



# FICE

8ª A FEIRA DE INICIAÇÃO  
CIENTÍFICA E EXTENSÃO

05 E 06 DE SETEMBRO

## APRENDIZADO DE CARACTERÍSTICAS PARA PROBLEMAS DE CLASSIFICAÇÃO DE UMA CLASSE APLICADOS À IMAGENS

*Arthur Zago<sup>1</sup>; Manassés Ribeiro<sup>2</sup>*

### 1. INTRODUÇÃO

A detecção de anomalias é um tema recorrente de pesquisa em Reconhecimento de Padrão (RP). Neste tipo de problema apenas as amostras de uma são totalmente conhecidas (classe dos *normais*). Para as demais amostras suas classes são totalmente desconhecidas (classes dos *anormais*). A maneira mais comum de abordar o problema de detecção de anomalias é por meio da Classificação de Uma Classe (CUC). A ideia geral da CUC é agrupar as amostras conhecidas e definir uma fronteira de decisão nas extremidades destes agrupamentos, de modo que as amostras conhecidas estejam localizados dentro desta fronteira, enquanto as demais estejam localizadas fora desta. A classificação se dá pela fronteira de decisão que é o limiar de classificação (RIBEIRO, 2018). A Figura 1 mostra um exemplo didático do problema de CUC, onde as amostras pertencentes ao conjunto normal  $N$  estão localizadas dentro da hiperesfera, o círculo que representa a fronteira de decisão (em vermelho) e as amostras  $o_1$ ,  $o_2$  e  $o_3$ , pertencentes ao conjuntos dos anormais, estão fora do círculo.

No geral, o desempenho de classificação em sistemas de RP são bastante dependentes da forma como as amostras são representadas no modelo. A representação do padrão acontece por meio de características (informações relevantes) que são extraídas das amostras. Em especial nos problemas de CUC a representação do padrão é bastante relevante, pois como só são conhecidos os dados da classe normal, espera-se que as características extraídas destas amostras possam representá-las de maneira mais coesa possível (RIBEIRO, 2018).

---

<sup>1</sup> Aluno do Instituto Federal Catarinense, Campus Videira. Curso técnico em Informática. E-mail: arthurzago284@gmail.com

<sup>2</sup> Professor Orientador do Instituto Federal Catarinense, Campus Videira. Curso técnico em Informática. Email: manasses.ribeiro@ifc.edu.br



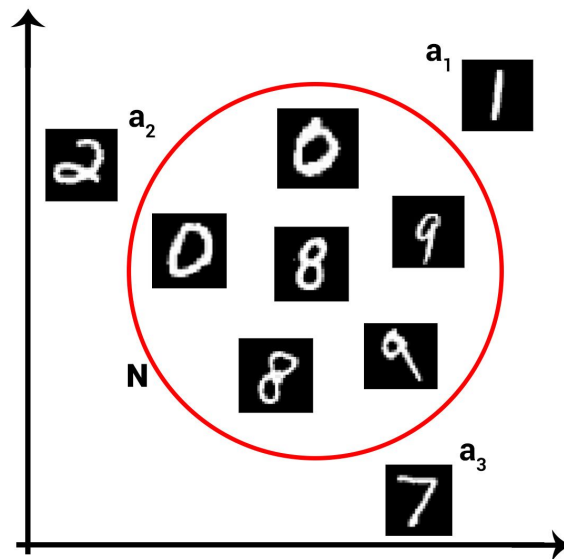
# FICE

8ª A FEIRA DE INICIAÇÃO  
CIENTÍFICA E EXTENSÃO

05 E 06 DE SETEMBRO

Tradicionalmente o processo de extração de características acontece utilizando abordagens feitas à mão (*hand-crafted*), que são métodos desenvolvidos para fins específicos. Entretanto, recentemente a representação do padrão vem sendo abordada por meio de modelos orientados a dados (*data-driven*). Estes modelos têm-se utilizados da robustez dos modelos de aprendizado profundo (também conhecidos em inglês como *deep learning*), que são um tipo especial de redes neurais artificiais com diversas camadas, e da surpreendente capacidade de processamento alcançada pelos computadores atuais.

Figura 1 - Exemplo didático do problema de CUC.



Fonte: O Autor

Atualmente, os modelos de aprendizado profundo têm alcançado estado da arte para a resolução de diversos problemas, especialmente relacionados à detecção de objetos em imagens, processamento de vídeos e detecção de anomalias (RIBEIRO et al., 2018; SUN et al., 2017 e ERFANI et al., 2016). Neste trabalho, o foco é o estudo dos Autocodificadores (AC), que são um tipo de rede neural artificial de aprendizado profundo otimizada por meio de treinamento não-supervisionado. O AC possui dois conjuntos de camadas, onde uma é a camada de codificação e outra é a camada de decodificação, assim como a camada latente intermediária (*bottleneck*). A principal característica do AC é sua capacidade de ao mesmo tempo aprender as características dos objetos (amostras) e reduzir a



# FICE

8ª A FEIRA DE INICIAÇÃO  
CIENTÍFICA E EXTENSÃO

05 E 06 DE SETEMBRO

dimensionalidade dos dados. Assim, as características aprendidas, já em baixa dimensionalidade, estarão representadas na camada latente do AC (HASAN et al., 2016; XU et al., 2015 e MASCI et al., 2011).

O AC é útil para modelar problemas de CUC, uma vez que seu treinamento não depende de rótulos para os dados de entrada, da mesma forma que não depende da contra-classe, que são necessárias para treinamentos de métodos supervisionados. Portanto, o problema que se pretende abordar neste trabalho consiste em utilizar o AC como ferramenta para realizar o aprendizado de características em problemas de classificação de uma classe. A principal hipótese de trabalho é que os métodos de aprendizado profundo, em especial o AC, podem ser uma ferramenta útil para o aprendizado de características e que podem ser aplicados eficientemente em problemas de CUC.

### 1.1. Objetivo Geral

Estudar o uso dos autocodificadores no aprendizado de características para problemas de classificação de uma classe em imagens.

### 1.2. Objetivos Específicos

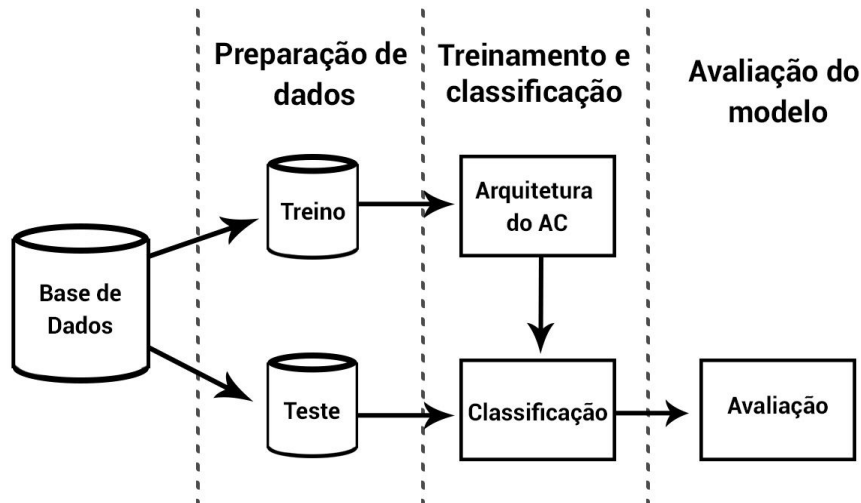
1. verificar um meio de utilizar o erro de reconstrução do autocodificador com pontuação para classificação de eventos anormais;
2. sugerir uma forma para definir o limiar de classificação para problemas de classificação de uma classe;
3. validar experimentalmente as abordagens sugeridas em uma base de dados disponível publicamente.

## 2. METODOLOGIA

O método sugerido para contornar o problema proposto é baseado nos trabalhos de Ribeiro et al. (2018) e Hasan et al. (2016), e é dividido em etapas que seguem desde a preparação dos dados até a avaliação do desempenho de classificação. A visão geral do método é apresentada na Figura 2.



Figura 2 - Visão geral do método proposto



Fonte: O Autor

## 2.1. Preparação de dados

A primeira etapa do método é a preparação dos dados, que consiste em converter as imagens para escala de cinza, redimensionar o tamanho das imagens, padronizar a escala dos dados em valores de mínimos e máximos e preparar os conjuntos de treinamento e teste. Como a entrada do AC é unidimensional de tamanho fixo, todas as imagens do conjunto de dados são redimensionadas para o tamanho 28x28 pixels e em escala de cinza. Os valores de cinza dos pixels são também escalados no intervalo [0..1], padronizando assim os valores de entrada de acordo com os valores de saída do AC (ver Seção 2.2). Por fim, são separados os conjuntos de treinamento e de teste. Para o primeiro são utilizados apenas amostras consideradas normais e sem rotulação. Para o conjunto de teste, amostras de ambas as classes, normais e anormais, devem estar presentes, assim como a rotulação direta de todas as amostras.

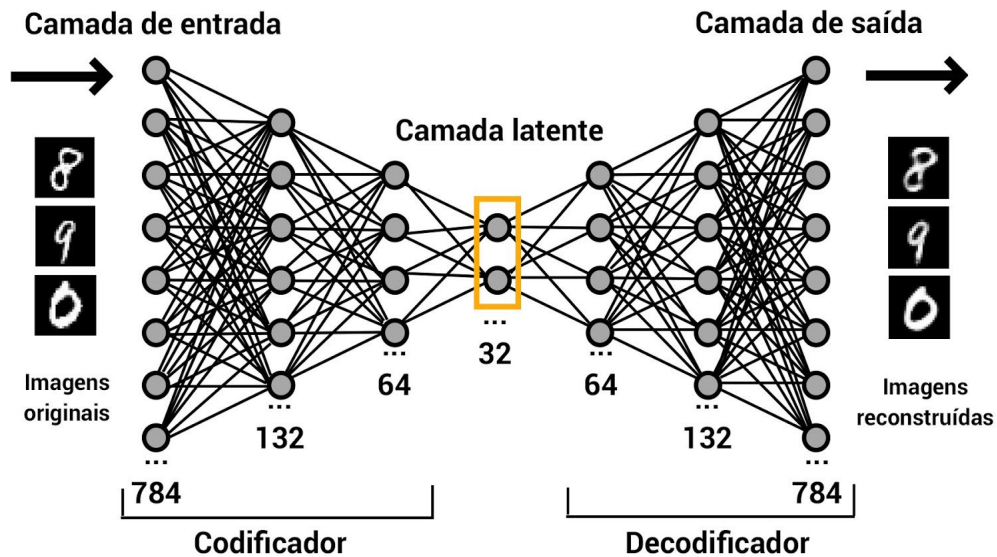
## 2.2. Treinamento e classificação

A arquitetura sugerida neste trabalho é de um AC profundo composto de seis camadas (três camadas de codificação e três camadas de decodificação), mais a camada latente (*bottleneck*). Todas as camadas são totalmente conectadas (*fully-connected*), o que significa que todos os neurônios da camada imediatamente anterior são ligados com todos os neurônios da camada imediatamente posterior.



Ambas as camadas de entrada e saída são compostas de 784 neurônios, enquanto as camadas internas são compostas de 128 e 64 neurônios, respectivamente, e a camada latente é composta por 32 neurônios. A arquitetura do AC escolhida para este trabalho pode ser visualizada graficamente na Figura 3.

Figura 3 - Arquitetura do AC profundo



Fonte: O Autor

O modelo é otimizado considerando o Erro Médio Quadrático (EMQ) entre os valores da camada de saída e a camada de entrada do AC. Conforme mostrado na Figura 3, as imagens originais são inseridas no AC pela camada de entrada e propagam pela rede neural do AC, da esquerda para a direita, ou seja, primeiro passando pelas camadas do *codificador* até atingir a camada latente, que representam as características da imagem de entrada em dimensão reduzida. Na sequência, estes valores da camada latente são propagadas pelas camadas do *decodificador* até atingir a camada de saída do AC.

O EMQ, por sua vez, é calculado considerando os valores reconstruídos na camada de saída e os valores originais dados como entrada na camada de entrada, que determinará o grau de aprendizado do modelo. Quanto menor for o EMQ, mais otimizado o modelo está, pois mais fiel será a reconstrução da imagem de saída em relação a imagem de entrada.



# FICE

8ª A FEIRA DE INICIAÇÃO  
CIENTÍFICA E EXTENSÃO

05 E 06 DE SETEMBRO

Após a otimização do modelo vem a etapa da classificação, que acontece apenas as amostras do conjunto de teste. Inspirado no que foi proposto por Hasan et al. (2016), neste trabalho é utilizado o próprio EMQ como pontuação para a classificação das amostras. A hipótese é que sendo uma amostra do tipo normal ela será *conhecida* do modelo e, portanto, o EMQ será menor. Por outro lado, para as amostras do tipo anormal, que não são conhecidas do modelo, é esperado que apresentem um EMQ maior. Assim, o próprio EMQ pode ser utilizado como um valor de classificação entre as amostras normais e anormais. A pontuação para classificação é obtida pelo EMQ Padronizado (EMQP), que é o EMQ padronizado no intervalo  $[0..1]$ , sendo o valor  $0$  para as amostras totalmente *normais* e  $1$  para as amostras totalmente *anormais*.

Para a classificação é necessário que seja definido um limiar de classificação a partir do qual as amostras podem ser rotuladas como normais ou anormais. Neste trabalho é proposto utilizar uma abordagem à *posteriori* que considera o EER que pode ser acessado por meio da curva ROC (ver Seção 2.3). Caso o EMQP seja maior do que o limiar de classificação, a amostra é classificada como anormal, caso contrário, a amostra é classificada como normal.

### 2.3. Avaliação do Modelo

Concluindo o processo de classificação é necessário avaliar o desempenho geral de classificação do modelo. Para tanto, é utilizada a matriz de confusão e calculado ambas as Taxas de Verdadeiro Positivos (TVP), também conhecido por *sensibilidade*, e de Verdadeiro Negativos (TVN), ou *especificidade*. O TVP avalia a taxa de acertos das amostras consideradas normais, enquanto a TVN avalia a taxa de acertos das amostras consideradas anormais. A ideia é que estas duas medidas sejam mais parecidas possível e também o mais próximo de 1.0. Sendo próximo os valores de TVP e TVN, significa que o classificador conseguiu classificar de maneira equilibrada as amostras de ambas as classes, sem priorizar uma das classe em desfavor de outra.

Com base na TVP e TVN é possível construir a curva de característica de operação do receptor (em inglês, *Receiver Operating Characteristic curve* ou ROC), calcular a área abaixo da curva ROC (em inglês, AUC) e com isto acessar o EER



# FICE

8<sup>A</sup> A FEIRA DE INICIAÇÃO  
CIENTÍFICA E EXTENSÃO

05 E 06 DE SETEMBRO

(em inglês, *Equal error rate*), ponto onde a taxa de falsa aceitação é igual a taxa de falsa rejeição.

### 3. EXPERIMENTOS, RESULTADOS E DISCUSSÃO

A hipótese de trabalho que se pretende explorar é de que métodos de aprendizado profundo, em especial o AC, podem ser uma ferramenta útil para o aprendizado de características em problemas de CUC. Como estudo de caso para a abordagem sugerida neste trabalho é utilizado o conjunto de imagens MNIST, que é um conjunto de dados disponível publicamente.

#### 3.1. Conjunto de dados MNIST

Para o estudo de caso deste trabalho foi utilizado o conjunto de dados MNIST (LECUN et al., 1998), que é composto por imagens dos 10 algarismos numéricos. Este conjunto de dados contém 60.000 imagens de treinamento e 10.000 de teste, e as imagens possuem um tamanho padrão de 28x28 pixels. Como o problema abordado neste trabalho se refere a CUC, foi necessário realizar a diferenciação entre as amostras, separando-as entre as amostras consideradas normais e as consideradas anormais. Assim, para o conjunto de treinamento (dos normais) foram utilizadas apenas as amostras que representam os algarismos “8”, “9” e “0”, que foram escolhidos pela similaridade de formas entre eles. Por outro lado, para o conjunto de teste optou-se por manter o conjunto original contendo todas as amostras, realizando, entretanto, a adequação de rotulação, ou seja, as amostras dos algarismos “8”, “9” e “0” foram rotuladas como *normais*, enquanto as demais como *anormais*. Por fim, o conjunto de treinamento ficou composto por 17.723 amostras e o conjunto de teste por 10.000 amostras, sendo 2.963 normais e 7.037 de anormais.

#### 3.2. Parâmetros de treinamento

O treinamento foi realizado por um total de  $50 + 1$  épocas, uma vez que a partir deste ponto o modelo convergiu e não apresentou ganho significativo, como pode ser observado na Figura 4. O método de otimização utilizado foi o Adam e o



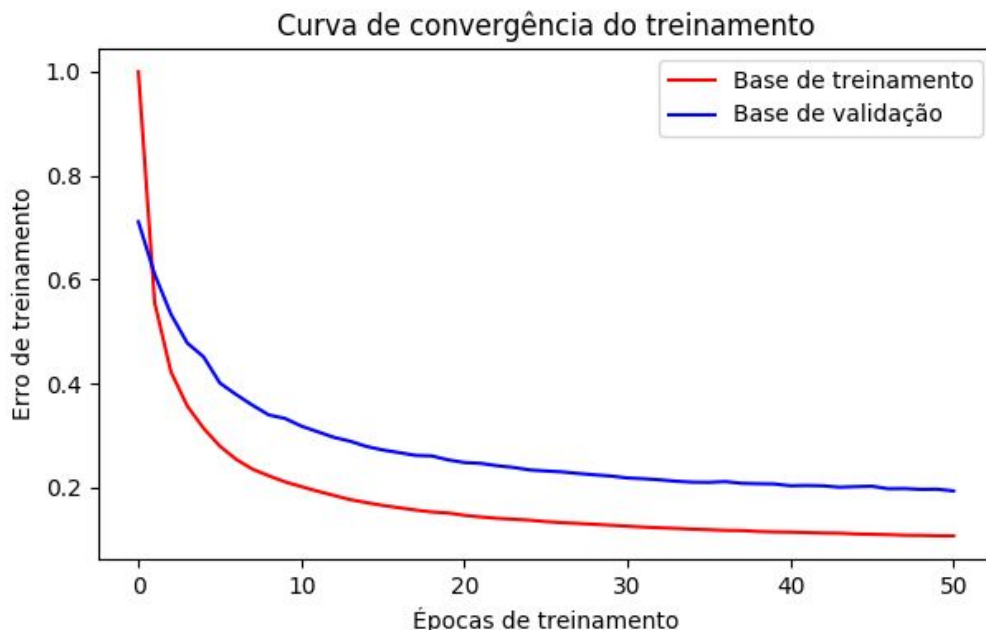
# FICE

8ª A FEIRA DE INICIAÇÃO  
CIENTÍFICA E EXTENSÃO

05 E 06 DE SETEMBRO

conjunto de treinamento foi dividido em lotes com tamanho de 256 imagens por lote. Durante o treinamento uma parte dos dados é utilizado para validação, que é importante para verificar se o modelo está otimizando sem que ocorra o sobreajustamento. O sobreajustamento ocorre quando o modelo fica muito ajustado para os dados de treinamento mas apresenta desempenho ruim na classificação do teste. A Figura 4 mostra a curva de convergência do treinamento tanto para a base de treinamento quanto para a base de validação. Na curva de treinamento (vermelha) é possível verificar que na época 50 o modelo atingiu a convergência e não é mais necessário estender o treinamento além deste ponto. Na curva de validação (azul) é possível verificar que o modelo foi otimizado de maneira adequada, uma vez que acompanhou a curva de treinamento em direção ao erro mínimo.

**Figura 4 - Curvas de convergência no treinamento do AC**



Fonte: O Autor

### 3.3. Experimentos

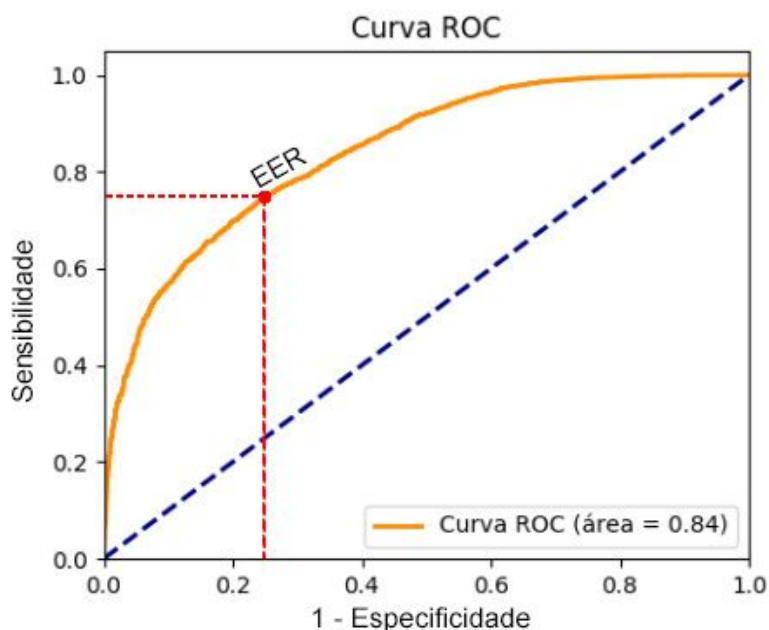
O experimento realizado têm o objetivo de verificar o desempenho de classificação considerando o aprendizado das características utilizando o AC, bem como a definição do limiar de classificação à *posteriori*. A abordagem sugerida utilizando o EER, que é acessado por meio da curva ROC (ver Subseção 2.3), mostrou-se adequada para o conjunto de dados estudado, uma vez que foi





observado equilíbrio de classificação para ambas as classes (considerando o TVP e TVN). A Figura 5 apresenta o EER, que foi o limiar de classificação escolhido para o experimento. Os resultados da classificação obtidos, considerando o EER, foram 2.253, 710, 1728 e 5.309, para verdadeiro positivo, falso positivo, falso negativo e verdadeiro negativo, respectivamente. Já os valores para TVP e TVN foram 0,7603 e 0,7544, respectivamente.

**Figura 5 - Limiar de classificação EER acessado pela curva ROC**



Fonte: O Autor

#### 4. CONSIDERAÇÕES FINAIS

Neste trabalho foram apresentados métodos para o aprendizado de características aplicados ao problema de classificação de uma classe com supervisão limitada, em especial o AC. Primeiro foi revisitada a literatura sobre o aprendizado de características, bem como os problemas de classificação de uma classe. Na sequência, com base na literatura, foi sugerido uma forma de abordar o problema de classificação para estes problemas utilizando o aprendizado de características por meio do AC. Também foi sugerido uma forma de utilizar o erro de reconstrução do AC como pontuação para a classificação das amostras discriminando-as entre normais e anormais. Uma questão importante em problemas



# FICE

8ª A FEIRA DE INICIAÇÃO  
CIENTÍFICA E EXTENSÃO

05 E 06 DE SETEMBRO

de classificação de uma classe é a definição do limiar de classificação. Neste sentido, com base na literatura, uma abordagem utilizando o EER acessado pela curva ROC foi sugerida. Por fim, experimentos utilizando o conjunto de dados MNIST foram realizados e os resultados foram analisados e discutidos.

Como possíveis trabalhos futuros, pode-se apontar a exploração de métodos diferentes de aprendizado profundo, como o AC convolucional, além das redes LSTM (*Long Short Term Memory*) para o aprendizado de características espaço-temporal. Além da utilização de novos métodos da aprendizagem profunda, pode-se abordar problemas de classificação de uma classe para detecção de anomalias em vídeos.

## AGRADECIMENTOS

O autor A. S. Zago agradece ao Instituto Federal Catarinense e o CNPq pela bolsa de estudos. Todos os autores agradecem à NVIDIA pela doação da GPU para este trabalho.

## REFERÊNCIAS

- ERFANI, S. M. et al. **High-dimensional and large-scale anomaly detection using a linear one-class SVM with deep learning**. *Pattern Recognition*, v. 58, p. 121–134, 2016.
- HASAN, M. et al. **Learning temporal regularity in video sequences**. In: *Conference on Computer Vision and Pattern Recognition*. Piscataway, NJ: IEEE, 2016. p. 733–742.
- LECUN, Y. et al. **Gradient-based learning applied to document recognition**. *Proceedings of the IEEE*, v. 86, n. 11, p. 2278–2324, Nov 1998.
- MASCI, J. et al. **Stacked convolutional auto-encoders for hierarchical feature extraction**. p. 54-55, 2011.
- RIBEIRO, M. **Deep Learning Methods for Detecting Anomalies in Videos: Theoretical and Methodological Contributions**. Thesis (Doctorate) — Federal University of Technology - Paraná, Curitiba, PR, Brazil, Abr 2018. Graduate Program in Electrical and Computer Engineering.
- RIBEIRO, M.; LAZZARETTI, A. E.; LOPES, H. S. **A study of deep convolutional auto-encoders for anomaly detections in video**. *Pattern Recognition Letters*, v. 105, p. 13-22, 2018. *Machine Learning and applications in Artificial Intelligence*.
- SUN, Q.; LIU, H.; HARADA, T. **Online growing neural gas for anomaly detection on changing surveillance scenes**. *Pattern Recognition*, v. 64, Supplement C, p. 187-201, 2017.
- XU, D. et al. **Learning deep representations of appearance and motion for anomalous event detection**. In: XIE, M. W. J. X.; TAM, G. K. L. (Ed.). *British Machine Vision Conference*. Swansea, UK: BMVA Press, 2015. p. 8.1-8.12.