From Impressionism to Expressionism: Automatically Identifying Van Gogh's Paintings

Guilherme Folego, Otavio Gomes, Anderson Rocha

RECOD Lab, Institute of Computing,
University of Campinas (Unicamp), Brazil

Introduction

The main research question we wanted to address was:
Given a painting of interest, was it portrayed by Vincent van Gogh?

Introduction

This is an interesting and important question because the authorship of a painting influences its artistic, social, historic, and monetary values [1].

[1] G. E. Newman and P. Bloom, “Art and authenticity: The importance of originals in judgments of value.,” *Journal of Experimental Psychology*, vol. 141, no. 3, pp. 558, 2012.

Introduction

This is an interesting and important question because the authorship of a painting influences its artistic, social, historic, and monetary values [1].

Art specialists often employ methods that are potentially invasive, and, consequently, may interfere with the painting [2].

[1] G. E. Newman and P. Bloom, “Art and authenticity: The importance of originals in judgments of value.,” *Journal of Experimental Psychology*, vol. 141, no. 3, pp. 558, 2012.

[2] J. Ragai, “The scientific detection of forgery in paintings,” *Proceedings of the American Philosophical Society*, vol. 157, no. 2, pp. 164–175, 2013.

Introduction

This is an interesting and important question because the authorship of a painting influences its artistic, social, historic, and monetary values [1].

Art specialists often employ methods that are potentially invasive, and, consequently, may interfere with the painting [2].

Machine learning and image processing techniques can help!

[1] G. E. Newman and P. Bloom, “Art and authenticity: The importance of originals in judgments of value.,” *Journal of Experimental Psychology*, vol. 141, no. 3, pp. 558, 2012.

[2] J. Ragai, “The scientific detection of forgery in paintings,” *Proceedings of the American Philosophical Society*, vol. 157, no. 2, pp. 164–175, 2013.

Related work

Johnson et al. [6] investigated attribution for van Gogh's paintings. Approaches: wavelet, Hidden Markov Model, Support Vector Machine, and Multidimensional Scaling.

[6] C. R. Johnson, E. Hendriks, I. J. Berezhnoy, E. Brevdo, S. M. Hughes, I. Daubechies, J. Li, E. Postma, and J. Z. Wang, "Image processing for artist identification," *IEEE Signal Processing Magazine*, vol. 25, no. 4, pp. 37–48, 2008.

Related work

Johnson et al. [6] investigated attribution for van Gogh's paintings. Approaches: wavelet, Hidden Markov Model, Support Vector Machine, and Multidimensional Scaling.

Li et al. [5] performed statistical hypothesis testing for distinguishing van Gogh from his contemporaries based on brush stroke analysis.

[6] C. R. Johnson, E. Hendriks, I. J. Berezhnoy, E. Brevdo, S. M. Hughes, I. Daubechies, J. Li, E. Postma, and J. Z. Wang, "Image processing for artist identification," *IEEE Signal Processing Magazine*, vol. 25, no. 4, pp. 37–48, 2008.

[5] J. Li, L. Yao, E. Hendriks, and J. Z. Wang, "Rhythmic brushstrokes distinguish van Gogh from his contemporaries: Findings via automated brushstroke extraction," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 6, pp. 1159–1176, 2012.

Related work

Khan et al. [11] used state-of-the-art image processing techniques to classify paintings within 91 artists.

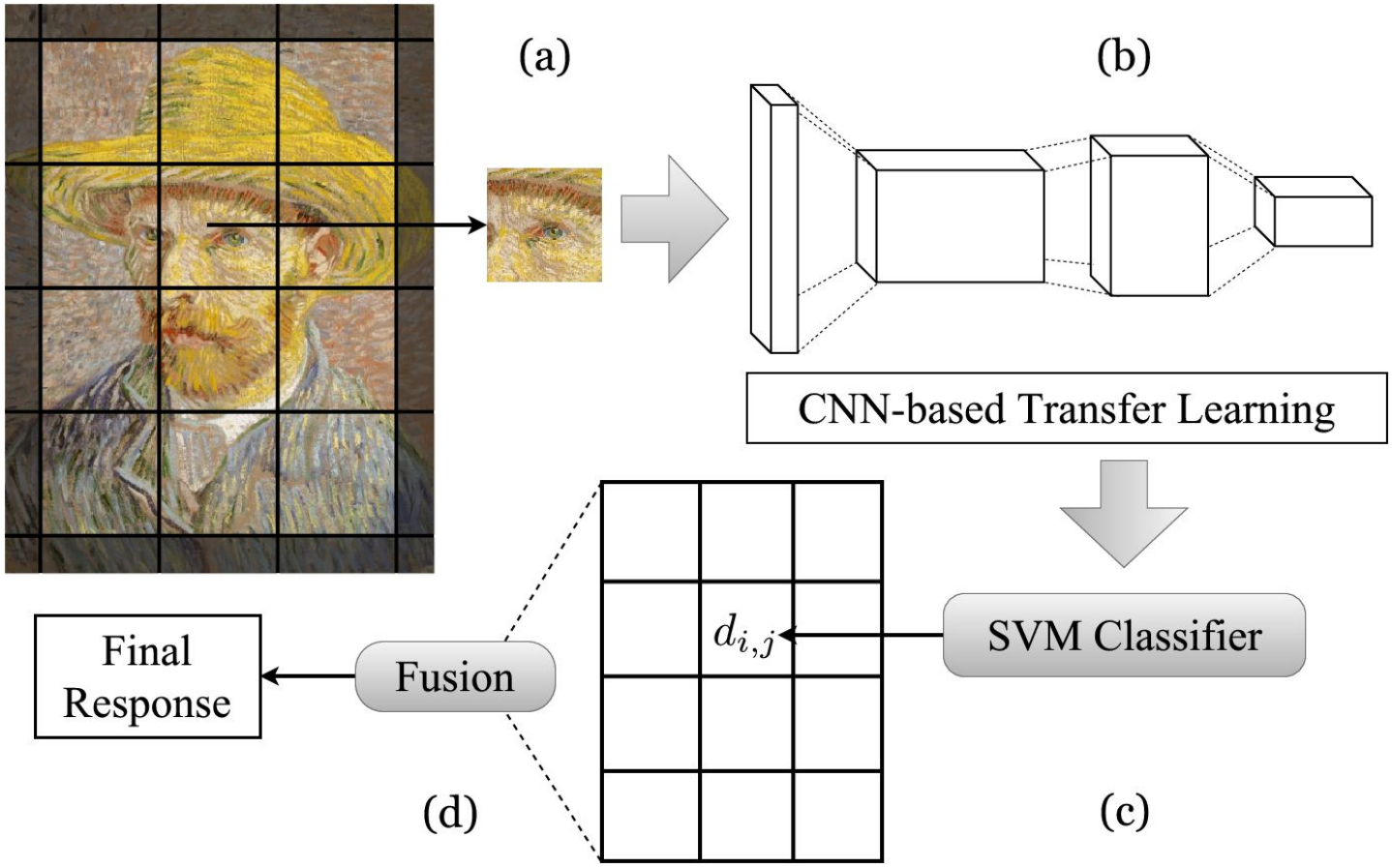
To the best of our knowledge, this work provided the only public dataset for painting identification to this day, although without standardization with respect to image density (in Pixels per Inch).

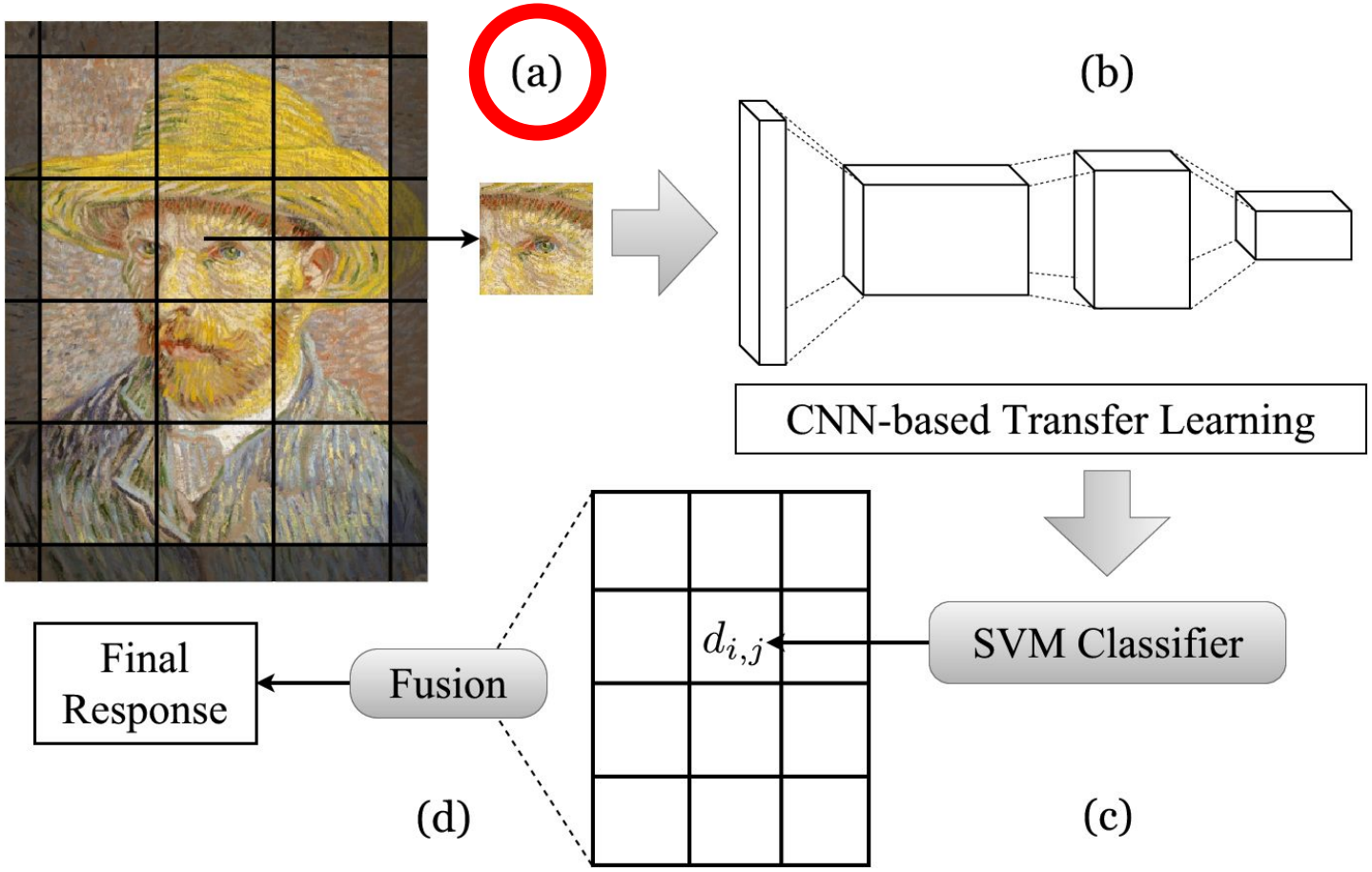
[11] F. S. Khan, S. Beigpour, J. Weijer, and M. Felsberg, “Painting-91: A large scale database for computational painting categorization,” *Machine Vision and Applications*, vol. 25, no. 6, pp. 1385–1397, 2014.

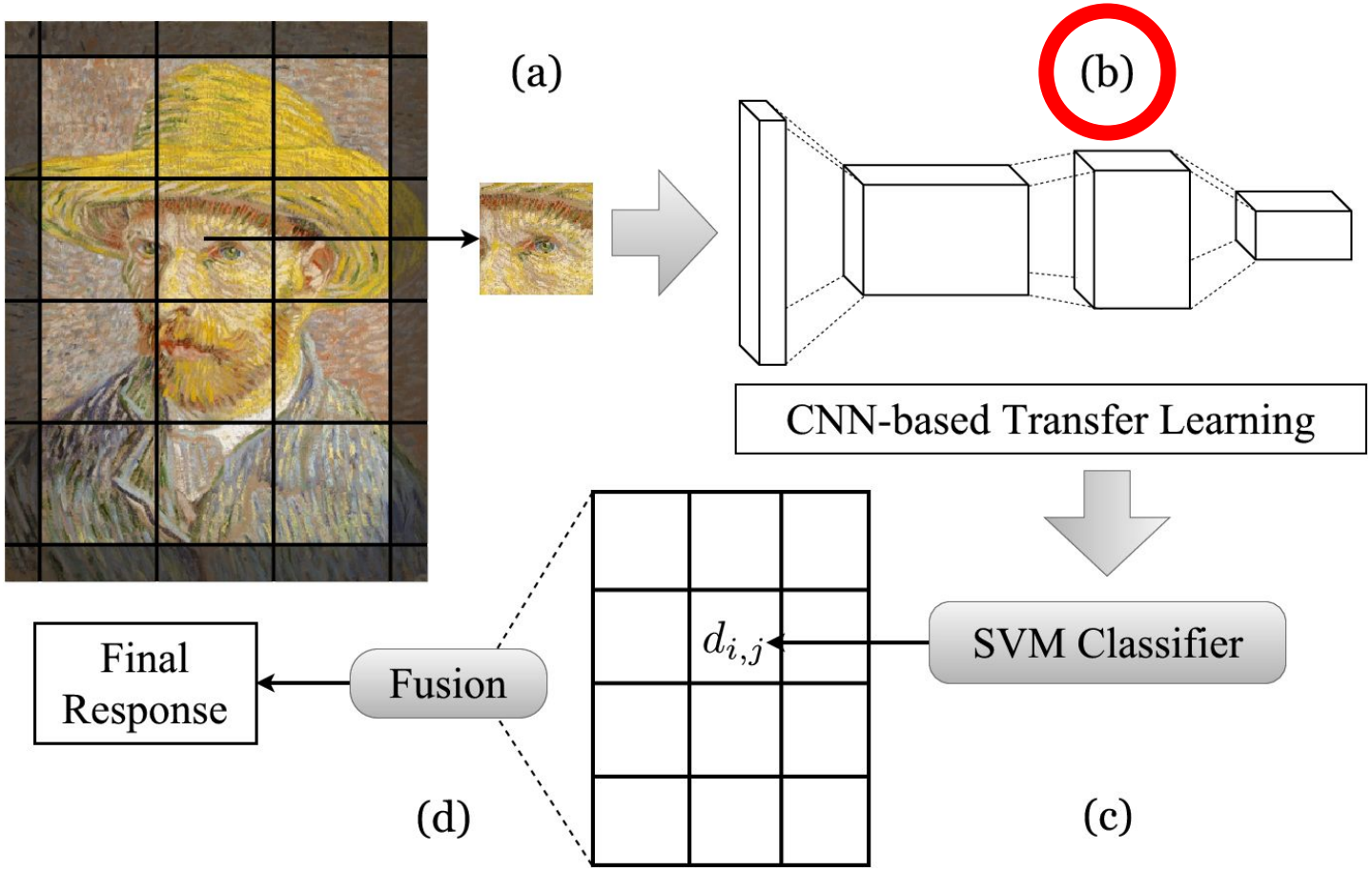
Our method

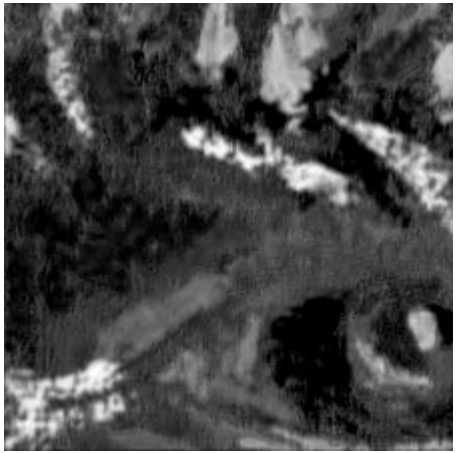
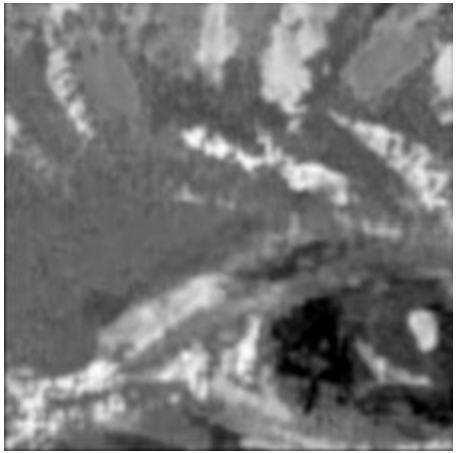
The proposed pipeline consists of four steps:

1. Divide each image into smaller patches
2. Extract features using a Convolutional Neural Network
3. Apply patch classifier
    a. At training time, generate the patch classifier model
    b. At test time, calculate the classification score for each patch
4. Use patch classification scores for a final response

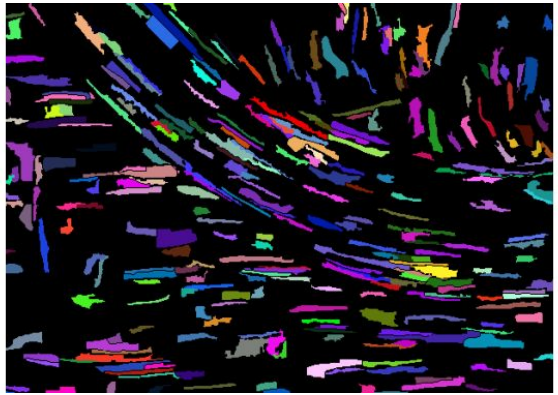
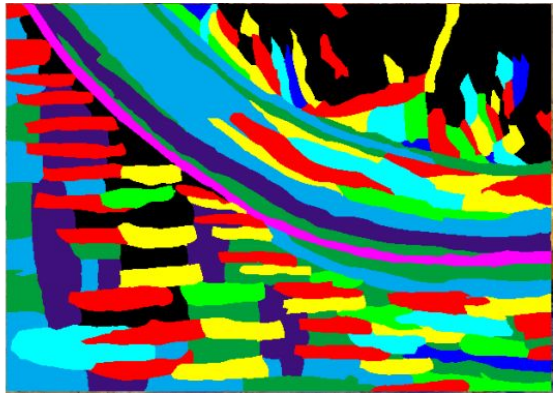

(a)
(b)
CNN-based Transfer Learning
SVM Classifier
$d_{i,j}$
Fusion
Final Response
(d)
(c)


(a)
(b)
CNN-based Transfer Learning
SVM Classifier
$d_{i,j}$
Fusion
Final Response
(d)
(c)

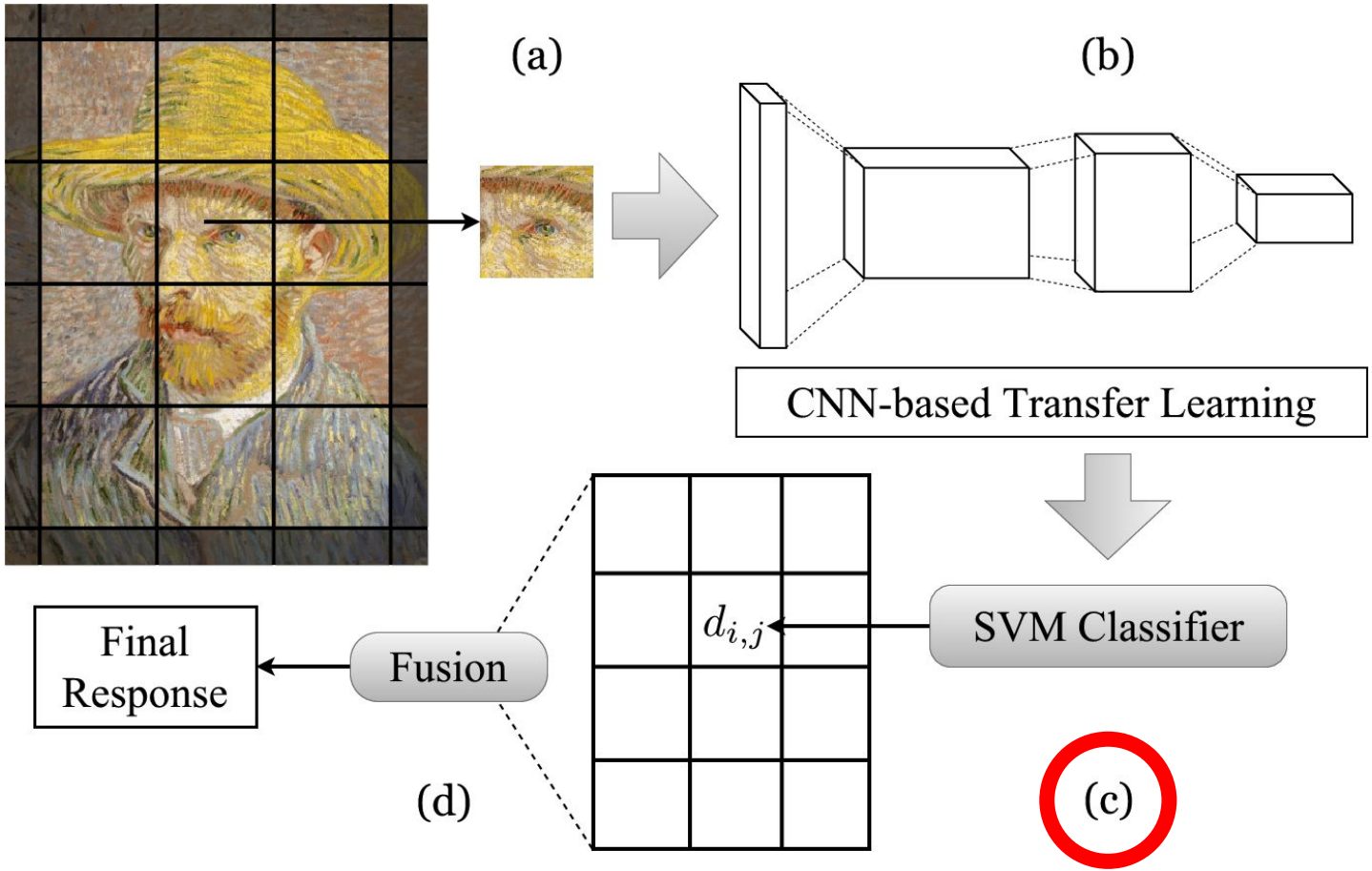

(a)
(b)
CNN-based Transfer Learning
SVM Classifier
$d_{i,j}$
Fusion
Final Response
(d)
(c)

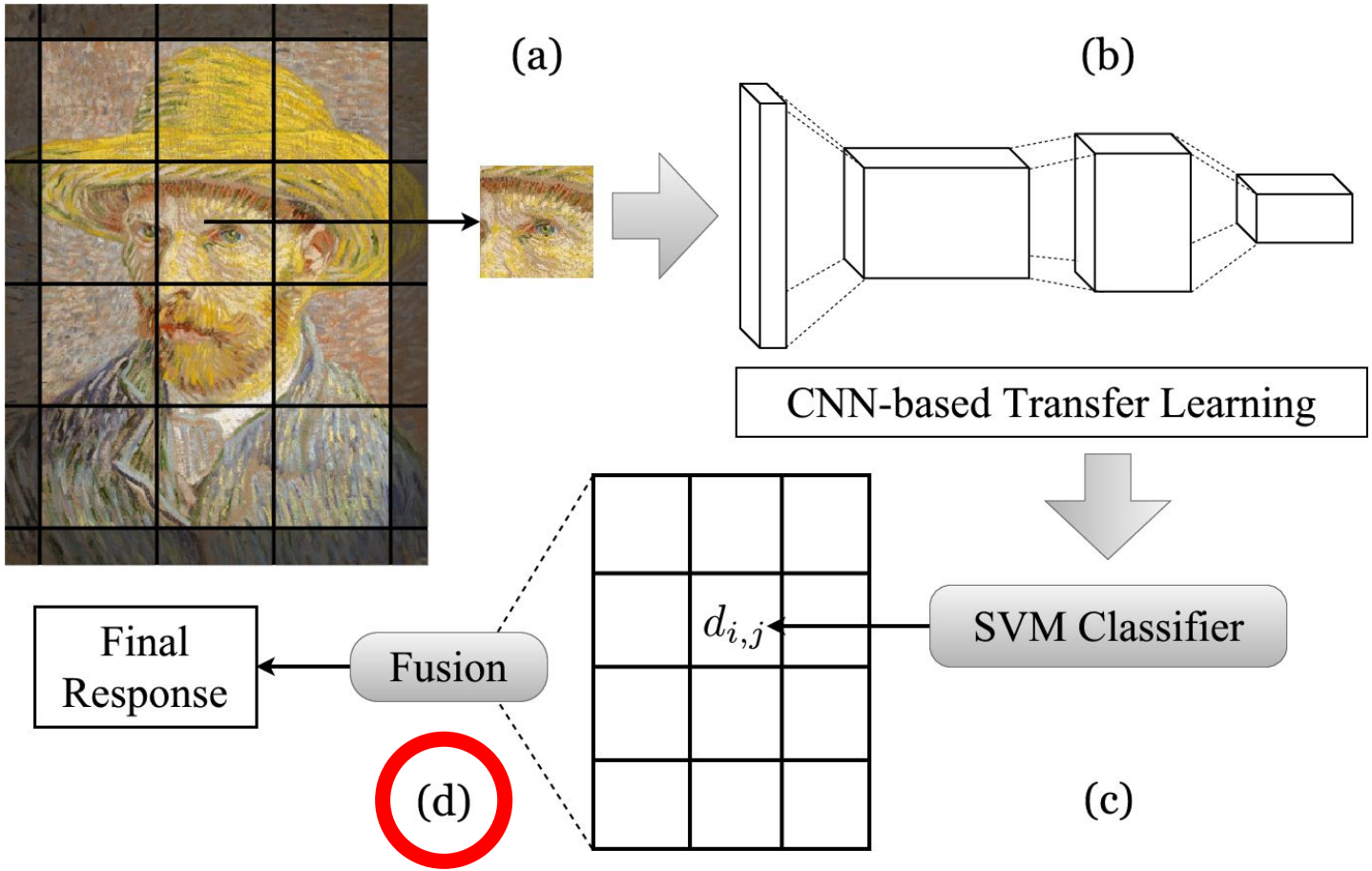




Features emphasizing brush strokes

Images from Li et al. [5]


(a)
(b)
CNN-based Transfer Learning
SVM Classifier
$d_{i,j}$
Fusion
Final Response
(d)
(c)


(a)
(b)
CNN-based Transfer Learning
SVM Classifier
$d_{i,j}$
(c)
Fusion
Final Response
(d)

Our method

Considering the two groups of positive and negative distances, we analyzed different fusion approaches:

- Mode
- Sum
- Far
- Mean
- Median

Our method

To the best of our knowledge, we created the very first public dataset for painting identification with high quality images and density standardization.

We gathered over 27,000 images from more than 200 categories in Wikimedia Commons [17].

The **VGDB-2016** dataset contains 207 van Gogh and 124 non-van Gogh paintings, which were randomly split, forming a standard evaluation protocol.

It also contains 2 paintings whose authorship are still under debate.

[17] Wikimedia, “Wikimedia Commons,” https://commons.wikimedia.org, [Online; accessed 2015-03].

Results

Mode

True Class	+	–
+	22	3
–	3	39

SMM

	+	–
+	24	1
–	4	38

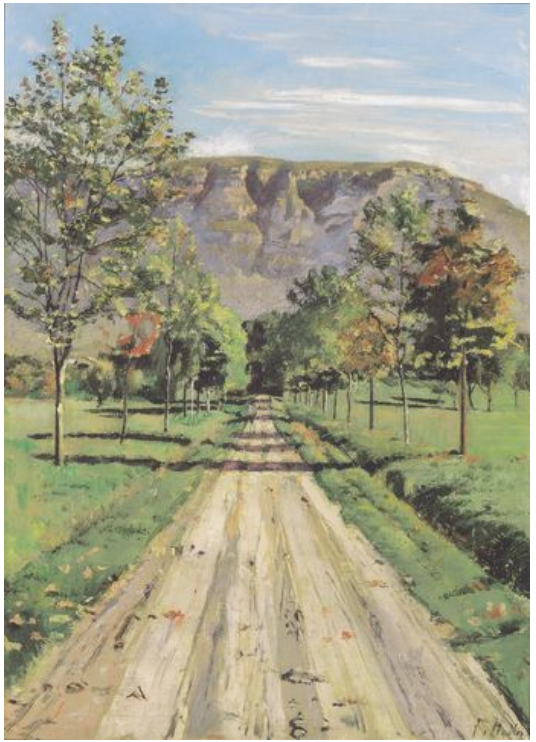
Far

	+	–
+	24	1
–	3	39

92.3%

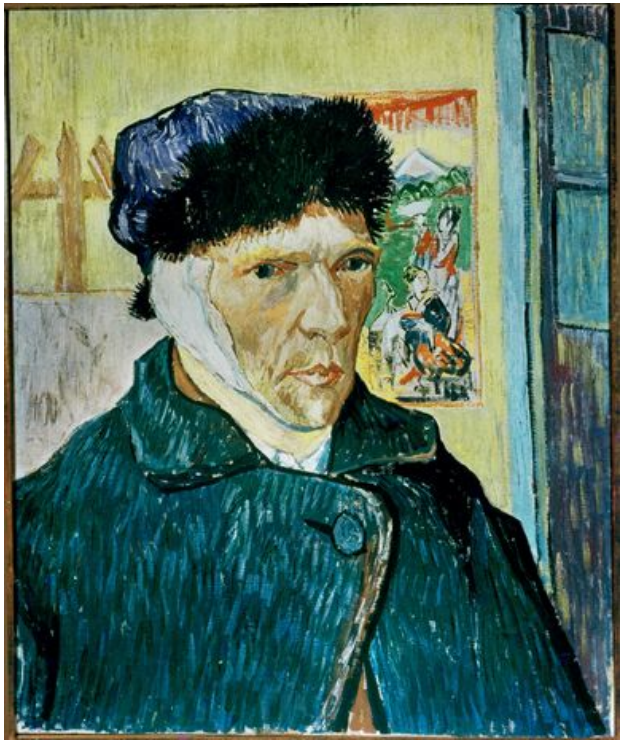
Best F1-Score (*Far*)

Portrait of a Girl
Paula Modersohn-Becker



The Road to Evordes
Ferdinand Hodler

The Palazzo Contarini, Venice
Claude Monet



Self-portrait with Bandaged Ear
Vincent van Gogh

Kingfischer

Portrait of a Woman

References

- The paper is available at IEEE Xplore (*free access until October 6, 2016*)
  - https://dx.doi.org/10.1109/icip.2016.7532335
- The dataset is available at figshare (*CC BY 4.0*)
  - https://dx.doi.org/10.6084/m9.figshare.3370627
- The source code is available at GitHub (*Apache 2.0*)
  - https://github.com/gfolego/vangogh
- There is also an entry in Kaggle Datasets
  - https://www.kaggle.com/gfolego/vangogh

You are free and encouraged to use and distribute the dataset and the source code developed in this paper!

Contact

Anderson Rocha
anderson.rocha@ic.unicamp.br

Thank you!