEMPO DO GRAFFITI E DA STREET ART, 2017. **Anais...** [S.l.: s.n.], 2017. v.34, p.1–16.

ODGERS, C. L. et al. Systematic social observation of children's neighborhoods using Google Street View: a reliable and cost-effective method. **Journal of Child Psychology and Psychiatry**, [S.l.], v.53, n.10, p.1009–1017, 2012.

OQUAB, M.; BOTTOU, L.; LAPTEV, I.; SIVIC, J. Learning and transferring mid-level image representations using convolutional neural networks. In: IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, 2014. **Proceedings...** [S.l.: s.n.], 2014. p.1717–1724.

PARRA, A.; BOUTIN, M.; DELP, E. J. Location-aware gang graffiti acquisition and browsing on a mobile device. **Proceedings of the IS&T/SPIE Electronic Imaging on Multimedia on Mobile Devices**, [S.l.], p.830402–1, 2012.

PARRA, A.; ZHAO, B.; KIM, J.; DELP, E. J. Recognition, segmentation and retrieval of gang graffiti images on a mobile device. In: TECHNOLOGIES FOR HOMELAND SECURITY (HST), 2013 IEEE INTERNATIONAL CONFERENCE ON, 2013. **Anais...** [S.l.: s.n.], 2013. p.178–183.

PORTUGAL, D.; ROCHA, R. A survey on multi-robot patrolling algorithms. In: DOCTORAL CONFERENCE ON COMPUTING, ELECTRICAL AND INDUSTRIAL SYSTEMS, 2011. **Anais...** [S.l.: s.n.], 2011. p.139–146.

PROCTOR, N. The Google Art Project: A new generation of museums on the web? **Curator: The Museum Journal**, [S.l.], v.54, n.2, p.215–221, 2011.

RUSSAKOVSKY, O. et al. ImageNet Large Scale Visual Recognition Challenge. **International Journal of Computer Vision**, [S.l.], v.115, n.3, p.211–252, 2015.

SAMPAIO, P. A.; SOUSA, R. D. S.; ROCHA, A. N. New Patrolling Strategies with Short-Range Perception. In: ROBOTICS SYMPOSIUM AND IV BRAZILIAN ROBOTICS SYMPOSIUM (LARS/SBR), 2016 XIII LATIN AMERICAN, 2016. **Anais...** [S.l.: s.n.], 2016. p.157–162.

SCHROFF, F.; KALENICHENKO, D.; PHILBIN, J. Facenet: A unified embedding for face recognition and clustering. In: IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, 2015. **Proceedings...** [S.l.: s.n.], 2015. p.815–823.

SHIN, H.-C. et al. Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. **IEEE transactions on medical imaging**, [S.l.], v.35, n.5, p.1285–1298, 2016.

SIMONYAN, K.; ZISSERMAN, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. **ImageNet Challenge**, [S.l.], p.1–10, 2014.

SONG, S.; XIAO, J. Deep sliding shapes for amodal 3D object detection in RGB-D images. In: IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, 2016. **Proceedings...** [S.l.: s.n.], 2016. p.808–816.

SRIVASTAVA, N. et al. Dropout: a simple way to prevent neural networks from overfitting. **Journal of machine learning research**, [S.l.], v.15, n.1, p.1929–1958, 2014.

SUPPORT, G. **Street View Service**. [Online; acessado em 19-Dezembro-2018], `https://developers.google.com/maps/documentation/javascript/streetview/`.

SZEGEDY, C. et al. Going deeper with convolutions. In: IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, 2015. **Proceedings...** [S.l.: s.n.], 2015. p.1–9.

TAIGMAN, Y.; YANG, M.; RANZATO, M.; WOLF, L. Deepface: Closing the gap to human-level performance in face verification. In: IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, 2014. **Proceedings...** [S.l.: s.n.], 2014. p.1701–1708.

THOMEE, B. et al. YFCC100M: The new data in multimedia research. **Communications of the ACM**, [S.l.], v.59, n.2, p.64–73, 2016.

TONG, W.; LEE, J.-E.; JIN, R.; JAIN, A. K. Gang and moniker identification by graffiti matching. In: ACM WORKSHOP ON MULTIMEDIA IN FORENSICS AND INTELLIGENCE, 3., 2011. **Proceedings...** [S.l.: s.n.], 2011. p.1–6.

VANDEVIVER, C. Applying google maps and google street view in criminological research. **Crime Science**, [S.l.], v.3, n.1, p.13, 2014.

WOOLDRIDGE, M. **An introduction to multiagent systems**. [S.l.]: John Wiley & Sons, 2009.

YANG, C.; WONG, P. C.; RIBARSKY, W.; FAN, J. Efficient graffiti image retrieval. In: ACM INTERNATIONAL CONFERENCE ON MULTIMEDIA RETRIEVAL, 2., 2012. **Proceedings...** [S.l.: s.n.], 2012. p.36.

YOSINSKI, J.; CLUNE, J.; BENGIO, Y.; LIPSON, H. How transferable are features in deep neural networks? In: ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS, 2014. **Anais...** [S.l.: s.n.], 2014. p.3320–3328.

YOSINSKI, J. et al. Understanding neural networks through deep visualization. **arXiv preprint arXiv:1506.06579**, [S.l.], 2015.

ZAMIR, A. R.; SHAH, M. Accurate image localization based on google maps street view. In: EUROPEAN CONFERENCE ON COMPUTER VISION, 2010. **Anais...** [S.l.: s.n.], 2010. p.255–268.

ZHANG, H.; MALCZEWSKI, J. Accuracy Evaluation of the Canadian OpenStreetMap Road Networks. , [S.l.], 2018.

ZHANG, X.; ZHAO, J.; LECUN, Y. Character-level convolutional networks for text classification. In: ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS, 2015. **Anais...** [S.l.: s.n.], 2015. p.649–657.

ANEXO A LINKS PARA DATASETS, MODELOS E APLICAÇÕES

Para esta dissertação foram utilizados imagens do Flickr e GSV para construção dos *datasets* usados no treinamento e testes, os quatro modelos modelos usados no MAS e a própria aplicação, todos disponibilizamos nos links abaixo.

A.1 Ambientes

A.1.1 Porto Alegre Centro Histórico - PCH

`https://github.com/glaucomunsberg/graphium/tree/master/data/datasets/city_porto_alegre_centro_historico`

A.1.2 Pelotas Região Metropolitana - PMe

`https://github.com/glaucomunsberg/graphium/tree/master/data/datasets/city_pelotas_full`

A.1.3 Pelotas Porto - PP

`https://github.com/glaucomunsberg/graphium/tree/master/data/datasets/city_pelotas_porto`

A.2 Datasets

Abaixo disponibilizamos os dois *datasets* construídos a partir do Flickr e GSV.

A.2.1 Flickr

`https://github.com/glaucomunsberg/kootstrap/tree/master/data/datasets/imagenet_flickr_full_cropped/`

Modelo	Nome do Arquivo
PRC	20170821191051.h5
PIRC	20170729205748.h5
PBC	20170802112022.h5
PRC+GSV	20190120203801.h5

Tabela 9 – Lista com os modelos e seus respectivos arquivos HDF5 no repositório.

A.2.2 Flickr usado nos modelos PRC/PIRC/PBC

`https://github.com/glaucomunsberg/kootstrap/tree/master/data/datasets/imagenet_flickr_cropped/`

A.2.3 Google Street View usado no modelo PRC+GSV

`https://github.com/glaucomunsberg/kootstrap/tree/master/data/datasets/imagenet_with_poa`

A.3 Modelos

Na Tabela 9 estão os quatro modelos para a identificação de grafite utilizadas nesta dissertação. Veja abaixo o link para cada um dos modelos.

A.3.1 Modelo PRC

`https://github.com/glaucomunsberg/graphium/blob/master/data/models/20170821191051.h5`

A.3.2 Modelo PIRC

`https://github.com/glaucomunsberg/graphium/blob/master/data/models/20170729205748.h5`

A.3.3 Modelo PBC

`https://github.com/glaucomunsberg/graphium/blob/master/data/models/20170802112022.h5`

A.3.4 Modelo PRC+GSV

`https://github.com/glaucomunsberg/graphium/blob/master/data/models/20190120203801.h5`

A.4 Aplicações

O código fonte e a documentação para o MAS desenvolvido neste trabalho estão disponíveis através do GitHub.

A.4.1 Sistema Multiagente

Código fonte da aplicação `https://github.com/glaucomunsberg/graphium/tree/master/applications/swarm`.

Documentação `https://glaucomunsberg.github.io/graphium/`

ANEXO B FLUXO DO TRABALHO

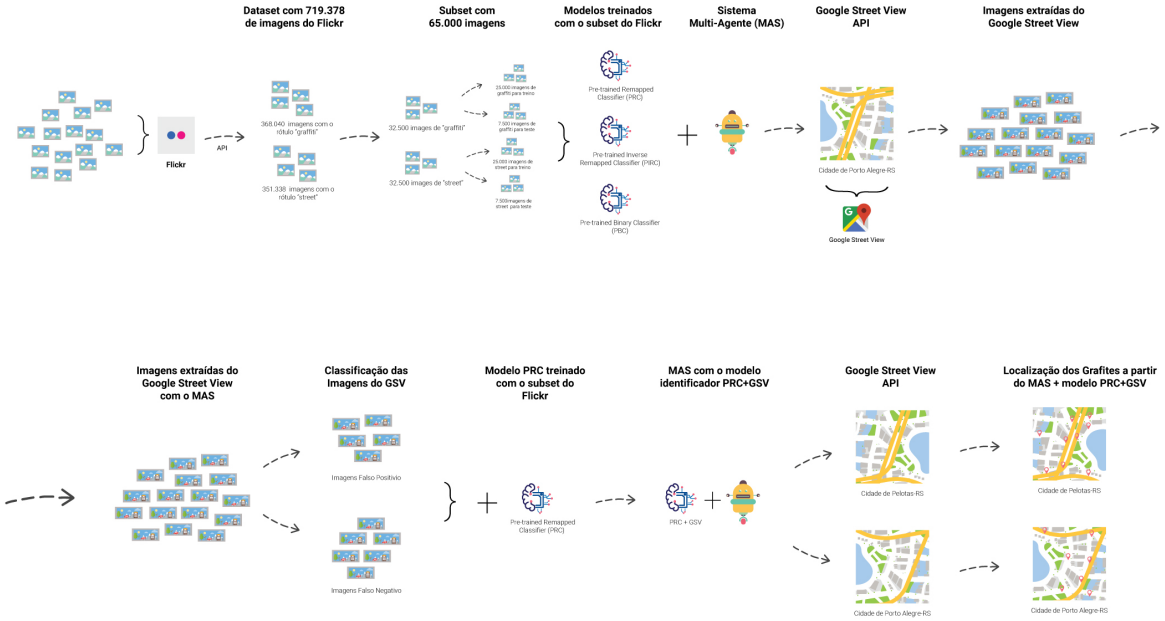


Figura 32 – A imagem ilustra o processo de coleta de imagens, treinamento dos modelos e identificação do grafite.

ANEXO C PSEUDO-CÓDIGO DO AGENTE

Abaixo descrevemos o pseudo-código do agente, o código completo poderá ser visto em https://github.com/glaucomunsberg/graphium/blob/master/applications/swarm/hive/Agent.py.

```
1  from threading import Thread
2
3  class Agent(Thread):
4
5      def run(self):
6
7          while get_street_no_visited() != 0:
8
9              street = get_street_no_visited()[0]
10
11             for node in street['nodes']:
12
13                 points = break_node_in_points(node)
14
15                 street_orientation = get_street_orientation(
                       node)
16
17                 heating_left = street_orientation - 90
18                 heating_right = street_orientation + 90
19
20                 for point in points:
21
22
23                     image_left = get_image_street_view(point['
                           lat'],point['lng'],heating_left)
```

```
24                      image_right = get_image_street_view(point['
                            lat'],point['lng'],heating_right)
25
26                      if identify_with_model(image_left):
27                          create_graffiti_on_map(point,
                                heating_left, image_left)
28
29
30                      if identify_with_model(image_rigth):
31                          create_graffiti_on_map(point,
                                heating_right, image_right)
32
33              set_street_visited(street)
```

ANEXO D IMAGENS DE GRAFITE E SUPORTES

Abaixo apresentamos uma série de imagens de grafite oriundas do Flickr e do Google Street View.

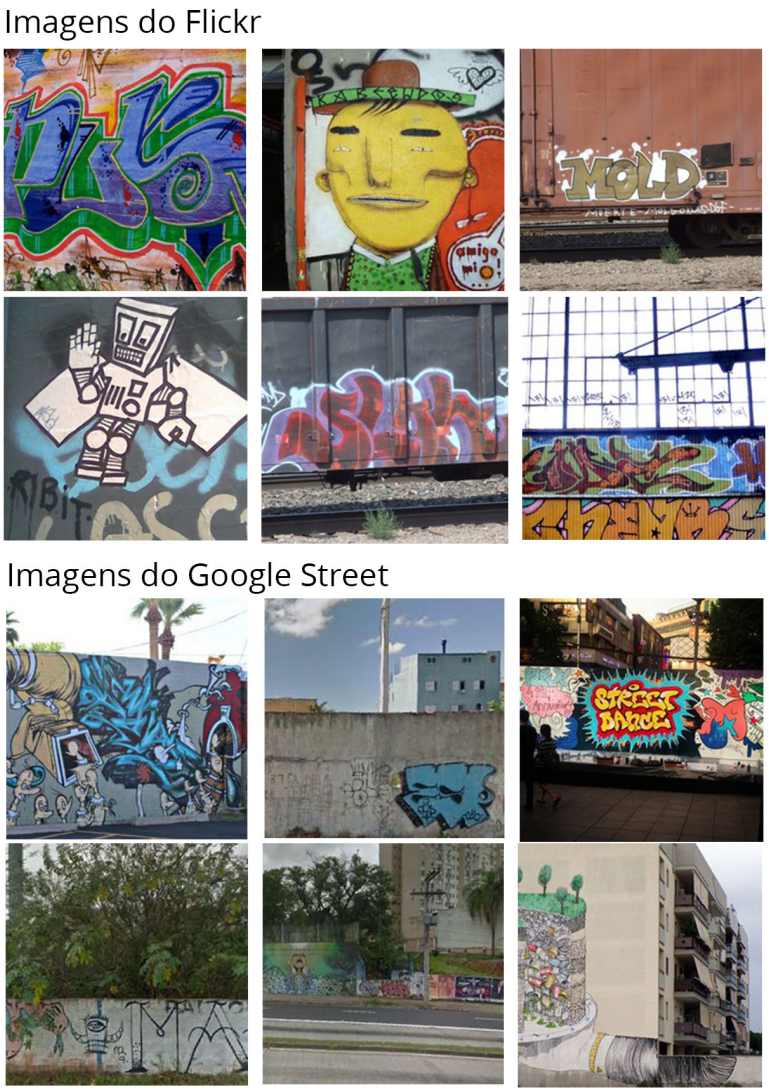


Figura 33 – Grafite em seus suportes com imagens do Flickr e GSV. Fonte: Flickr

Uma abordagem de extração de grafite com multi-agente e identificação por CNNs – Glauco Roberto Munsberg dos Santos

UNIVERSIDADE FEDERAL DE PELOTAS
Centro de Desenvolvimento Tecnológico
Programa de Pós-Graduação em Computação

Dissertação

Uma abordagem de extração de grafite com multiagente e identificação por CNNs

GLAUCO ROBERTO MUNSBERG DOS SANTOS

Pelotas, 2019