

## Automatic Monitoring for Cyberspace Applications

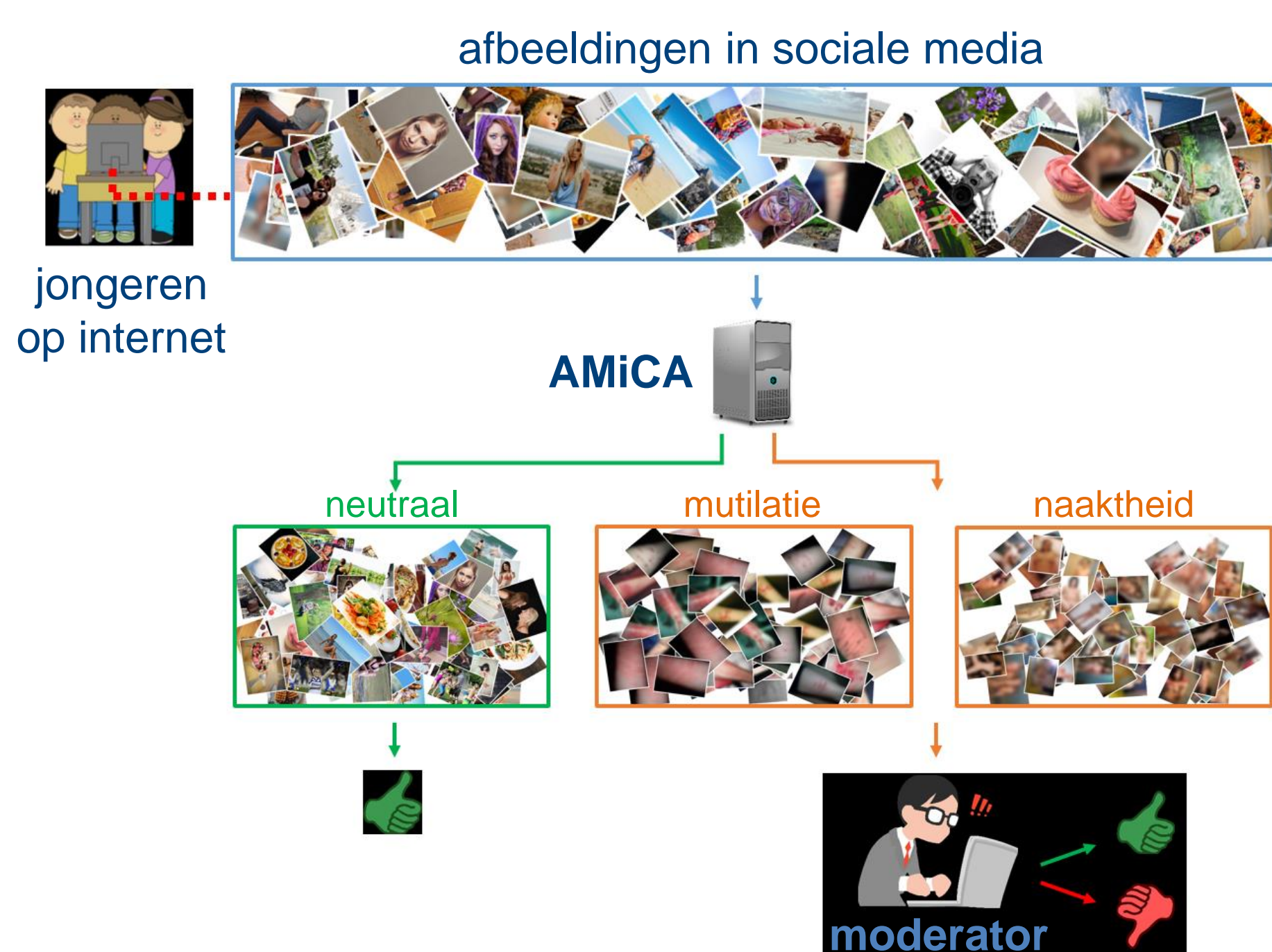
### Inleiding

Het **AMiCA**-project heeft tot doel om bedreigende inhoud op sociale media platformen op te sporen.

Voor beeldmateriaal wordt ernaar gestreefd om beelden te karakteriseren op basis van hun inhoud, meer bepaald:

- ▷ beelden van potentieel pornografische aard
- ▷ beelden van (zelf)mutilatie (wat zou kunnen wijzen op intenties tot zelfdoding)

### Doel

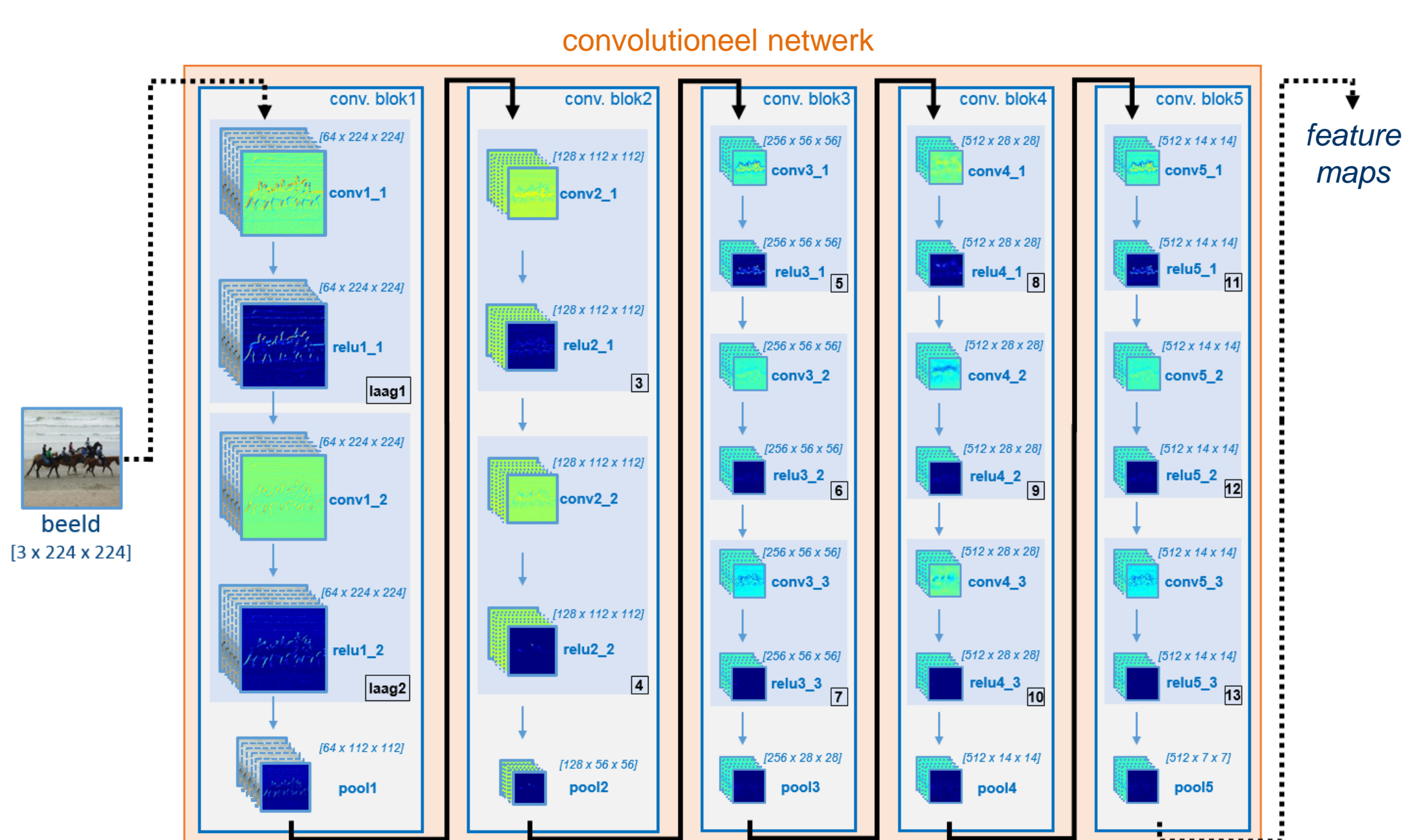


### Dataset

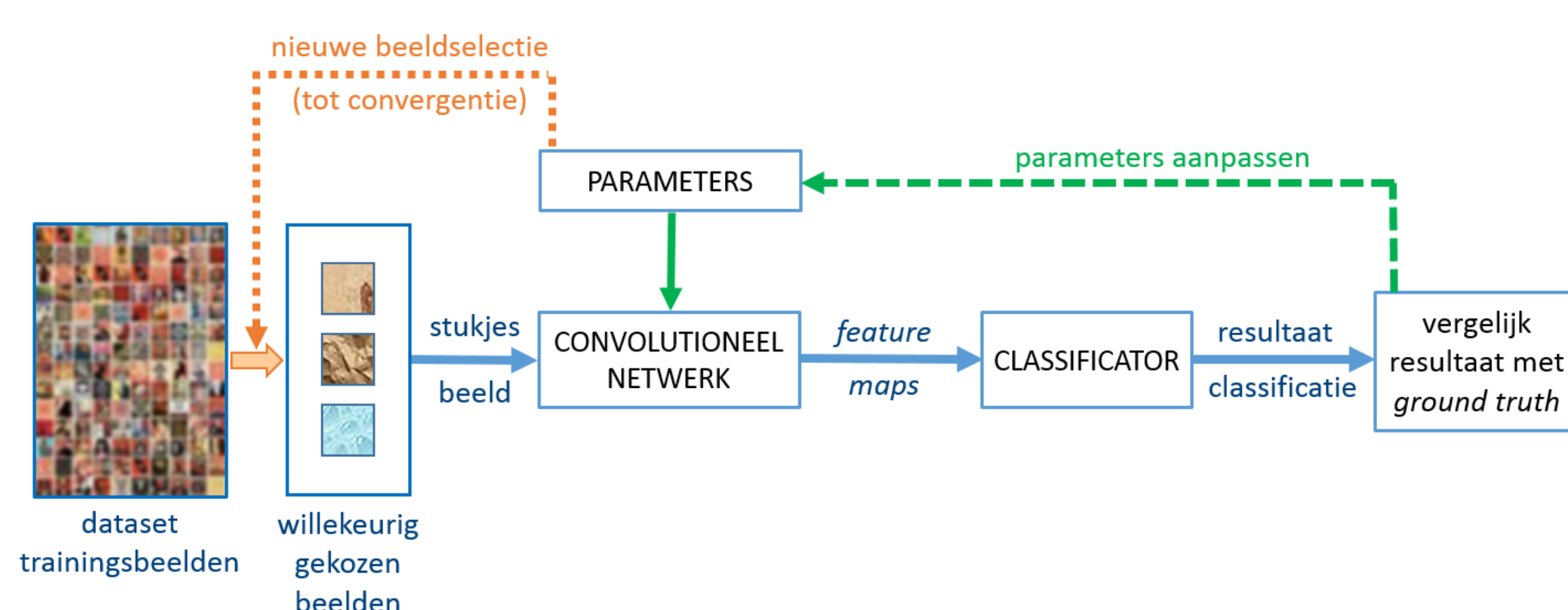


### Methodie ⇒ diepe convolutieele netwerken

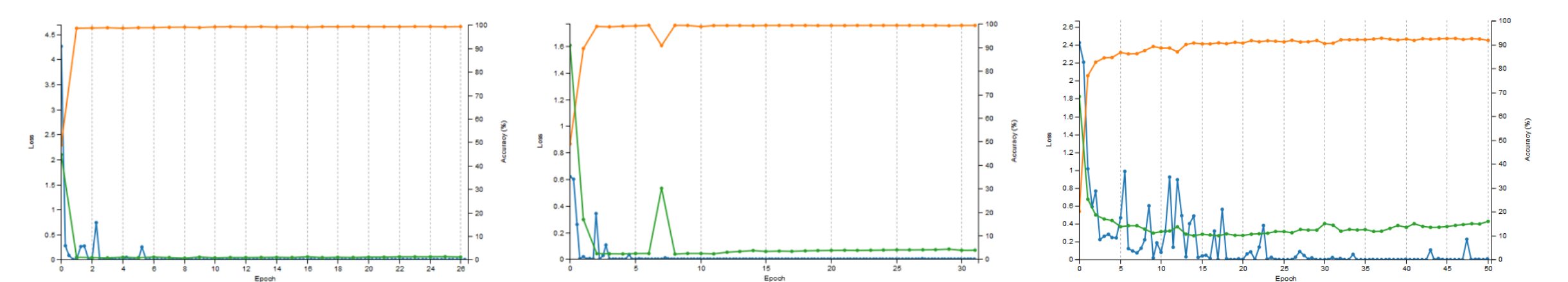
#### a) architectuur van het netwerk: een opeenvolging van verschillende convolutieele en pooling lagen



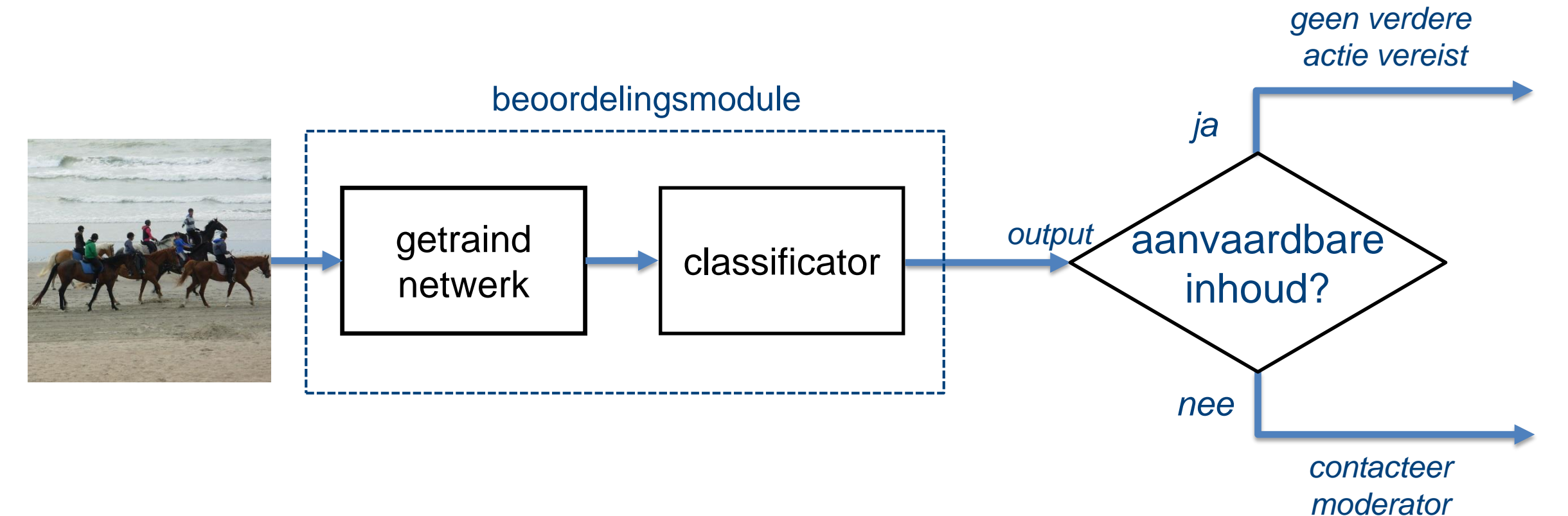
#### b) leerfase: berekenen van geschikte convolutiefilters



### evolutie van loss/accuracy tijdens het trainen van een netwerk



#### c) operationele fase: classificeren van nieuwe beelden



### Voorbeelden

Example 1: *De drie Gratiën naar het antieke Romeinse Beeld in de Borghese verzameling te Rome, Jean-Etienne Liotard, 1737 (Collectie Rijksmuseum)*  
 mutilatie: 3,4e-10  
 naaktheid: 0,999999  
 neutraal: 1,8e-08  
**naaktheid!**

Example 2: *Krijgsman met getrokken zwaard., Joannes Echarius Carolus Alberti, 1808 (Collectie Rijksmuseum)*  
 mutilatie: 1,6e-07  
 naaktheid: 0,999995  
 neutraal: 4,8e-06  
**naaktheid!**

Example 3: *[Close-up of a face]*  
 mutilatie: 0,999999  
 naaktheid: 4,1e-14  
 neutraal: 6,2e-15  
**mutilatie !**

Example 4: *[Person on a beach]*  
 mutilatie: 2,1e-10  
 naaktheid: 8,1e-07  
 neutraal: 0,999999  
**neutraal**

### Experimentele resultaten

#### Trainingsdataset

- 20.000 beelden die een onaanvaardbare graad van naaktheid bevatten
- 20.000 neutrale beelden

#### Validatiedataset

- 10.000 beelden die een onaanvaardbare graad van naaktheid bevatten
- 10.000 neutrale beelden

#### Testdataset

- 4.000 beelden die een onaanvaardbare graad van naaktheid bevatten (en niet gebruikt werden voor training)
- 70.000 neutrale beelden (die niet gebruikt werden voor training)

⇒ beoordeling:

ground truth	geklasseerd als	
	naaktheid	neutraal
naaktheid	97,63%	2,37%
neutraal	0,48%	99,52%

#### Evaluatietijd (FYI: PC met Intel(R) Core(TM) i7, NVIDIA GeForce GTX 970)

- beelden [256x256] → op CPU ± 830 ms/beeld; op GPU ± 40 ms/beeld
- beelden [640x640] → op CPU ± 890 ms/beeld; op GPU ± 105 ms/beeld