

## **Campus inteligente: uma proposta de segurança**

### **Smart Campus: a Security Proposal**

Jorge R. B. Garay<sup>1</sup>  
Isaac G. G. Vizcarra<sup>2</sup>  
Lincoln Marcellos<sup>3</sup>  
Moacyr Martucci<sup>4</sup>  
Sergio T. Kofuji<sup>5</sup>

**Resumo:** Por meio deste texto descreve-se a proposta de desenvolvimento de um sistema de detecção e reconhecimento facial a ser aplicado dentro do contexto de um *Campus* Inteligente. O sistema, ainda em fase de desenvolvimento, já se apresenta promissor em sua primeira versão e suas funcionalidades podem ser facilmente estendidas para outros cenários de aplicação. O sistema desenvolve-se como fruto de uma parceria entre a Huawei International e a Universidade de São Paulo.

**Palavras-chave:** Cidades Inteligentes; *Campus* Inteligente; Internet das Coisas; Segurança; Reconhecimento Facial.

**Abstract:** The text describes the proposal under development of a facial detection and recognition system to be applied within the context of an Intelligent Campus. The system, still under development, is already promising in its first version and its functionalities can easily be extended to other application scenarios. The system is developed as a result of a partnership between Huawei International and the University of São Paulo.

**Keywords:** Smart Cities; Smart *Campus*; Internet of Things; Security; Face Recognition.

### **Introdução**

A segurança é um fator determinante dentro de qualquer sistema computacional assim como na vida do indivíduo em uma sociedade. Situações de risco são observadas todos os dias e, como consequência, o número de soluções e pesquisas em segurança são cada vez maiores. A inovação tecnológica produz sistemas cada vez mais complexos e as cidades inteligentes encaixam-se dentro desse paradigma de múltiplas arquiteturas embarcadas e altamente heterogêneas,

<sup>1</sup> Universidade de São Paulo (USP), Departamento de Engenharia de Sistemas Eletrônicos, Av. Professor Luciano Gualberto, tr.3, n.158, São Paulo, SP. Universidade Mogi Das Cruzes (UMC), Núcleo de Pesquisas Tecnológicas, *Campus* Mogi das Cruzes, Av. Doutor Cândido Xavier de Almeida e Souza, 200, Mogi das Cruzes, Mogi das Cruzes, SP. E-mail: jorge@Isi.usp.br

<sup>2</sup> Universidade de São Paulo (USP), Departamento de Engenharia de Sistemas Eletrônicos, Av. Professor Luciano Gualberto, tr.3, n.158, São Paulo, SP. E-mail: isaac.gonzalez@pad.Isi.usp.br

<sup>3</sup> *Ibidem*. E-mail: lincoln.marcellos@pad.Isi.usp.br

<sup>4</sup> *Ibidem*. E-mail: mmartucc@usp.br

<sup>5</sup> *Ibidem*. E-mail: kofuji@pad.Isi.usp.br

em que a segurança gera grande interesse e abordagens mais desafiadoras são propostas (S. Srivastava *et. al.*, 2017) (Gharaibeh *et.al.*, 2017). No contexto das cidades inteligentes, a segurança surge como um desafio e uma meta, considerando a diversidade de dispositivos sensores e atuadores que interagem com o ambiente e disponibilizam suas informações na nuvem através da internet. Esse grande volume de dados fornece informações do contexto e do indivíduo, em alguns casos altamente sensíveis, e precisam ser protegidos, considerando-se todas as suas características (L. Cagliero *et al.*, 2015). A percepção do cidadão sobre o ambiente é o ponto chave no gerenciamento de uma cidade inteligente, já que todos os eventos e processos são direcionados para melhorar a qualidade de vida das pessoas. É difícil abordar a segurança dentro de um amplo cenário de uma Cidade Inteligente (A. Gharibeh, *et. al.*, 2016), sendo assim, acredita-se que seja necessária uma caracterização dos diversos ambientes e que seja feita uma abordagem individual e pormenorizada, com o intuito de maximizar a eficiência de qualquer modelo a ser proposto. Neste trabalho aborda-se a descrição de um contexto menor, um *campus* universitário inteligente, aberto não apenas para estudantes como, também, para o trânsito da comunidade em geral.

Por meio deste texto, têm-se por objetivo descrever de forma macro o desenvolvimento e a aplicação de um sistema de detecção e reconhecimento a ser implantado dentro de uma cidade universitária, atendendo conceitos de cidades e *campus* inteligentes; no contexto de um *campus* inteligente e considerando a integração de uma infraestrutura de sensores e sistemas multicâmeras com capacidade de detecção e reconhecimento facial. O sistema hospeda-se dentro de uma infraestrutura em nuvem, conhecida como “fiware”, sobre a qual são realizados o processamento e a análise dos dados em tempo real.

### **Campus inteligente**

Em um *campus* universitário inteligente é observado o mesmo comportamento de uma cidade inteligente (K. Zhang *et. al.*, 2017), porém, possivelmente com um número de incidências menor. No *campus* inteligente o estudante é o centro de quase todos os eventos. No entanto, descreve-se aqui um *campus* universitário aberto, no qual pessoas alheias à comunidade acadêmica

podem circular, fazendo com que eventos comuns a uma cidade sejam trasladados a esse ambiente, como por exemplo roubos, acidentes de carro, circulação por lugares restritos e outras atividades irregulares.

### **Descrição do cenário**

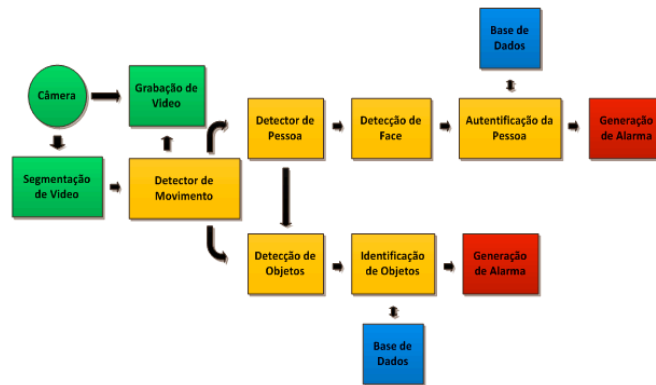
O cenário a ser descrito corresponde a uma cidade universitária com acesso livre para a comunidade externa. A cidade universitária, denominada daqui em diante de *Campus* Inteligente, possui uma área de 7.443.770 m<sup>2</sup> e conta com aproximadamente 30 unidades de ensino, entre faculdades, institutos e centros de pesquisa universidade-empresa. Encontra-se resguardada por um conjunto efetivo denominado de guarda universitária, formada por um agrupamento orientado a manter a segurança do patrimônio institucional e que, como auxílio, utiliza-se de estrutura de monitoramento com aproximadamente duas mil câmeras, apenas para observação do *campus*. Sobre essas condições, é descrita a utilização de um sistema de monitoramento, detecção e reconhecimento em tempo real. O sistema apresenta uma estrutura modular, o que permite a sua integração parcial com sistemas já existentes. Sobre esse cenário deverá operar o sistema proposto, utilizando-se da atual infraestrutura de câmeras e integrando-se uma plataforma de sensores (movimento, luminosidade e som), que deverá atuar de forma coesa em todos os processos do sistema, com o intuito de maximizar a precisão do sistema de detecção e reconhecimento.

### **Visão da arquitetura**

A arquitetura é modular e representada em cinco módulos, alguns desses independentes entre si. A descrição das funcionalidades dos módulos (Figura 1) que compõem a arquitetura de um sistema de segurança, proposto para o contexto de *Campus* Inteligente, inicia-se com as especificações iniciais do sistema, baseadas em características de soluções existentes e acrescentando-se atuais desafios para *Security Cities*, como escalabilidade, processamento e geração de alertas, aplicando-se inteligência aos sistemas de reconhecimento facial, entre outras. Nesse ponto, são inúmeros os desafios, não apenas com uma simples especificação do

sistema, mas com os algoritmos e técnicas de reconhecimento que permitam abstrair a complexidade do sistema e o grande volume de dados gerados.

**Figura 1:** Fluxo básico de operação do sistema.



A arquitetura é representada basicamente por cinco módulos, descritos a seguir:

- a) *Módulo de Registro*: o acesso é atribuído ao usuário com privilégios de administrador. Nesse módulo é realizado o cadastro das pessoas, informações básicas e captura de foto. As informações são armazenadas no servidor e acessíveis para o sistema de reconhecimento.
- b) *Módulo de Detecção*: nesse módulo é recepcionado o fluxo de *streamings* de vídeo proveniente do servidor de mídias, disponibilizando os mesmos para os algoritmos de detecção. Imediatamente depois, a sequência de vídeo (*streamings* de interesse) é armazenada em *buffer* para simular a previsão de detecção.
  - Detecção de movimento: função que é executada como evento de pré-processamento de detecção de pessoas. Caso o movimento seja detectado, o sistema prossegue com a *pipeline* de processamento. Caso contrário, nenhum processamento extra é efetuado. Nesse último processo, obtém-se o *frame* de vídeo referente ao momento da detecção de movimento, o alerta para sistema de processamento e o *trigger* para sistema de armazenamento.
  - A detecção de pessoas é utilizada como pré-processamento para o sistema de alerta de invasão. A detecção de faces é utilizada como entrada para o sistema de autenticação/classificação. Seu objetivo é detectar as faces e pré-processar as imagens.

- Na geração de alertas, o sistema gera alertas de acordo com o tipo de evento ocorrido (detecção de pessoa, face, movimento, entre outros) e os envia para o centro de monitoramento. No envio de alertas, a informação é encapsulada e enviado um alerta para o sistema de controle de mídia (servidor de aplicação).
- c) *Módulo de Autenticação*: neste módulo analisam-se as imagens oriundas do sistema de câmeras quando existir alguma face na imagem. Quando a face é detectada, aplicam-se filtros para ressaltar padrões e, imediatamente depois, é realizado o processamento de imagem para identificar a quem pertencem os rostos detectados. O processo ocorre por meio de uma rápida análise comparativa das faces capturadas com o banco de dados de cadastros preexistentes. Finalmente, a informação é encapsulada e encaminhada para o sistema de controle de mídias, a partir do que é publicada a identidade do indivíduo como conhecido ou desconhecido, no caso de não existir na base de dados.
- d) *Módulo de Controle e Armazenamento de Mídia*: neste módulo é criada a estrutura de *pipeline* para conexão com os componentes do Kurento, necessários para o processamento, exibição e armazenamento da mídia. Origina-se a *pipeline* contendo módulo de ponto de mídia referente à câmera inicializada. Disponibiliza-se um ponto de conexão com módulos de outros recursos (WebRTC, processamento, armazenamento, *streaming*, entre outros).
- Inicializam-se os componentes de processamento de imagens contendo os algoritmos de detecção. Adiciona-se o módulo à *pipeline*. *Pipeline* de vídeo para inserção do módulo de detecção e metadados de configuração, como: tipo de algoritmo a ser utilizado para as detecções, tempo de *bufferização*, taxa de *frames* desejada, largura de banda máxima alocada. *Pipeline* contendo módulo de detecção. Disponibiliza-se um ponto de conexão com módulos de outros recursos (WebRTC, armazenamento, streaming, entre outros).
  - Inicializa-se o sistema de armazenamento. Inicializam-se os módulos de armazenamento de vídeo. Adiciona-se o módulo à *pipeline*. Inicializa-se o módulo de *streaming* WebRTC. Permite-se transmitir o fluxo de vídeo proveniente da *pipeline* associada. Inicializa-se o módulo de *streaming*

WebRTC. Permite-se transmitir o fluxo de vídeo proveniente da *pipeline* associada.

- e) *Módulo de Monitoramento e Controle*: neste módulo inicia-se o *WebSocket* para conexão dos usuários com o sistema de controle de mídia. Validam-se usuário e senha. Identificam-se perfil e suas permissões no sistema. Permite-se que o usuário inicialize o fluxo de mídia de uma câmera IP. O usuário pode configurar os metadados da câmera, como localização e tipo de filtro a ser utilizado (tipo de algoritmo para detecção). Permite-se que o usuário elimine uma câmera, já inicializada no sistema, e libere os recursos de *hardware*. Permite-se que o usuário elimine todas as câmeras alocadas no sistema. Libera-se todo o recurso de *hardware* utilizado pelo serviço. Permite-se que o usuário selecione as câmeras que deseja visualizar o vídeo.

### **O sistema de processamento e *Deep Learning***

O uso de técnicas de *Deep Learning*, associadas ao serviço de processamento analítico de dados dos sensores (*Recognition*), aumenta a eficácia do sistema para identificação de pessoas em ambientes não controlados e reduz a quantidade de dados armazenados por identificação para tratamento do *Big Data*, como o processo de rastreamento entre as câmeras. Isso permite escalar a capacidade de identificação de milhares de pessoas, como no caso encontrado em grandes cidades com grande movimentação, como ocorre dentro do *Campus Inteligente*. Assim, contribui-se para criar um modelo escalável para Cidades Inteligentes e para o conceito de *Security Cities*.

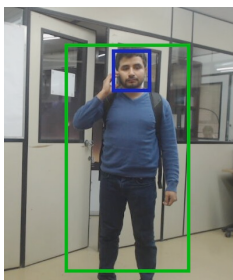
É preciso salientar que, dentro do conceito de Cidades Inteligentes, todas as áreas necessitam ter informações dos seus processos e, com isso, tem-se a necessidade de monitoração. Essa monitoração ocorre por meio de pequenos sensores, com *hardware* limitado, que normalmente realiza uma única função, por exemplos: monitorar temperatura, movimento, ruído, luminosidade e outras características que sejam necessárias para compreender o ambiente. A disposição dos sensores é realizada em redes locais, para que seja possível caracterizar uma determinada área para análise de uma medida específica. Dentro desta especificação, sensores de presença são integrados ao sistema em operação.

## Casos de uso

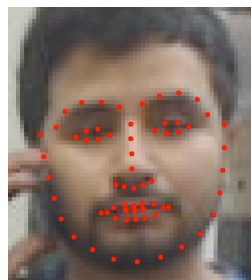
O sistema de segurança com funções de detecção e reconhecimento poderá atuar tanto em ambientes controlados (ambientes com limitações de acesso) quanto em ambientes abertos, dentro da extensão do *Campus* Inteligente, inserindo-se como um sistema inteligente de monitoramento e segurança. Inicialmente são descritas duas aplicações e cenários para o sistema proposto.

- a) *Detecção de pessoas*: no cenário da universidade faz-se necessário o controle de acesso a determinados ambientes restritos, sendo nesses ambientes que o sistema alcança sua maior extensão de funcionalidades. O sistema inicia-se com o pré-cadastro de usuários no sistema. O sistema detecta pessoas nos ambientes controlados (Figura 2a) e imediatamente acessa a base de dados para identificação do indivíduo. Quando o indivíduo não existe na base de dados, o sistema gera um alerta e armazena a imagem da detecção (Figura 2b). A mesma imagem pode ser utilizada em uma próxima identificação do mesmo indivíduo, sendo possível fazer um acompanhamento dos lugares onde ele esteve e onde se encontra. Isso ocorre porque, no processo, são extraídas características únicas da face de cada indivíduo e acessadas posteriormente, utilizando-se técnicas de *Deep Learning*, e um processo comparativo ocorre por meio de *machine learning* (J. Li *et al.*, 2017). Para o reconhecimento de objetos, são usadas algumas “Redes Neurais Convolucionais” (W. Rawat e Z. Wang, 2017). Para isso, a imagem é dividida em várias regiões e introduzida em uma rede neural que vai predizer em que regiões da imagem é possível encontrar objetos, indicando a probabilidade de classificação (Figura 2c).

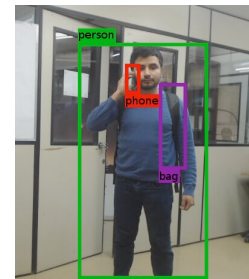
- b) **Figura 2**: Descrição do sistema de detecção e reconhecimento.



**Figura 2a:** Detecção de pessoa e face.

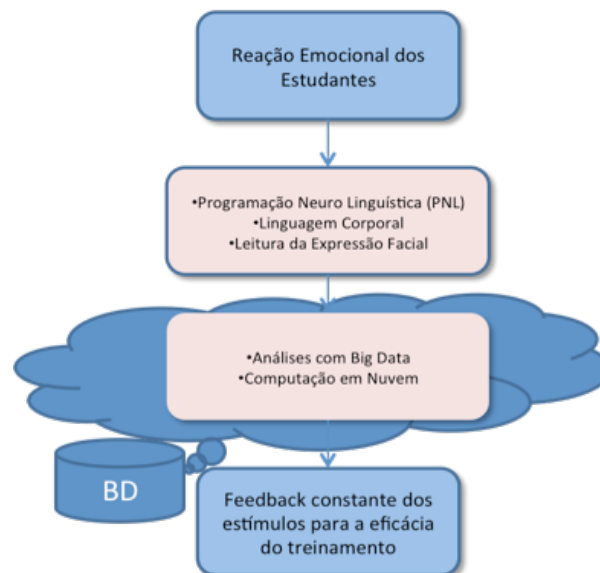


**Figura 2b:** Reconhecimento de face.



**Figura 2c:** Detecção e identificação de objetos.

- c) *Detecção de estilos comportamentais*: o sistema, em um segundo nível, ao detectar um indivíduo que não seja reconhecido na base de dados, de acordo com a captura gerada em acompanhamento dos lugares onde esteve, possibilita a análise comportamental desse indivíduo. A partir da digitalização do processo analógico capturado, é possível obter um grande volume de dados complexos que possam ser verificados via soluções de *analytics* de *Big Data*, vinculados ao poder computacional do ambiente de computação em nuvem, seguindo o princípio da computação em grade. A coleta desses dados em tempo possibilita a inferência por parâmetros comportamentais propostos pela Programação Neurolinguística (PNL) e por outros referenciais, desde postura, rotina de respiração, tensão muscular, relaxamento muscular, comportamento físico, até mesmo mudanças cor da pele, os quais permitem a análise dos dados em nível acurado. A detecção do estilo comportamental do indivíduo anônimo pode prever algumas condutas potenciais, partindo do reconhecimento do nível de empatia, ansiedade e características corporais que no momento sejam diagnosticadas.
- d) **Figura 3**: Modelagem do processo de interação e reconhecimento de emoções.

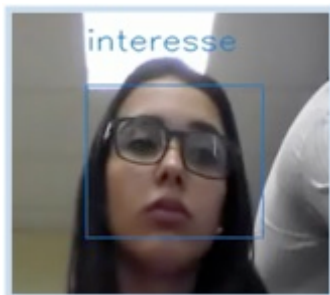


A partir do momento em que o indivíduo não for identificado na base de dados, o sistema gera um alerta e armazena a imagem da detecção. Nesse instante, a imagem capturada pelas câmeras é processada por intermédio de algoritmo de leitura de face e de linguagem corporal (Figura 3), utilizando os referenciais da

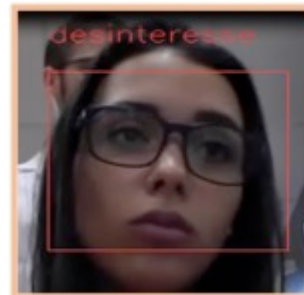


programação neurolinguística e do catálogo de emoções (GUR, Ruben C. *et al.*, 2002). O sistema promove a comparação da imagem com um banco de dados previamente definido e reconhece o estado emocional do indivíduo naquele momento (RICHARDSON, D. C. *et al.*, 2004), podendo, até mesmo, identificar a alternância de estados emocionais do indivíduo no período em que estiver em trânsito dentro do *campus* (Figura 4).

**Figura 4:** Sistema de detecção e reconhecimento de emoções. Aplicação em aula.



**Figura 4a:** Detecção de estado emocional.



**Figura 4b:** Detecção de estado emocional alternado.

## Conclusão

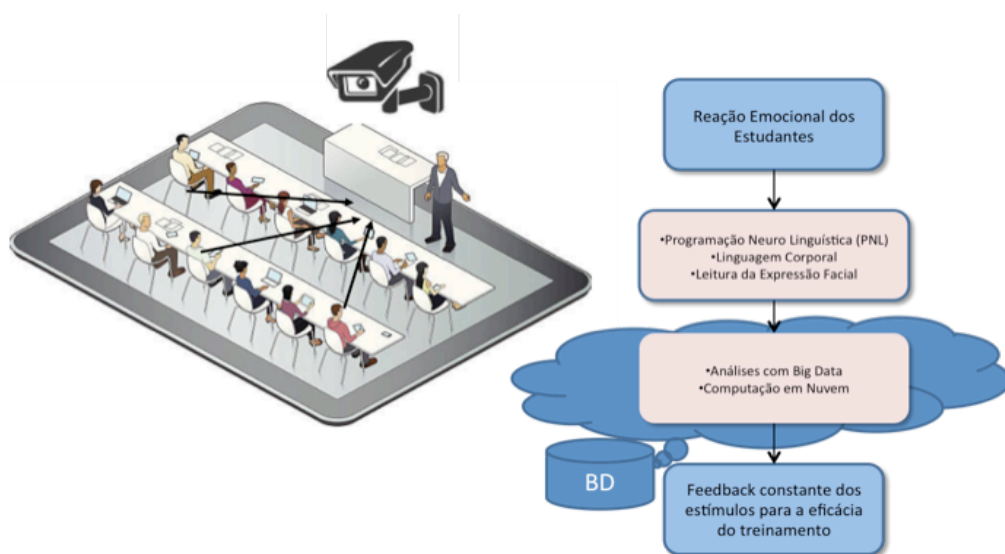
A proposta em andamento já possui a sua primeira versão em operação e testes dentro dos ambientes controlados pelo LSI-USP (Laboratório de Sistemas Integráveis – Universidade de São Paulo). No contexto do *Campus* Inteligente, é gerado um grande volume de dados sobre os quais é necessário inferir com o maior número de técnicas para o gerenciamento de dados e para garantir consistência, interoperabilidade, granularidade e reutilização dos dados gerados pelo IoT em cidades inteligentes. No sistema de detecção, reconhecimento e monitoramento, um fator importante é o ciclo de vida dos dados. O tempo é uma característica acoplada com todas as camadas transversais de segurança, privacidade de dados e infraestrutura de suporte.

No desenvolvimento do sistema proposto, o módulo de pré-processamento é resultado de uma especificação inicial do sistema e do levantamento dos principais requisitos funcionais e não funcionais identificados. Módulos de cache de vídeo temporário, módulo de detecção de eventos e registro (operando na nuvem), processamento analítico dos sensores, interfaces de usuário, serviços de controle de

administração, seleção dos algoritmos e treinamento da estrutura *Deep Learning* e a implantação do protótipo em ambiente controlado.

Embora em versão inicial, o sistema apresenta-se promissor e busca ser aplicado em cenários cada vez mais extensos e abertos, expandindo e maximizando cada uma das suas funcionalidades. Novas funcionalidades devem ser incluídas, os algoritmos de detecção e reconhecimento devem ser otimizados e, como consequência, uma nova versão deve ser gerada. Acredita-se que o sistema pode ser utilizado não apenas para monitoramento de ambientes, como também para reconhecimento de veículos e rastreamento, como ferramenta de acompanhamento do aprendizado em sala de aula (Figura 5) (caso de uso de reconhecimento de emoções) e em muitos outros cenários de aplicação. O sistema em fase de teste é financiado por uma parceria entre a Huawei International e a Universidade de São Paulo.

**Figura 5:** Aplicação em sala de aula.



## Referências

- CAGLIERO, L. *et al.* Monitoring the Citizens' Perception on Urban Security in Smart City Environments. In: IEEE INTERNATIONAL CONFERENCE ON DATA ENGINEERING WORKSHOPS, 31<sup>th</sup>, 2015, Seoul. **Anais**, Seoul, 2015, p.112-116.
- GHARAIBEH; SALAHUDDIN, M. A.; HUSSINI, S. J.; KHREISHAH, A.; KHALIL, I.; GUIZANI, M.; AL-FUQAHA, A. Smart Cities: A Survey on Data Management, Security and Enabling Technologies. In: IEEE **Communications Surveys & Tutorials**, vol.PP, n.99, p.1-1, 2017.

GHARIBEH, A.; SALAHUDDIN, M. A.; HUSSINI, S. J.; KHREISHAH, A.; KHALIL, I.; GUIZANI, M.; AL-FUQAHA, A. Smart Cities: A Survey on Data Management, Security and Enabling Technologies. In: IEEE **Communications Surveys & Tutorials**, vol.PP, no.99, p.1-1, 2016.

GUR, Ruben C. *et al.* A Method for Obtaining 3-Dimensional Facial Expressions and its Standardization for Use in Neurocognitive Studies. In: **Journal of Neuroscience Methods**, v.115, n.2, p.137-143, 2002.

LI, J. *et al.* Safety Helmet Wearing Detection Based on Image Processing and Machine Learning. In: 2017 INTERNATIONAL CONFERENCE ON ADVANCED COMPUTATIONAL INTELLIGENCE (ICACI), 9<sup>th</sup>, 2017, Doha. **Anais**, Doha, 2017, pp.201-205.

RAWAT, W.; WANG, Z. Deep Convolutional Neural Networks for Image Classification: A Comprehensive Review. In: **Neural Computation**, v.29, n.9, p.2352-2449, Sept. 2017.

RICHARDSON, D. C.; SPIVEY, Michael J. Eye Tracking: Research Areas and Applications. In: **Encyclopedia of Biomaterials and Biomedical Engineering**, p.573-582, 2004.

SRIVASTAVA, S.; BISHT, A.; NARAYAN, N. Safety and Security in Smart Cities Using Artificial Intelligence: a Review. In: INTERNATIONAL CONFERENCE ON CLOUD COMPUTING, DATA SCIENCE & ENGINEERING - CONFLUENCE, 7<sup>th</sup>, 2017, Noida. **Anais**, Noida: 2017, p.130-133.

ZHANG, K.; NI, J.; YANG, K.; LIANG, X.; REN, J.; SHEN, X. S. Security and Privacy in Smart City Applications: Challenges and Solutions. In: IEEE **Communications Magazine**, vol. 55, n.1, p.122-129, Jan. 2017.