Improving Skin Lesion Segmentation with Generative Adversarial Networks

Federico Pollastri*, Federico Bolelli*, Roberto Paredes [†], Costantino Grana*

*DIEF, Università degli Studi di Modena e Reggio Emilia, Italy
[†]PRHLT Research Center, Universitat Politécnica de Valencia, Spain

Problem Statement

Malignant melanoma is the most dangerous skin cancer, with a substantial death rate. Automated skin lesion segmentation is a fundamental step to help experts in early diagnosis, but requires a huge amount of data to be performed. Unfortunately, manual image segmentation is a very time consuming job that demands the work of a competent specialist.
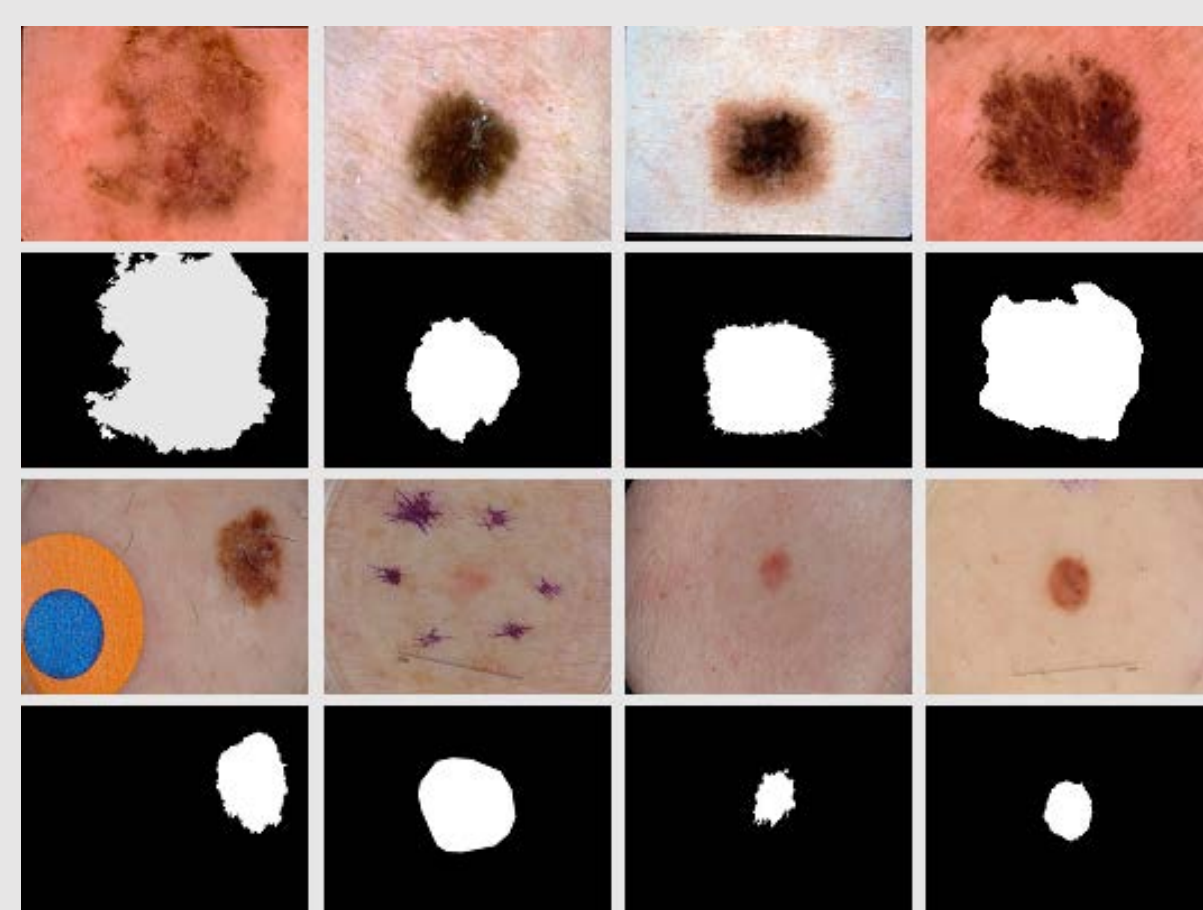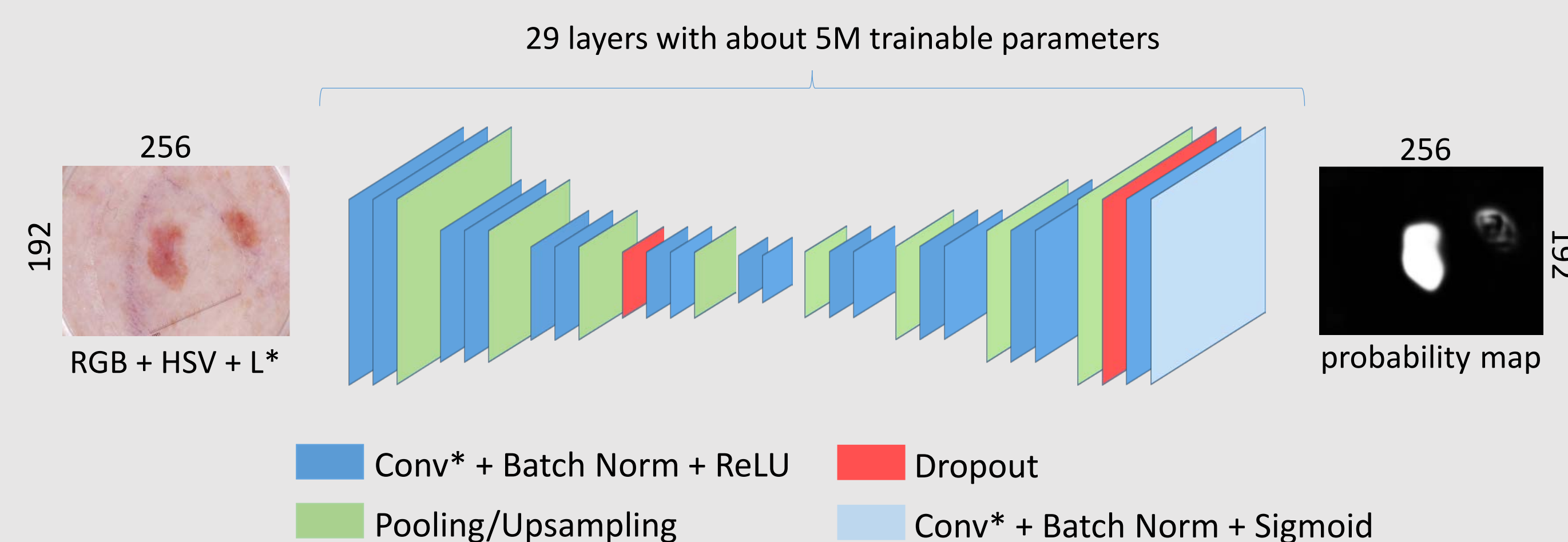

Samples from the ISIC dataset.

Goals:

- **decrease the amount of manually segmented images required by automated analysis;**
- **support the clinical decision making.**

Hyperparameters Analysis

The baseline CDNN is mainly affected by three hyperparameters, which are stressed in our analysis:

1) Input image size:
- 96x128
- 192x256
- 384x512

2) Image channels:
- RGB
- RGB + HSV + L*
- RGB + HSV + CIELAB

3) Loss function:

$$L = 1 - \frac{\sum_{i,j} t_{ij} p_{ij}}{\sum_{i,j} t_{ij}^2 + \sum_{ij} p_{ij}^2 - \sum_{i,j} t_{ij} p_{ij}} \quad (Eq.\ 1)$$

$$d_j = 1 - \frac{\sum_{i,j} t_{ij} p_{ij}}{\sum_{i,j} t_{ij} + \sum_{ij} p_{ij} - \sum_{i,j} t_{ij} p_{ij}} \quad (Eq.\ 3)$$

$$L = \frac{1}{n} \sum_{i,j} (t_{ij} - p_{ij})^2 \quad (MSE)$$

Baseline Architecture

Our model maps the input dermoscopic image into a posterior probability map, exploiting an architecture based on the CDNN which won the International Skin Imaging Collaboration (ISIC) challenge in 2017 *[1]*.
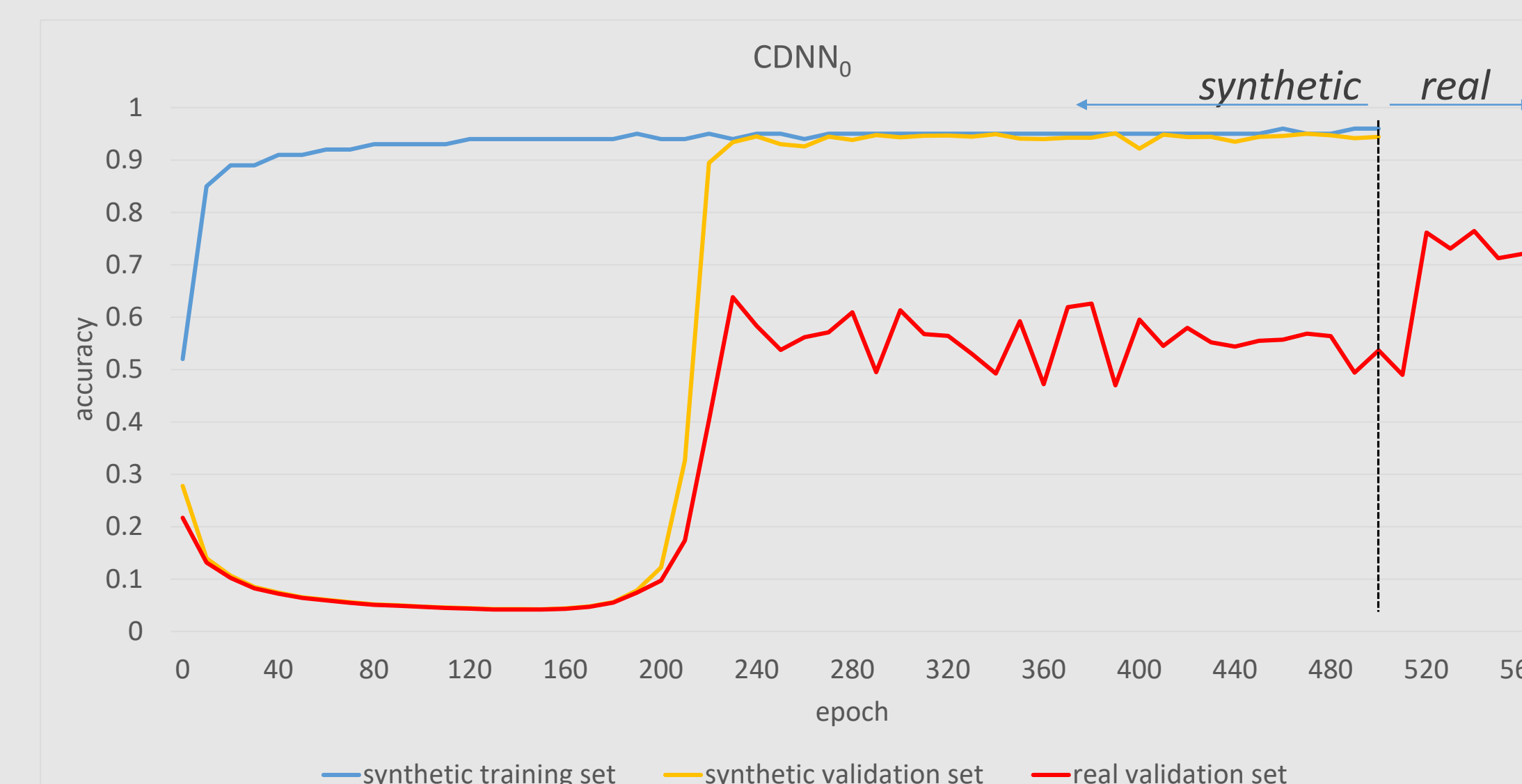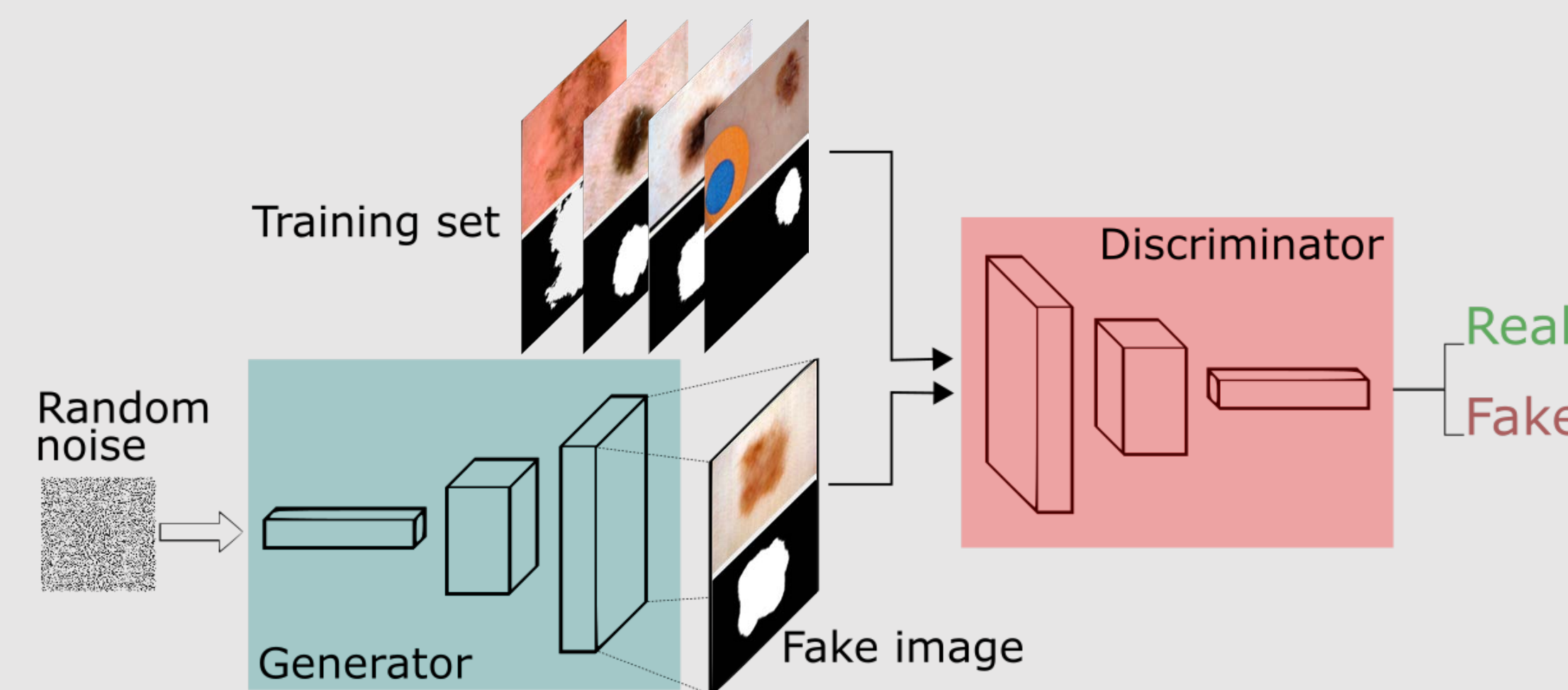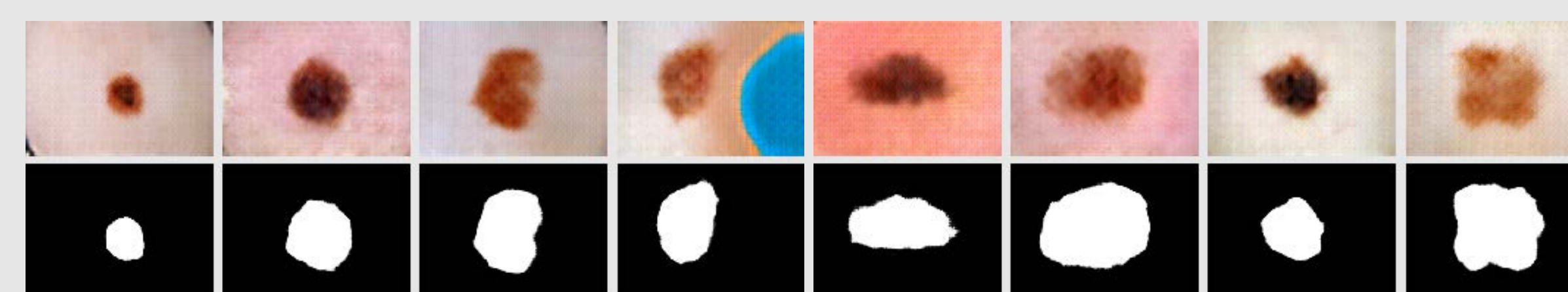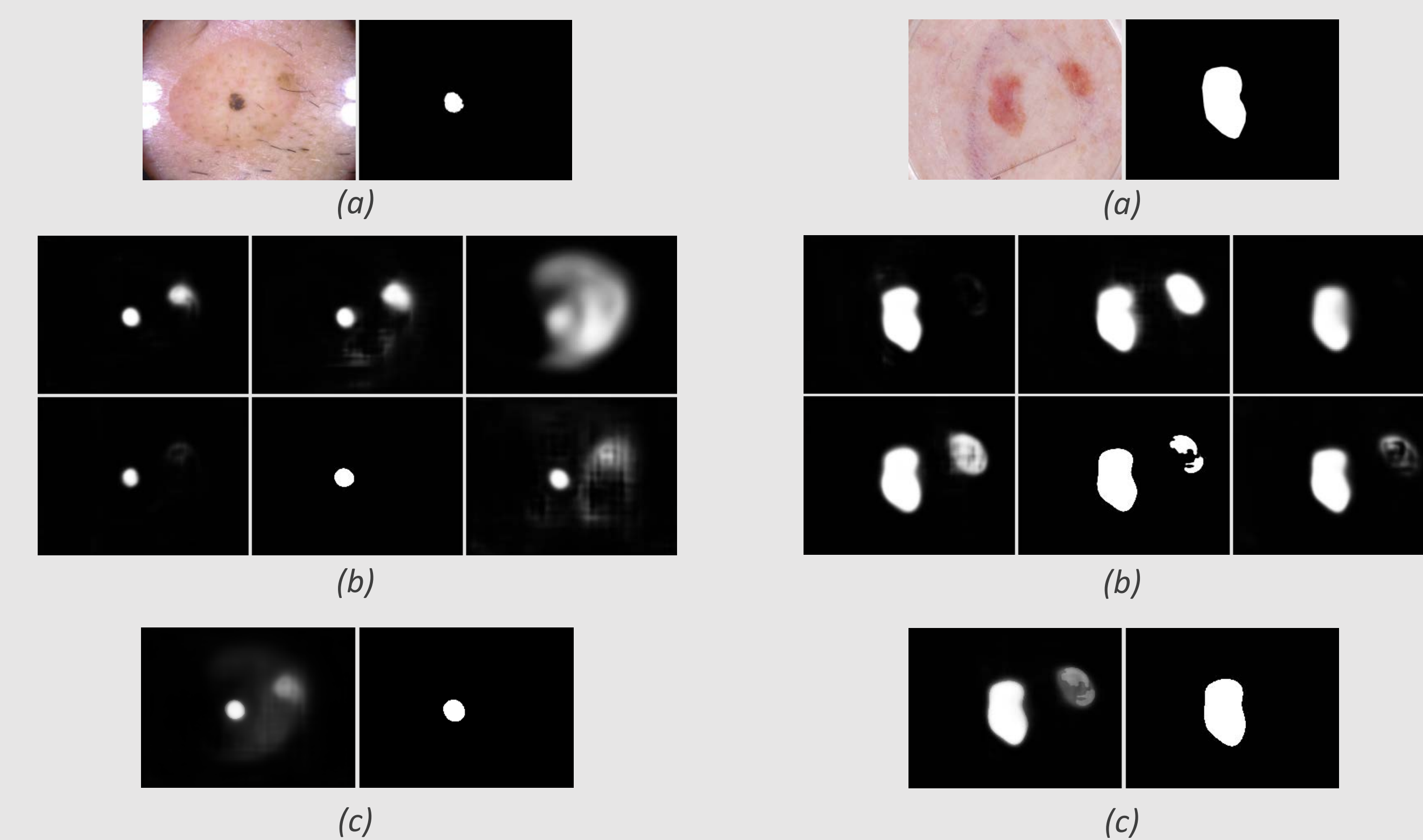

29 layers with about 5M trainable parameters

256 / 192 RGB + HSV + L* → 256 / 192 probability map

- Conv* + Batch Norm + ReLU
- Pooling/Upsampling
- Dropout
- Conv* + Batch Norm + Sigmoid

* fixed stride of 1 pixel.

[1] Yuan, Yading, Ming Chao, and Yeh-Chi Lo. "Automatic skin lesion segmentation with fully convolutional-deconvolutional networks." arXiv preprint arXiv:1703.05165 (2017).

GAN Augmented Data

Training set — Discriminator → Real / Fake

Random noise → Generator → Fake image

A DCGAN is employed to generate both the skin lesion image and its segmentation mask, improving the data augmentation process.

Samples from the generated dataset.

Experimental Results

Neural Network	Input Size	Number of Channels	Loss Function	Without Augmentation	With Augmentation
CDNN$_0$	192x256	7	Eq. 1	0.731	0.743
CDNN$_1$	192x256	3	Eq. 1	0.732	0.753
CDNN$_2$	192x256	9	Eq. 1	0.734	0.743
CDNN$_3$	96x128	7	Eq. 1	0.735	0.750
CDNN$_4$	384x512	7	Eq. 1	0.700	-
CDNN$_5$	192x256	7	Eq.3	0.738	0.738
CDNN$_6$	192x256	7	MSE	0.738	0.739
			Ensemble:		**0.781**

CDNN$_0$ — *synthetic* | *real*

Visualization of the training process: synthetic data are exploited to performed an "initialization" of the CDNN, which is then fine-tuned using the real data.

- synthetic training set
- synthetic validation set
- real validation set

(a) Input image and ground truth. (b) Left-to-right then top-to-bottom, output prediction of CDNN$_1$, CDNN$_2$, CDNN$_3$, CDNN$_0$, CDNN$_5$ and CDNN$_6$. (c) Outputs ensemble before and after binarization.