M Ű E G Y E T E M 1 7 8 2

Budapest University of Technology and Economics
Faculty of Electrical Engineering and Informatics
Department of Control Engineering and Information Technology

Melanoma detection using deep learning technology

Kőszegi Balázs

Table of contents

1 Abstract:

Melanoma, also known as malignant melanoma, is the most dangerous type of skin cancer. It develops form melanocytes, pigment containing skin cells. The number of cases rise rapidly and researchers around the world are trying to create more efficient, automated methods of screening. One promising field is using neural networks to judge the risk of a skin lesion. The benefits include fast and fairly accurate results on large amounts of data. In my project, I'm trying to create a network with these goals in mind, using deep learning.

2 Melanoma

2.1 What is a melanoma?

Melanoma is a malignant skin lesion, a type of skin cancer. The primary cause is exposure to UV light, which damages the DNA, although there are some other risk factors such as genetic defect, poor immune system and the high number of moles. About 25% of melanomas develop from moles and it is more frequent in men than women. It usually develops in middle-aged people.

Early detection is very important, because survivability rates decrease rapidly with time. A stage 0 melanoma (only local lesion with no metastasis, involving only the skin surface) have a 99.9% survival rate, while advanced stages quickly drop survival chances below 20% on a 5 year term. [1]

2.2 Detection

Skin lesions should be checked regularly. This is done by visual inspection and in case of suspicion, the problem is confirmed by biopsy. As the first defence line is the individual, self-check is highly recommended. This can be simplified by creating tools that can automatically rank the lesion and calculate the risk. Neural networks are well suited for this purpose.

The usual method of visual detection is using the ABCDE rule:

- A: Assymmetrical – The shape is assymmetrical
- B: Border – The border is irregular
- C: Color – Multiple color means higher risk
- D: Diameter – Mole diameter more than 6 mm means higher risk
- E: Enlarging – Growing over time

The difficulty is that these point are guidelines, and they can also apply to other kind of lesions. However, using these guidelines, the high risk moles can be filtered efficiently.

3 Neural networks

3.1 The concept

The concept of neural networking is to model nature’s most complex computer, the brain. Even if we don’t fully understand how it works, we can use the connected node approach to create something similar.

The basic element of a neural network is a neuron. It is a unit with multiple inputs and outputs and it performs a simple task (transfer function). These neurons are organized into layers. Different layers are connected, but nodes in the same layer are not connected. The first layer is called the input layer which passes on information to the so called hidden layers.

The hidden layers process the information and compute the output.

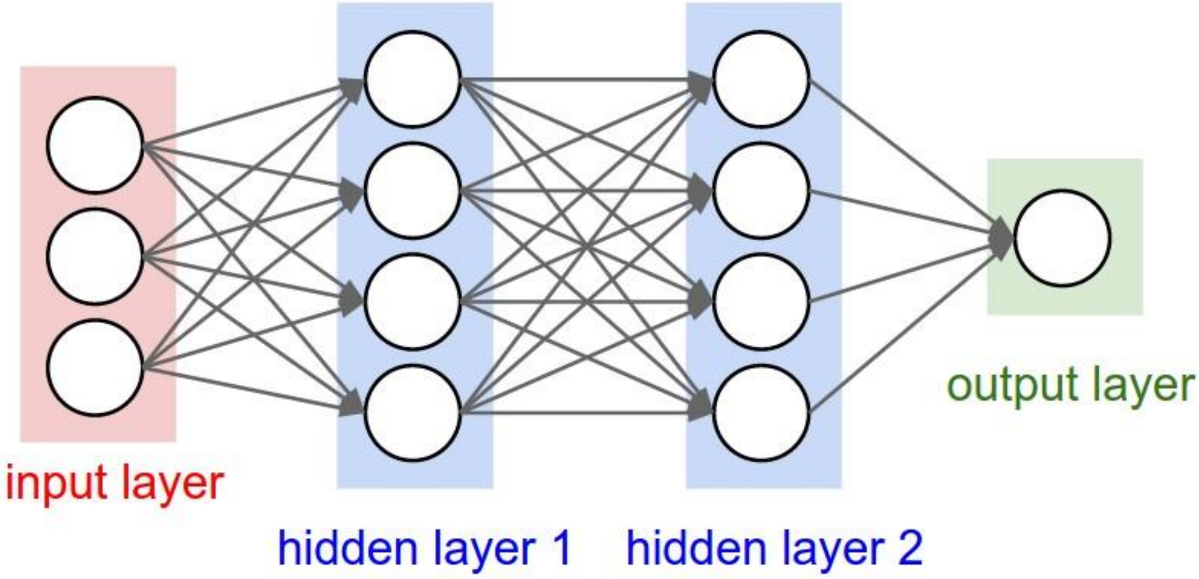

Figure 1: A neural network

The most interesting feature of the neural networks is that they are capable of learning. Teaching a network means that it is fed a large amount of data and the network is trying to find an optimal solution to accomplish its task. This is done by a so called cost function. Cost functions measure how far is the solution from being optimal. As learning progresses, the cost should converge to the minimum, which means the solution is optimal.

There are several cases where a neural network outperforms traditional programming methods. [3] Problems which are difficult to formulate, pattern recognition, classification, generalization or problems with insufficient theoretical background but large sample data are such cases. The advantage that the network figures out the best way to the task on its own, however, it needs a well constructed dataset. The disadvantage is that we can’t really see the inner workings of the network which makes debugging very hard.

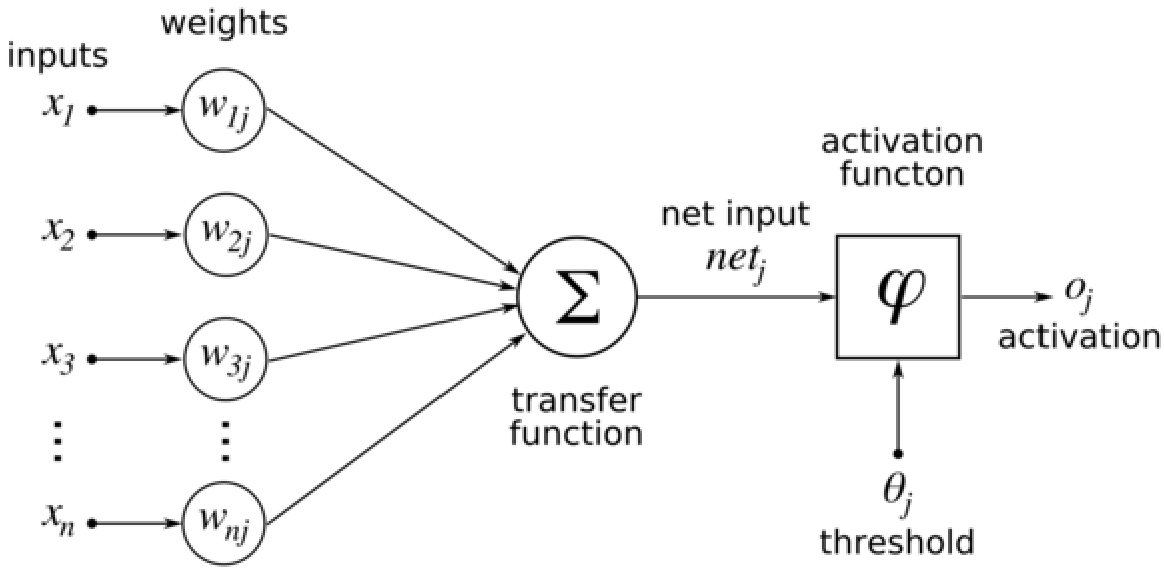

Figure 2: A neuron

The above picture shows the structure of a neuron. The inputs are weighed and summed. The net input is fed to the so called activation function which is the transfer function of the neuron. Neurons may also have a threshold to disable their operation under a certain limit. If the input is above the threshold, the neuron „fires", produces an output between 0 and 1.

Learning is a process where the network model corrects itself to achieve a smaller error with each cycle. This is achieved by data exchange between the neurons. The data flow in the input-to-output direction is called feedforward. If the network doesn't form a cirlce, it is called a feedforward network. Self correction happens because of the concept called backpropagation. That means when the input is processed and the output is calculated, the loss function calculates the error compared to the desired output. This error is then used to update the weights of the neurons in attempt to lower the loss.

3.2 Convolutional Neural Networks

Convolutional neural networks are inspired by animal visual cortexes. The pictures are divided into small regions which are processed by different neurons. These image parts partially overlap to give the full visual field. Convolution operation is used to simulate the response of a biological neuron to an input image.

A convolutional neural network has multiple receptive layers which process parts of the input image. A typical process chain include:

- **convolutional layers:** the basic block of the network. They have learnable filters (kernels) which are small windows, scanning through the whole image. Each filter creates an output showing certain features of the image.
- **max pooling layers:** they examine a partial image and find the maximum value. Only this value is passed through, which helps to reduce data size. With each pooling layer, the image gets more downsampled.
- **ReLU layers:** Rectified Linear Units layer. It increases the nonlinearity of the decision function.
- **fully connected layers:** All neurons are fully connected with the previous layers, like in regural neural networks.
- **loss layer:** It computes the error of the output. For classification uses, this layer is usually a softmax layer.

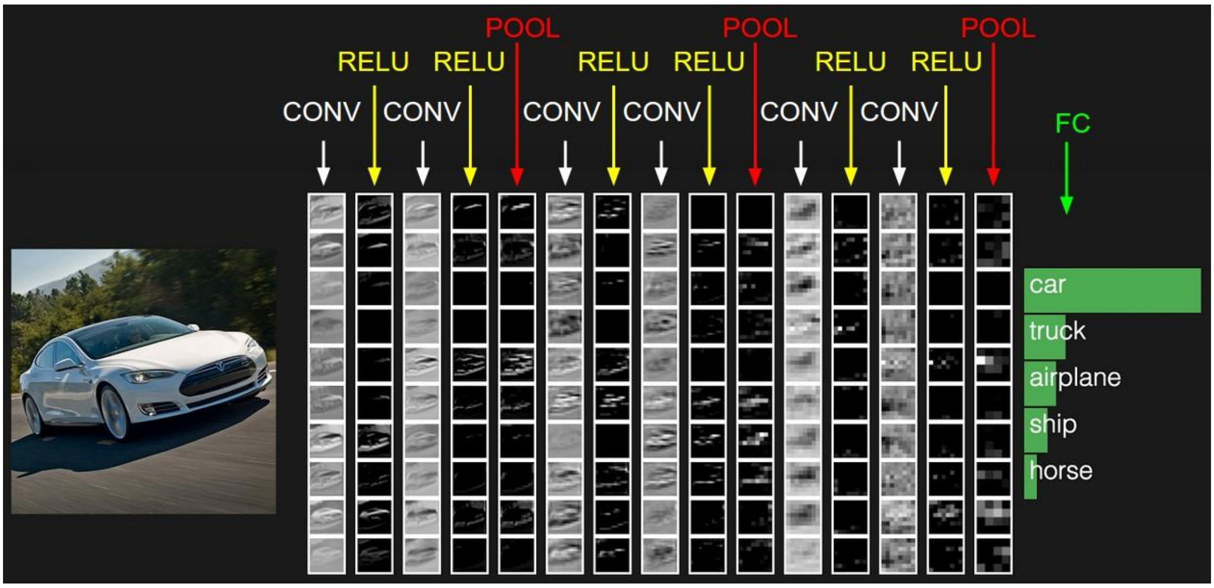

Figure 3: Convolutional network [6]

In this classification example, the output is a probability for the best matches.

3.3 Deep learning

My neural network is used for pattern recognition. The performance a network can achieve is connected to the depth of the network, which means the more layers you have, the better will be the performance. Early researches lacked the computation power to create complex networks, but nowadays powerful hardware and GPU computation allows that.

Deep learning is a term used for machine learning specialized in certain tasks, image, sound or text recognition. [5] It is a promising technology in automated vehichles, artificial intelligence and other fields. I am using it because it suits my needs of loosely defined pattern recognition stated above.

To create a neural network, I needed 2 things: Torch and a dataset

3.4 Torch

Torch [7] is a computational framework mainly intended for machine learning and neural networks, available for Linux, iOS and Android. It features many official and user contributed libraries for different tasks.

Torch can be programmed by either Lua or Python language. The high number of built in functions make it easy to create neural networks. It supports CUDA, which uses GPU computing for accelerated parallel computation.

3.5 My network

The network I will train is a convolutional network. Without prior knowledge on network performance, the initial structure will be the following:

- Convolution layer
- ReLU layer
- Max pooling
- Convolution layer
- ReLU layer
- Max pooling
- Linear layer
- ReLU layer
- Linear layer
- ReLU layer
- Linear layer
- Softmax layer

I will experiment with other structures as well when I gain some experience.

4 Dataset

A dataset is the information used for training and testing the neural network. It is crucial to use an adequate sized database, as it is the component that will shape the network and its operation.

A dataset consists of two parts:

- training set
- test set

Training data is the database used for training. It should consist of at least several thousand images. My training set consists of more than 9300 images.

Test data is for performance testing. After training test data can be used for accuracy and processing speed measurements, needed for further improvements to the network. Test data should not overlap with the training set. My test set consists of 1250 images.

There can be issues concerning the training:

- **Overfitting:** Overfitting occurs when the network memorized the training samples but didn't learn to generalize. This means that using the training datafor tests (which is bad practice, that is the reason a separate test set should be used), very small error is achievable, but when the network is fed unknown data, the error is large. There are several methods to solve this, for example retraining, averaging the output of multiple networks or regularization.
- **Underfitting:** The model can't model the training data nor generalize. A new algorithm may be needed.

My dataset includes healthy moles, benign and malignant lesions. I used the ISIC archive [8] as my image source, which is an online database with the aim to help advancements in automatic lesion recognition.

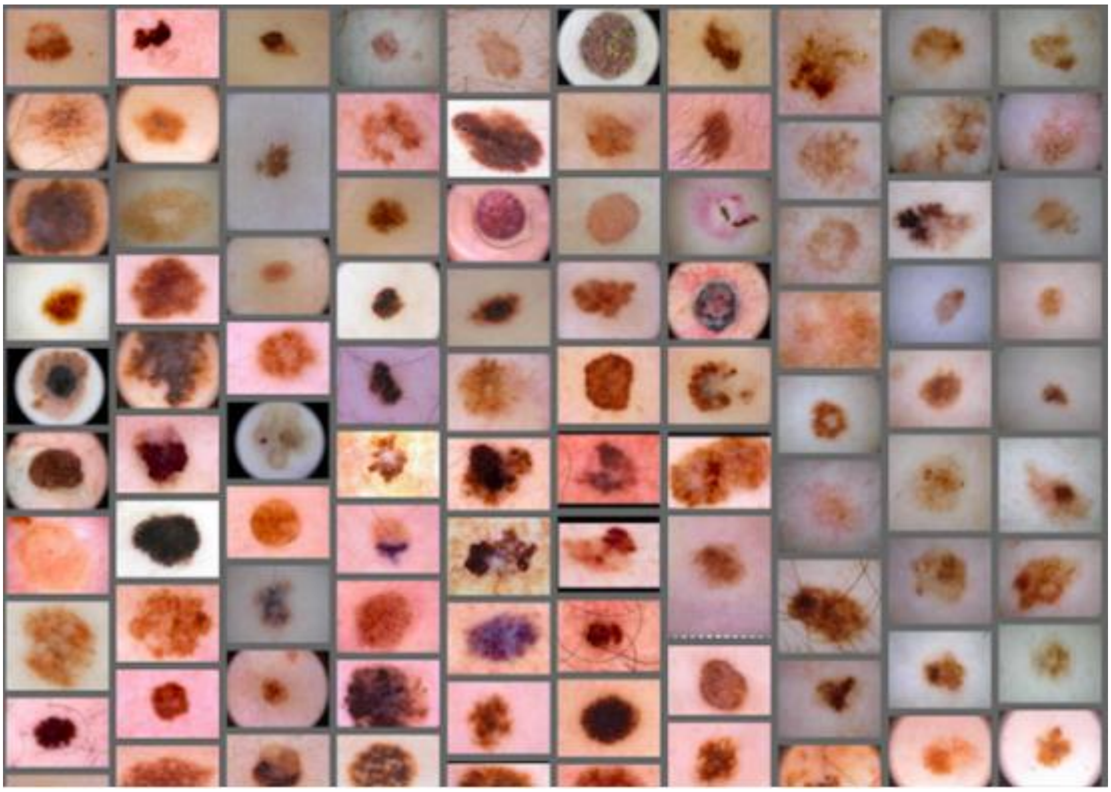

Figure 4: ISIC archive

ISIC archive doesn't specialize in melanoma, it stores pictures of many types of lesions, both malignant and benign. At this stage of the project, I didn't separate the types. The network should identify if the lesion poses a significant risk or not.

The database originally consists of 5 subsets:

- **ISIC_UDA-1:** 557 lesion pictures contributed by 12 lesion clinics. Cutaneous melanocytic lesions from adults, mostly benign.
- **ISIC_UDA-2_1:** 44 high quality lesions of melanomas and benign lesions. Lots of metadata are available, which are not relevant to the project.
- **ISIC_MSK-1:** Pictures of various types of lesions, both benign and malignant
- **ISIC_MSK-2:** Pictures of various types of lesions, both benign and malignant.
- **ISIC_SONIC:** 9251 pictures of moles in children, mostly benign.

I organized them into only 2 categories: malignant and non-malignant. Then I separated part of the dataset as a test set so there is no overlap with my training images.

The images need processing to be usable. Most of them are high resolution, but until the network is proven usable, a very low resolution is preferred. Early teaching is done on 64x64 pixel pictures, which allows faster results, while showing if the network is learning anything. Also, in fully connected networks the image resolution heavily impacts the number of parameters in the network, which arises the danger of overfitting. Convolutional networks are safer in this regard. Some pictures also have to be cropped, because of irrelevant information next to the lesion.

Future plans:

As this is a 2 semester project, I have a list of the work to be done. I have a yet untrained network so far and the dataset. My plans for the next semester:

- Training the network with the current topology
- Try different topologies (different set of layers) and compare their performance
- Increase the size of my datasets by modifying the existing pictures: adding noise and slight modifications (for example hairs)

5 References

[1] https://en.wikipedia.org/wiki/Melanoma

[2] https://en.wikipedia.org/wiki/Artificial_neural_network

[3] https://en.wikibooks.org/wiki/Artificial_Neural_Networks/Print_Version

[4] https://www.nvidia.com/object/what-is-gpu-computing.html

[5] https://www.nvidia.com/object/deep-learning.html

[6] https://cs231n.github.io/convolutional-networks/

[7] http://torch.ch/

[8] https://isic-archive.com/#