WEAKLY-SUPERVISED LOCALIZATION OF DIABETIC RETINOPATHY LESIONS IN RETINAL FUNDUS IMAGES

<u>JAN M. KÖHLER</u>, WALEED M. GONDAL
BOSCH CENTER FOR ARTIFICIAL INTELLIGENCE (BCAI), STUTTGART, GERMANY

RENÉ GRZESZICK, GERNOT A. FINK
DPTM. OF COMPUTER SCIENCE, TU DORTMUND UNIVERSITY, GERMANY

MICHAEL HIRSCH
MAX PLANCK INSTITUTE FOR INTELLIGENT SYSTEMS, TÜBINGEN, GERMANY

CONTACT: JAN.KOEHLER@DE.BOSCH.COM

BOSCH

IEEE International Conference on Image Processing 2017

Motivation: Working on retinal images to classify for diabetic retinopathy.

Retinal image

Algorithm

Classification Output

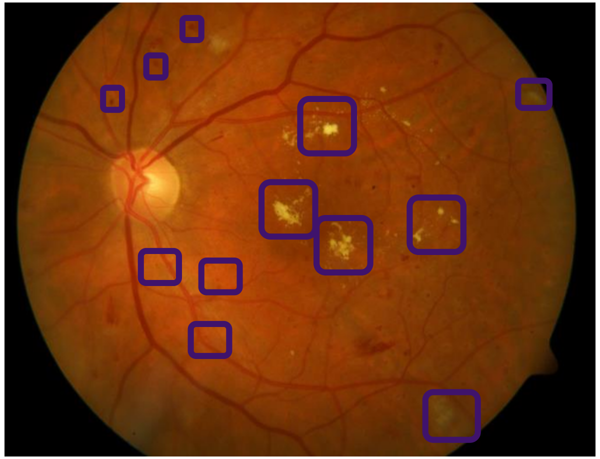
Healthy

Diseased

+

+

Lesion marking output



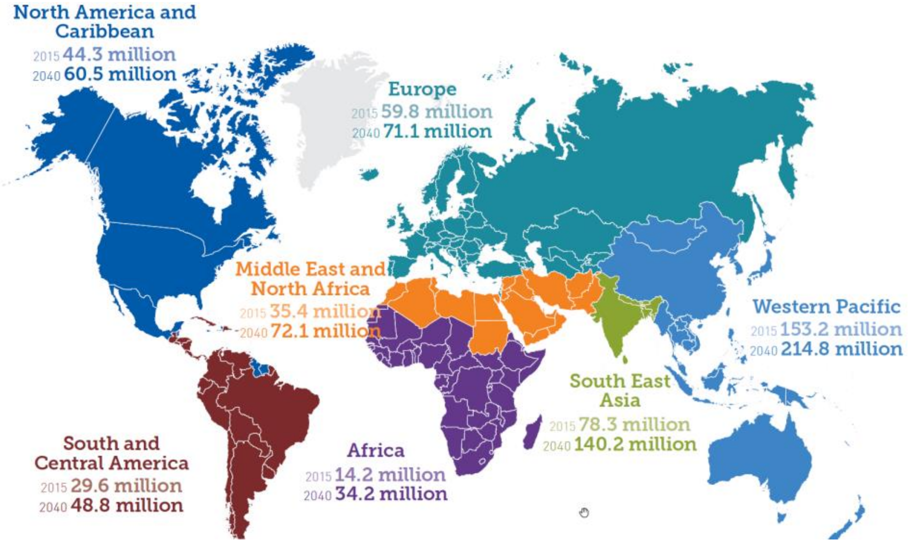
- Given a retinal image we want to classify the entire image for diabetic retinopathy.
- Additionally the individual lesions in the image responsible for the classification should be highlighted to increase trust of medical experts.
- The training of the algorithm has only image-level labels, i.e. no information about pixel-wise lesions (semi-supervised object localization).

Image source: [1]

IEEE International Conference on Image Processing 2017

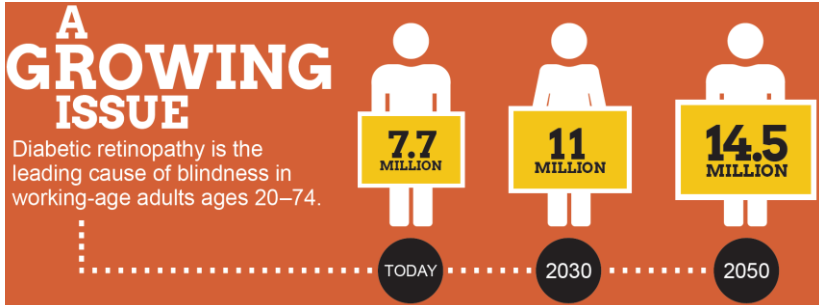
Diabetic retinopathy (DR): Some facts

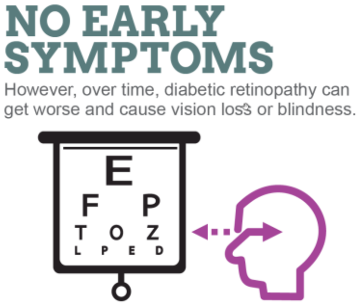
- In 2014 there have been 415m adults living with diabetes. About 145m (35%) had some form of diabetic retinopathy (DR). Among these 45m (11%) had vision-threatening DR. In 2040 about 642m adults will have diabetes.



- Low- and middle-income countries account for about 75% of the global diabetes cases. But medical infrastructure is lacking to identify and treat this disease.

- About 7m of people with diabetes are blind due to DR.





- There are no early symptoms, but early detection and treatment can reduce the risk of vision loss by 95%.

Sources: [2, 3, 4]

IEEE International Conference on Image Processing 2017

Why Bosch is in this topic? Bosch Eye Care Solution*

- Bosch India provides a non mydriatic, handheld camera for fundus imaging.
- It is easy to use, portable and suitable for mass screening.
- The main target market is mass screening in developing countries.
- Mass screening will be supported by machine learning algorithms for classification and object localization.
- Bosch Center for Artificial Intelligence (BCAI) supports with algorithmic development.

Handheld camera for fundus imaging

Prime minister of India Narendra Modi and Chancellor of Germany Angela Merkel

BOSCH

* More information at http://www.bosch-eyecare.com/

IEEE International Conference on Image Processing 2017

Method: Class Activation Maps with Global average pooling

⇨ Within a CNN, the last layer is set as a Global average pooling (GAP) layer.

1. **GAP:** From each of the K activation maps $A^k \in R^{u \times v}$ (width *u* and height *v)* the average value is calculated:

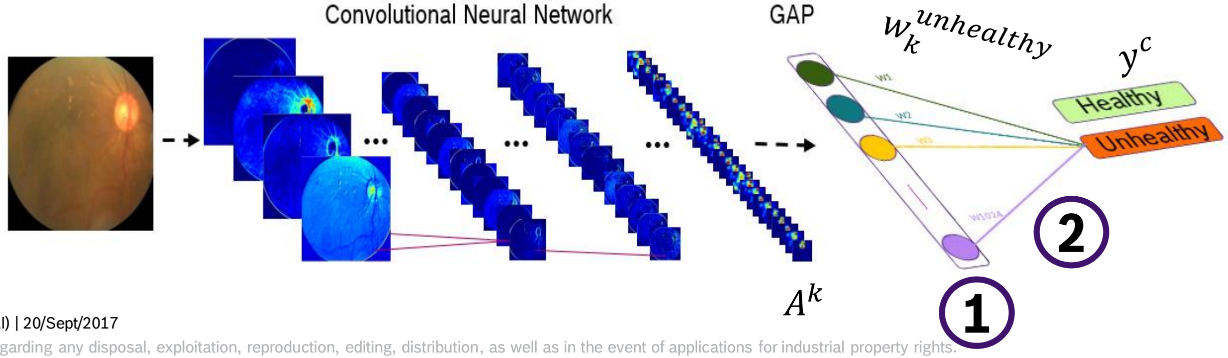
$$GAP: R^{u \times v} \to \mathrm{R}, \qquad \mathrm{A}^k \to \frac{1}{UV} \sum_x \sum_y A_{xy}^k$$

2. **Weights $\boldsymbol{w_k^c}$:** The GAP layer is fully connected to output neurons via w_k^c; c $\in \{healthy, unhealthy\}$

$$y^c = \sum_k w_k^c \; GAP\left(A^k\right)$$

Weights $\boldsymbol{w_k^c}$ encode the importance of each feature map A^k with respect to class *c*.

3. **Training:** The network is trained in a <u>weakly-supervised fashion</u>, i.e. only labels on image level (healthy, unhealthy) are available. No label information about any lesion type is given.



Jan Köhler | Bosch Center for Artificial Intelligence (BCAI) | 20/Sept/2017

BOSCH

Based on Zhou et al. „Learning Deep Features for Discriminative Localization“ [5]

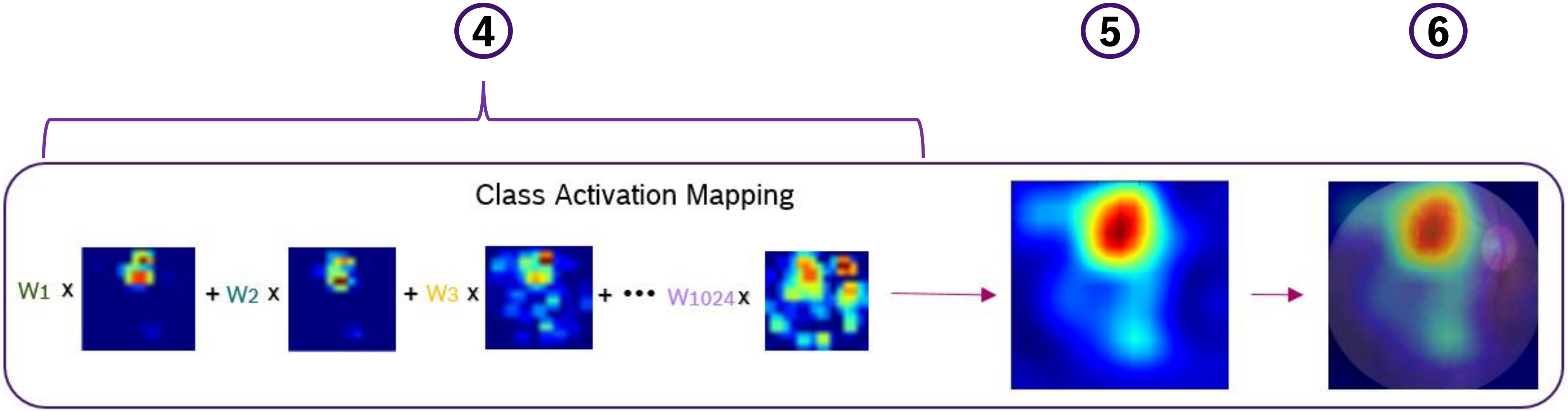
IEEE International Conference on Image Processing 2017

Method: Class Activation Maps with Global average pooling

4. **Localization map L:** Given an image, the weighted sum of the activation maps of the last convolutional layer forms the localization map.

$$L^c = \sum_k w_k^c \; A^k$$

5. **Upscaling:** The localization map is bilinear upscaled to image resolution
6. **Overlaying:** The original image is overlayed with the localization map



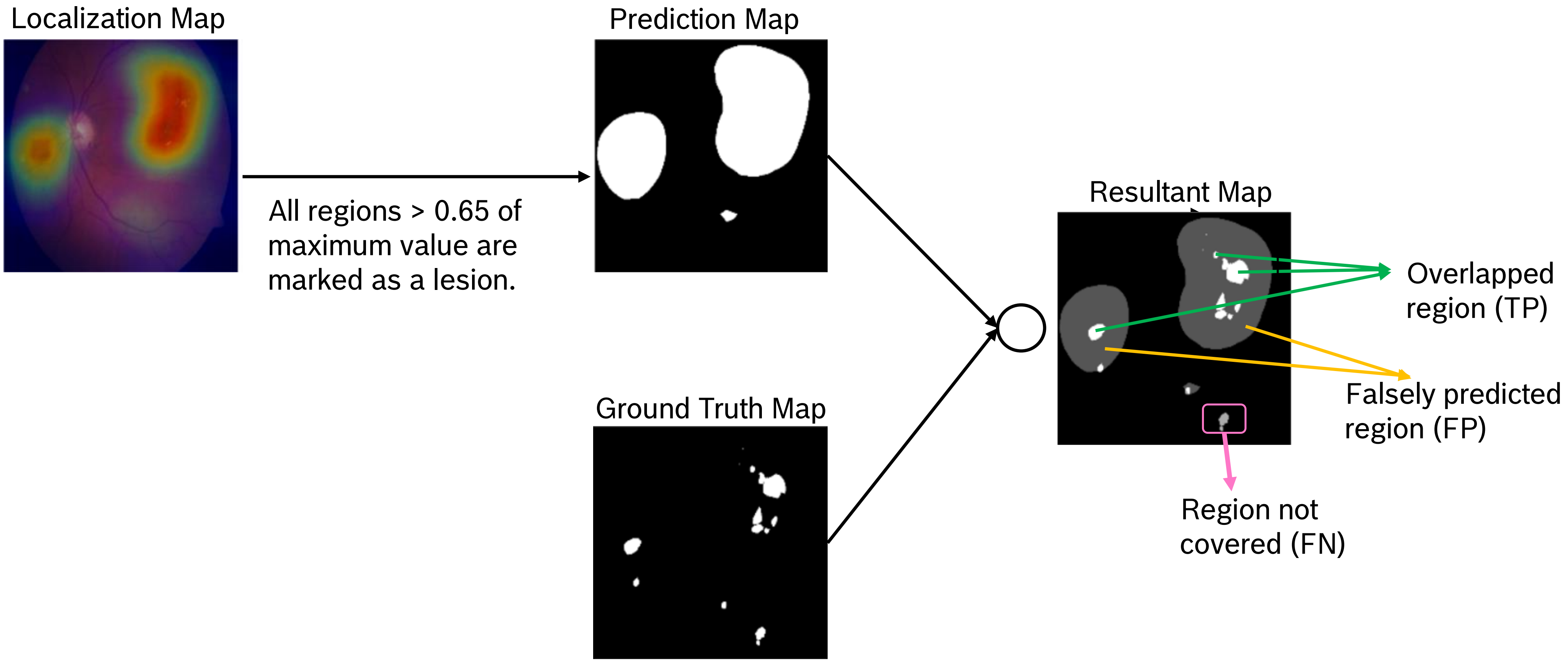
Localization map L *L + original image*

Jan Köhler | Bosch Center for Artificial Intelligence (BCAI) | 20/Sept/2017

Based on Zhou et al. „Learning Deep Features for Discriminative Localization“ [5]

IEEE International Conference on Image Processing 2017

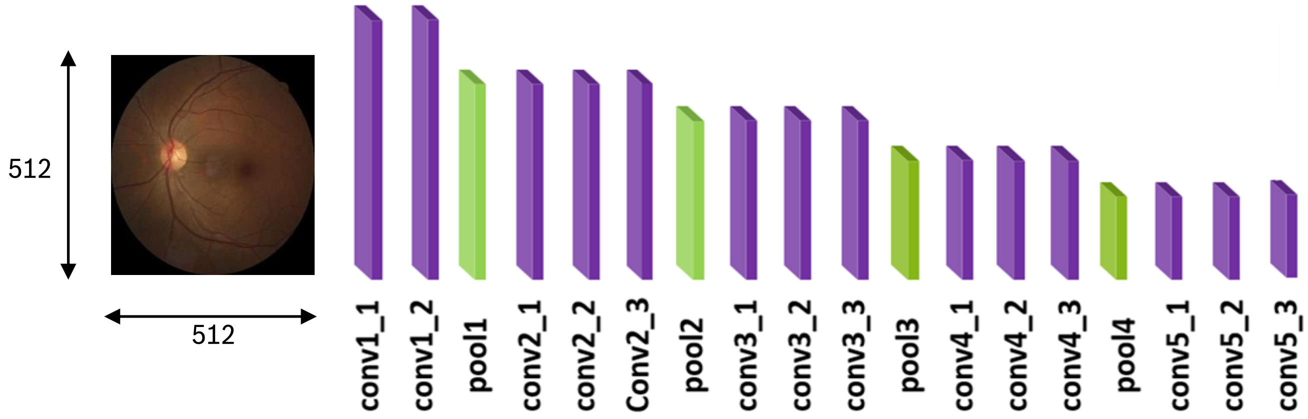
From localization map to True positive / False positive / false negative



Jan Köhler | Bosch Center for Artificial Intelligence (BCAI) | 20/Sept/2017

IEEE International Conference on Image Processing 2017
The network structure

- Network and training configuration:
  - VGG-16 based architecture
  - Batch normalization
  - L2 Regularization with weight decay factor of 0.0005
  - Gradient descent with momentum 0.8
  - Initial learning rate of 0.01 decayed by 1% after each epoch
  - 150 epochs for training

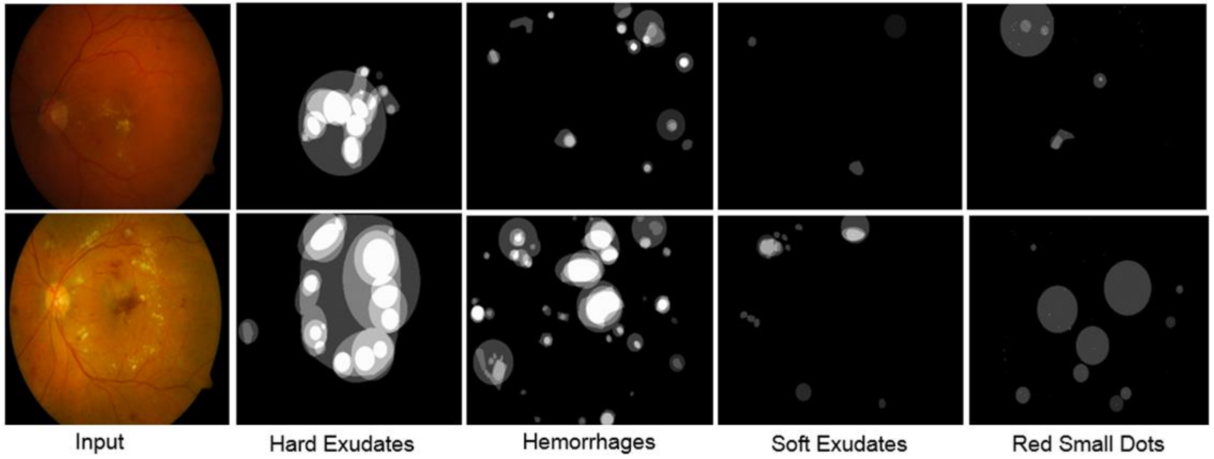


Experiments: Data sets used

Training data set:
Kaggle data set on diabetic retinopathy [6]

- 88,702 images of which 80% are used for training and 20% for validation
- No information about lesions given
- Collected in a clinical setting with high-end, stationary equipment.
- Five stages of diabetic retinopathy: For classification the first two classes were grouped into non-referable DR and the remaining three classes into referable DR.

Test data set:
DiaretDB1 data set [1]



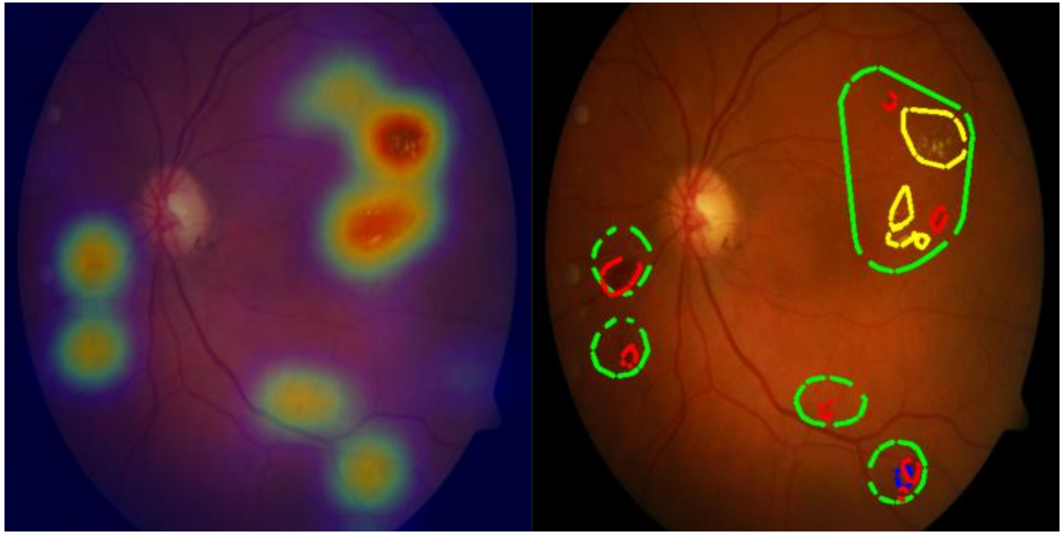
- 89 high resolution images used for testing
- Lesions marked by four experts.
- Regions with more than 75% consensus among the experts are considered as positive.

IEEE International Conference on Image Processing 2017

Results for object localization on diaretDB1 – example images

Test image [1]

Ground truth [1]



Our result:
localization map

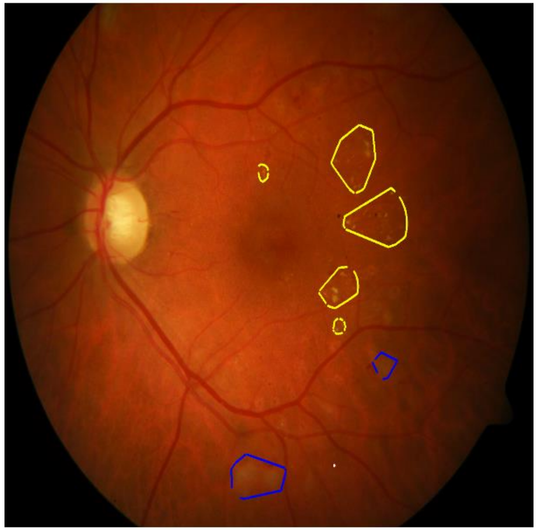
Ground truth and
our predicted regions

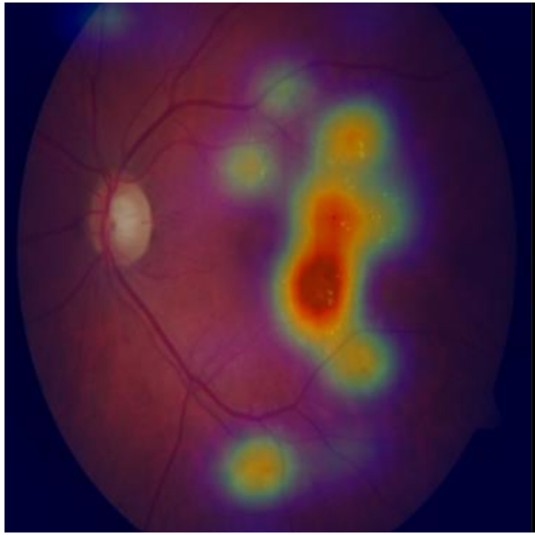
Ground truth: Yellow -- Hard Exudates   Blue -- Soft Exudates   Red -- Hemorrhages

Our result: Green -- Our predicted lesion regions

IEEE International Conference on Image Processing 2017
Results for object localization on diaretDB1 – example images


Test image [1]


Ground truth [1]


Our result: localization map


Ground truth and our predicted regions

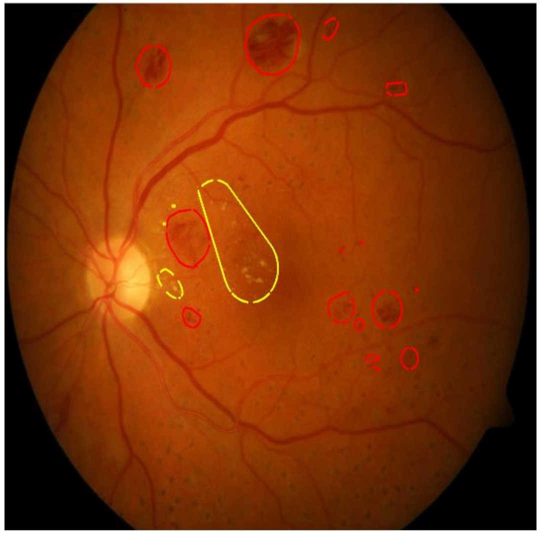
Ground truth: Yellow -- Hard Exudates Blue -- Soft Exudates Red -- Hemorrhages
Our result: Green -- Our predicted lesion regions

Jan Köhler | Bosch Center for Artificial Intelligence (BCAI) | 20/Sept/2017

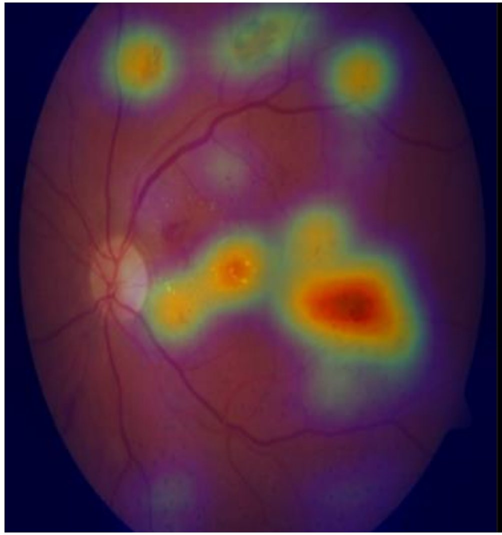
IEEE International Conference on Image Processing 2017

Results for object localization on diaretDB1 – example images

Test image [1]



Ground truth [1]



Our result:
localization map

Ground truth and
our predicted regions

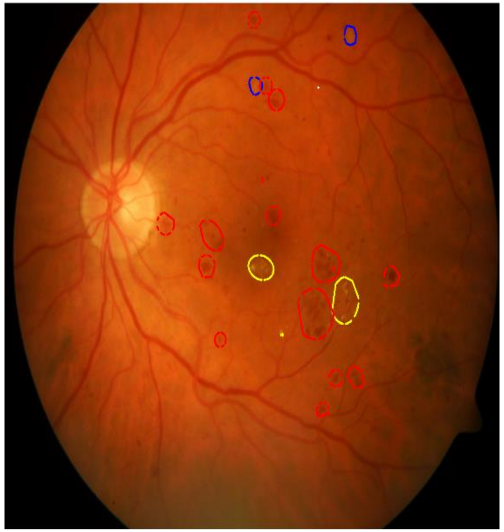
Ground truth: Yellow -- Hard Exudates Blue -- Soft Exudates Red -- Hemorrhages

Our result: Green -- Our predicted lesion regions

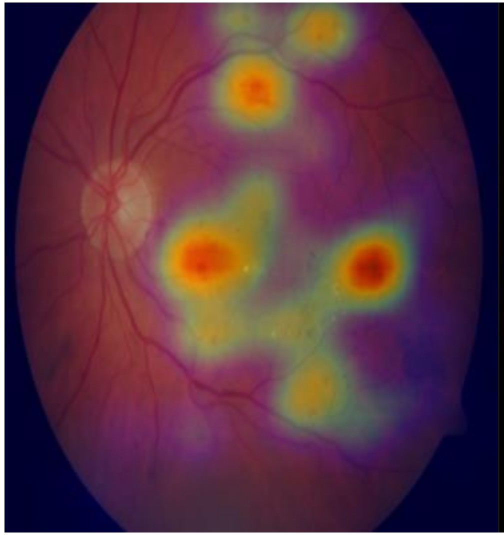
Jan Köhler | Bosch Center for Artificial Intelligence (BCAI) | 20/Sept/2017

Results for object localization on diaretDB1 – example images

Test image [1]



Ground truth [1]



Our result:
localization map

Ground truth and
our predicted regions

Ground truth: Yellow -- Hard Exudates Blue -- Soft Exudates Red -- Hemorrhages

Our result: Green -- Our predicted lesion regions

IEEE International Conference on Image Processing 2017
Lesion detection - performance at image level

Image level sensitivity [%]

Method	Type**	H*	HE*	SE*	RSD*
Zhou *et al.*[7]	S.	94.4	-	-	
Liu *et al.*[8]	S.	-	83.0	83.0	-
Haloi *et al.*[9]	S.		96.5	-	-
Mane *et al.*[10]	S.	-	-	-	96.4
Ours (50% Overlap)	W. S.	97.2	93.3	81.8	50
Ours (OnePixel Overlap)	W. S.	97.2	100	90.9	50

* H, HE, SE, RSD: Hemorrhages, Hard Exudates, Soft-Exudates and Red Small Dots.
** S.= supervised, W.S. = weakly-supervised

Image level sensitivity

- Detecting at least one lesion of type T on an image is counted as a True Positive.
- A lesion is detected if either one pixel or 50% of its pixels are covered by the prediction map.
- $\text{Sensitivity} = \frac{Number\ of\ images\ lesion\ type\ T\ is\ detected}{Number\ of\ total\ images\ showing\ lesion\ type\ T}$

Binary image classification (healthy vs. unhealthy) yields 93.6% sensitivity and 97.6% specificity on DiaretDB1 dataset with AUC of 0.954.

IEEE International Conference on Image Processing 2017

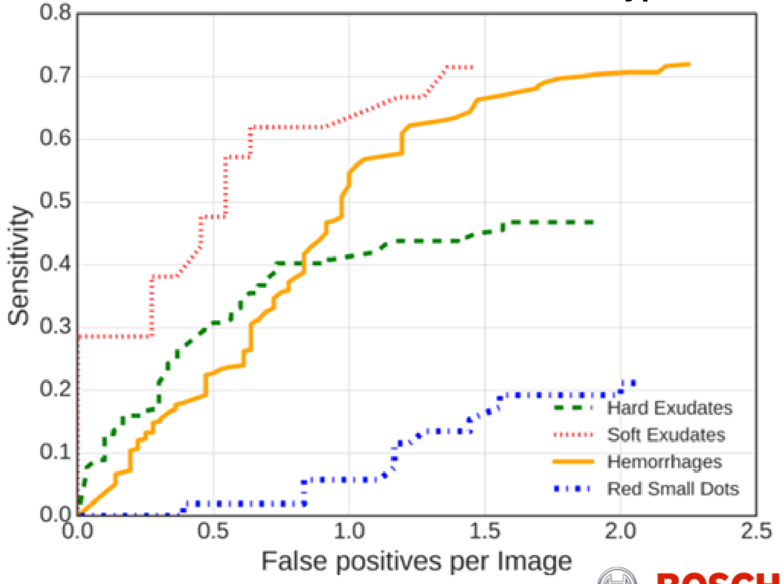
Performance at lesion level

Lesion level sensitivity [%]

Type	Method	Hemorrhages		Hard Exudates		Soft Exudates		RSD	
		SE%	FPs/I	SE%	FPs/I	SE%	FPs/I	SE%	FPs/
W.S.	Quellec *et al.* [11]	71	10	80	10	90	10	61	10
S.	Dai *et al.* [12]	-	-	-	-	-	-	29	20.30
W.S.	Ours (50% Overlap)	72	2.25	47	1.9	71	1.45	21	2.0
W.S.	Ours (OnePixel Overlap)	91	1.5	87	1.5	89	1.5	52	1.5

Type: S.= supervised, W.S. = weakly-supervised
SE: sensitivity
FPs/I: false positives per image

FROC curves for all four lesion types



Lesion level sensitivity

- Detecting at least one pixel (OnePixel Overlap) or 50% of the pixels (50% Overlap) of a lesion of type T is counted as a **True Positive**.
- A **False Positive (FP)** is a predicted region not containing any lesion type or a predicted region with mIOU<0.5.
- $\text{Sensitivity} = \dfrac{Number\ of\ detected\ regions\ of\ lesion\ type\ T}{Number\ of\ total\ regions\ of\ lesion\ type\ T}$

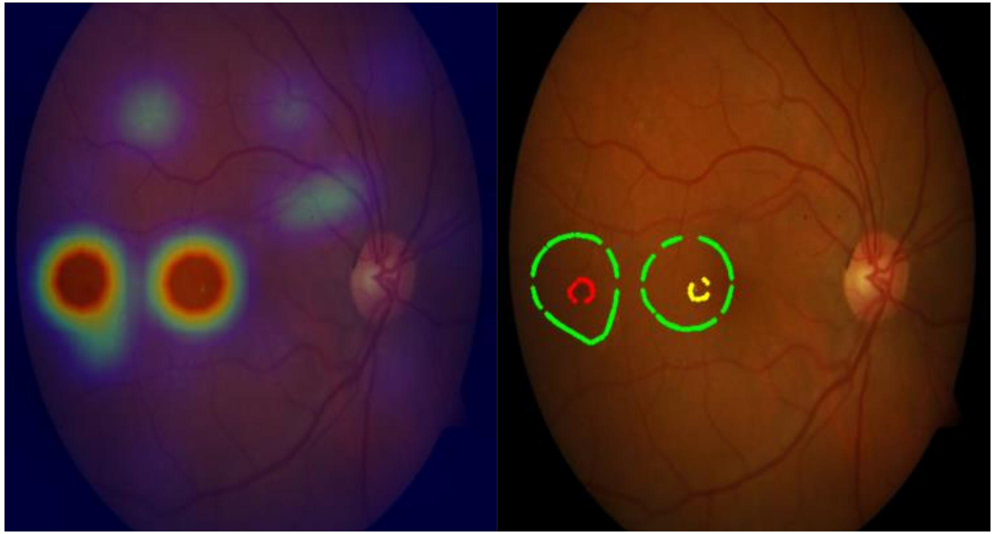
IEEE International Conference on Image Processing 2017
Summary

- Given only image level labels we could identify lesion regions which are important for the CNN for classification of retinal images.
- The network is trained for binary classification and classification accuracy is high, though introducing a GAP layer, with sensitivity of 93,6% and specificity of 97.6% on the DiaretDB1 test data.
- The sensitivity for detecting the lesion regions is beating or competitive to supervised methods.
- Red small dots are hard to localize. A reason might be the small resolution of the feature maps which are the basis of the localization map.

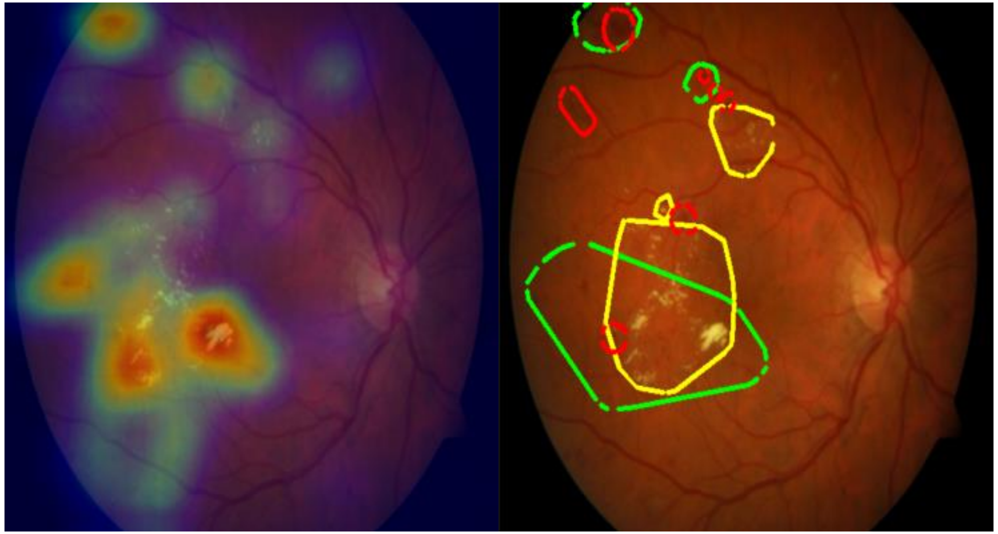
References

- [1] T. Kauppi, V. Kalesnykiene, J.-K. Kamarainen, L. Lensu, I. Sorri, A. Raninen, R. Voutilainen, H. Uusitalo, H. Kälviäinen, and J. Pietilä, “The diaretdb1 diabetic retinopathy database and evaluation protocol.,” in British Machine Vision Conference (BMVC), 2007, pp. 1–10.
- [2] World Health Organization, “Global report on diabetes,” 2016.
- [3] National Eye Institute https://www.nei.nih.gov/sites/default/files/nehep-pdfs/NEI_DR_Infographic.pdf
- [4] IDF Diabetes Atlas Seventh Edition, International Diabetes Federation, 2015
- [5] B. Zhou, A. Khosla, A. Lapedriz, A. Oliva, and A. Torralba, “Learning deep features for discriminative localization,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 2921–2929.
- [6] “https://www.kaggle.com/c/diabetic-retinopathy-detection,” assessed on 2017-08-14.
- [7] L. Zhou, P. Li, Q. Yu, Y. Qiao, and J. Yang, “Automatic hemorrhage detection in color fundus images based on gradual removal of vascular branches,” in IEEE International Conference on Image Processing (ICIP), 2016, pp. 399–403.
- [8] Q. Liu, B. Zou, J. Chen, W. Ke, K. Yue, Z. Chen, and G. Zhao, “A location-to-segmentation strategy for automatic exudate segmentation in colour retinal fundus images,” Computerized Medical Imaging and Graphics, vol. 55, pp. 78–86, 2017.
- [9] M. Haloi, S. Dandapat, and R. Sinha, “A gaussian scale space approach for exudates detection, classification and severity prediction,” preprint arXiv:1505.00737, 2015.
- [10] V. M. Mane, R. B. Kawadiwale, and D. Jadhav, “Detection of red lesions in diabetic retinopathy affected fundus images,” in IEEE International Advance Computing Conference (IACC), 2015, pp. 56–60.
- [11] G. Quellec, K. Charrière, Y. Boudi, B. Cochener, and M. Lamard, “Deep image mining for diabetic retinopathy screening,” Medical Image Analysis, 2017.
- [12] B. Dai, X. Wu, and W. Bu, “Retinal microaneurysms detection using gradient vector analysis and class imbalance classification,” PloS one, vol. 11, no. 8, pp. e0161556, 2016.

IEEE International Conference on Image Processing 2017
Thank you



Questions?



Remarks?

Contact: Jan.Koehler@de.bosch.com

BOSCH