III LATIN AMERICAN WORKSHOP ON COMPUTATIONAL NEUROSCIENCE
(LAWCN 2021)

Towards Loop Closure Detection for SLAM Applications using Bag of Visual Features: Experiments and Simulation

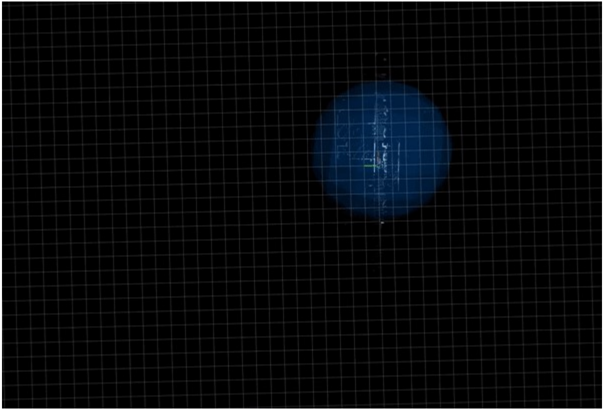
Alexandra Raibolt
raibolt@ime.eb.br

Gustavo Casqueiro
gustavo.casqueiro@ime.eb.br

Paulo Rosa
rpaulo@ime.eb.br

Alberto Angonese
aangonese@faeterj-petropolis.edu.br

Loop Closure Detection (LCD)



https://www.youtube.com/watch?v=OV6wNr62nqQ

Loop Closure Detection (LCD)

LCD problem are based on:

- Bag of Visual Features (BoVF).

Not fully explored in the LCD problem:

- Deep Learning (DL).

Objective

Integration of a hybrid model of Artificial Neural Network (ANN) architecture — the Long Term Recurrent Convolutional Network (CNN-LSTM) with a vision-based Simultaneous Localization and Mapping (SLAM) system, capable of solving Loop Closure Detection (LCD) problem.

Proposed Approach

- Evaluation of a Bag of Visual Features (BoVF) approach;
- Extracting features through the Local Feature Descriptors and Local Binary Descriptors;
- Recognition and classification tasks on visual datasets through Multilayer Perceptron (MLP):
  - M NIST, JAFFE, Extended CK+, FEI, CIFAR-10, and FER-2013.

Proposed Approach

We expect that the experiment and the preliminary simulations lead us in a right choice of a Descriptor that will be addressed in future work that consists of reformulation of Convolutional Filters through a Descriptor that constitutes the hybrid model of Artificial Neural Network (ANN) architecture.

Comparative Evaluation of Feature Descriptors

Local Binary Descriptors:

- BRIEF
- ORB
- BRISK
- FREAK
- AKAZE

Local Feature Descriptors:

- SIFT
- SURF
- KAZE

Comparative Evaluation of Feature Descriptors

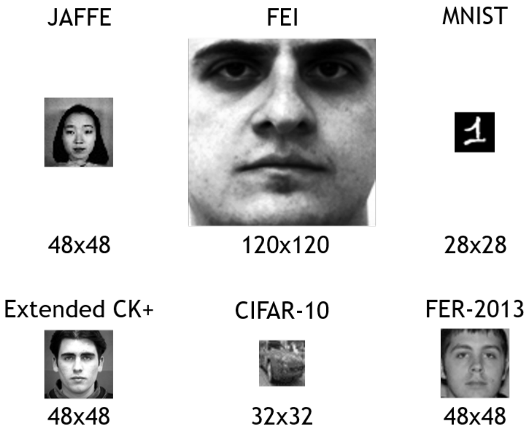


Figure: Size (pixels) of the visual datasets used in these experiments.

Training Step

Table: Processing time (min.) in the Feature Representation step of the Bag of Visual Feature (BoVF) approach of the training step.

Feature Representation step					
Algorithms	***Dataset***	**Visual Datasets**			
		MNIST	**Extended CK+**	**CIFAR-10**	**FER-2013**
SIFT	**Training-set**	02:37	01:01	02:59	02:51
SURF		00:49	**00:16**	00:54	**00:25**
KAZE		02:59	01:02	02:59	02:52
ORB		**00:32**	00:17	**00:46**	00:39
BRISK		-	00:29	00:57	01:00

Training Step

Table: Processing time (min.) in the Visual Vocabulary Generation step of the Bag of Visual Feature (BoVF) approach of the training step.

Visual Vocabulary Generation step					
Algorithms	***Dataset***	**Visual Datasets**			
		MNIST	**Extended CK+**	**CIFAR-10**	**FER-2013**
SIFT	**Training-set**	02:30	00:48	05:14	02:50
SURF		**00:01**	**00:05**	00:03	00:20
KAZE		01:45	00:39	05:09	04:13
ORB		02:39	01:46	03:22	02:43
BRISK		-	00:06	**00:00**	**00:06**

Training Step

Table: Processing time (min.) in the Image Representation step of the Bag of Visual Feature approach of the training step.

Image Representation step					
Algorithms	**Dataset**	**Visual Datasets**			
		MNIST	**Extended CK+**	**CIFAR-10**	**FER-2013**
SIFT	**Training-set**	01:04	00:11	00:56	00:33
SURF		**00:13**	**00:10**	00:20	00:29
KAZE		01:04	00:11	00:56	00:33
ORB		01:02	00:12	00:55	00:33
BRISK		-	**00:10**	**00:01**	**00:19**

Test Step

Table: Processing time (min.) in the Feature Representation and Image Representation steps of the Bag of Visual Feature approach of the test step.

Feature Representation and Image Representation steps					
Algorithms	**Dataset**	**Visual Datasets**			
		MNIST	**Extended CK+**	**CIFAR-10**	**FER-2013**
SIFT	**Test-set**	00:35	00:04	00:46	00:25
SURF		**00:09**	**00:01**	**00:08**	**00:06**
KAZE		00:39	00:04	00:46	00:25
ORB		00:16	**00:01**	00:20	00:09
BRISK		-	00:02	00:11	00:09

Test Step

Table: Accuracy rate (%) in the test steps of the BRIEF, AKAZE and FREAK descriptors in the FEI visual dataset.

Algorithms	**Dataset**	**Multilayer Perceptron Model**					
		MLP1	**MLP2**	**MLP3**	**MLP4**	**MLP5**	**MLP6**
BRIEF	**FEI**	0.78	0.74	0.77	0.76	0.82	0.85
AKAZE		**0.85**	**0.87**	**0.83**	**0.84**	**0.83**	**0.86**
FREAK		0.47	0.47	0.47	0.51	0.51	0.54

Test Step

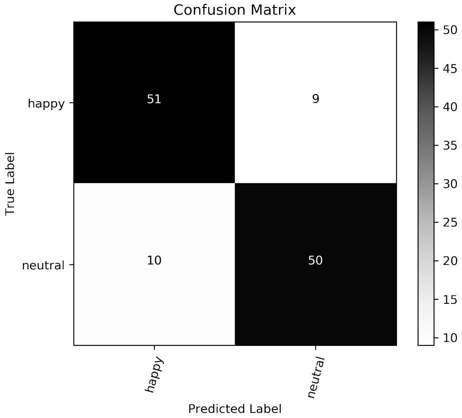


Figure: Confusion Matrix over FEI visual dataset with BRISK descriptor with MLP5 model.

Reformulation of Convolutional Filters through a Local Binary Descriptor

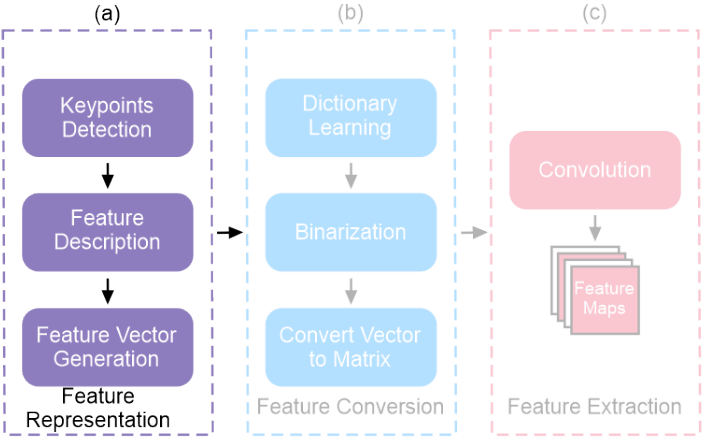


Figure: Flowchart of the proposed reformulation.

Reformulation of Convolutional Filters through a Local Binary Descriptor

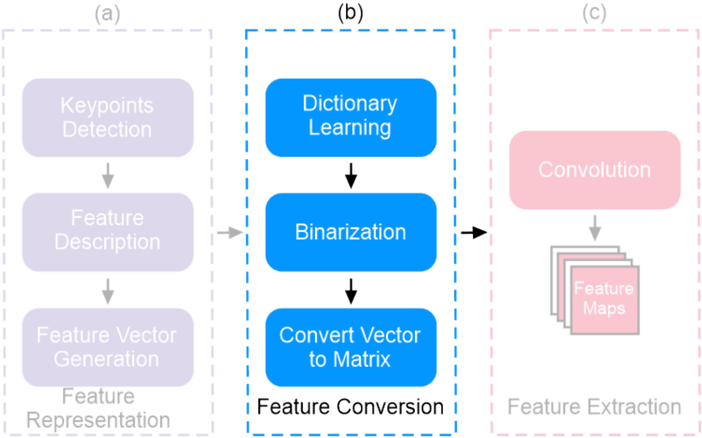


Figure: Flowchart of the proposed reformulation.

Reformulation of Convolutional Filters through a Local Binary Descriptor

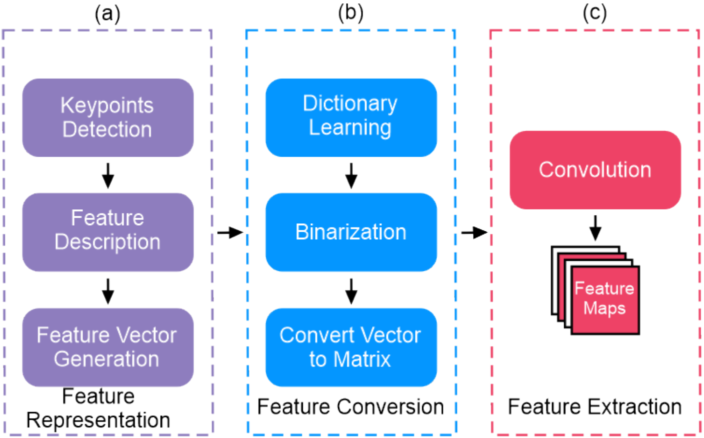


Figure: Flowchart of the proposed reformulation.

Behavior of the convolution operation

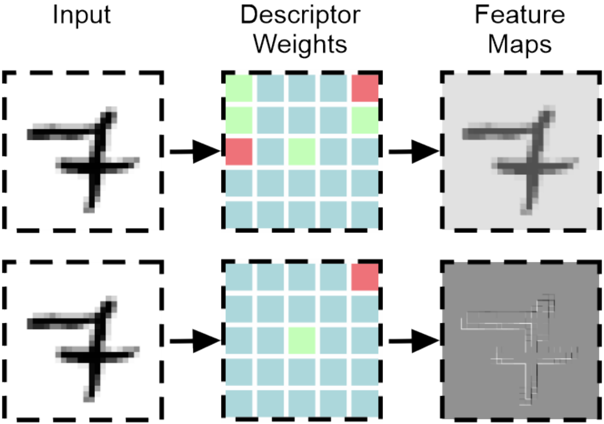


Figure: Behavior of the convolution operation.

Conclusion

The descriptor to be addressed in the reformulation of Convolutional Filters of a traditional Convolutional Neural Network through a Local Binary Descriptor of a hybrid model of Artificial Neural Network architecture was defined and implemented in this work, and now we can move towards the next steps, that consists in feature extraction and image classification tasks with the proposed methods.

Our approach is promising, where we hope in the next steps of this work to demonstrate that the proposed methods reduce the computational complexity and have the potential to perform the Loop Closure Detection problem for a simultaneous localization and mapping system.

Acknowledgments

CAPES

PROCAD-DEFESA
PROGRAMA DE COOPERAÇÃO
ACADÊMICA EM DEFESA NACIONAL

III LATIN AMERICAN WORKSHOP ON COMPUTATIONAL NEUROSCIENCE
(LAWCN 2021)

Towards Loop Closure Detection for SLAM Applications using Bag of Visual Features: Experiments and Simulation

Alexandra Raibolt
raibolt@ime.eb.br

Gustavo Casqueiro
gustavo.casqueiro@ime.eb.br

Paulo Rosa
rpaulo@ime.eb.br

Alberto Angonese
aangonese@faeterj-petropolis.edu.br