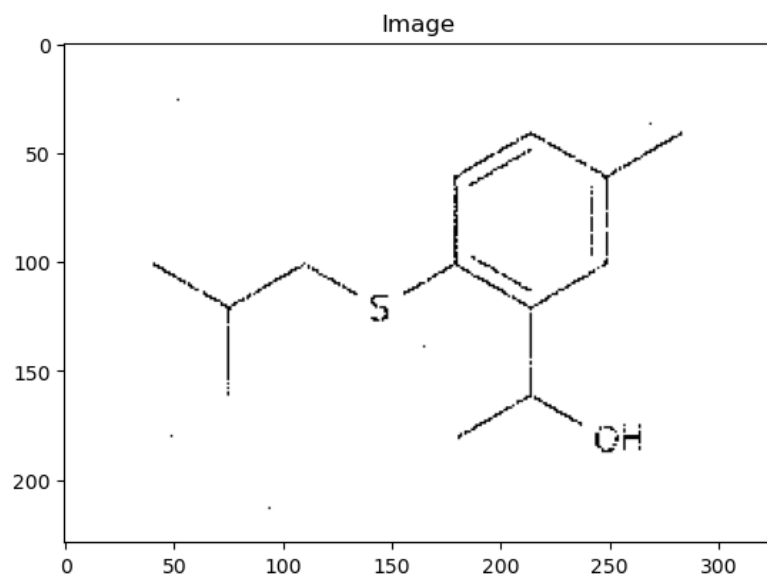# Chemistr AI – Architecture (Target)

Description

- The target is to predict the INCHI from the skeletal chemical formula

Image

InChI=1S/C13H20OS/c1-9(2)8-15-13-6-5-10(3)7-12(13)11(4)14/h5-7,9,11,14H,8H2,1-4H3

The solution is composed of 3 parts and in the following order:
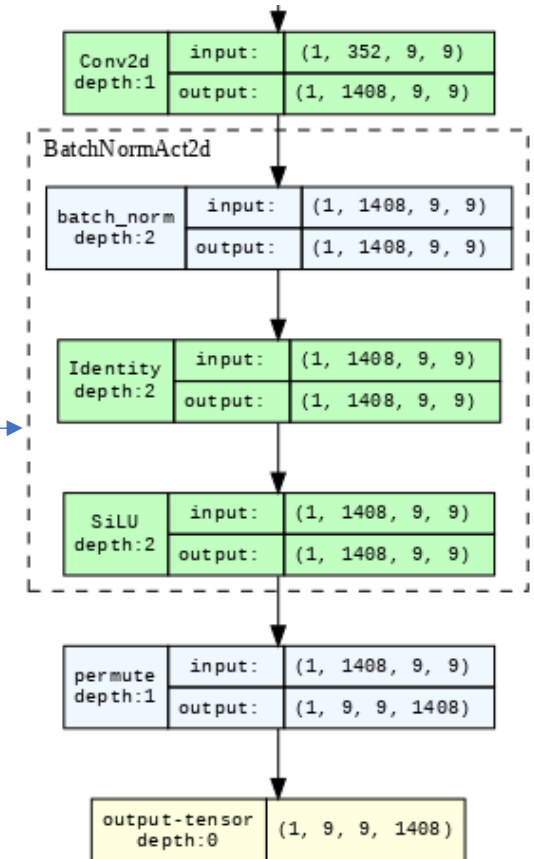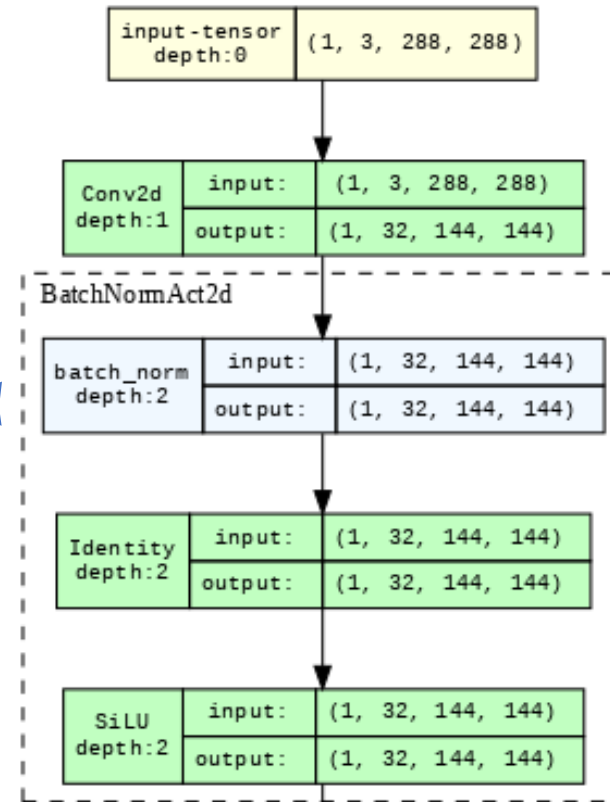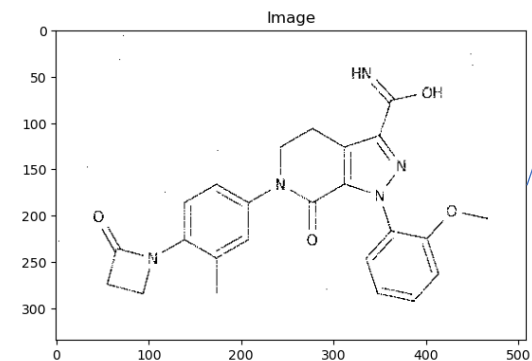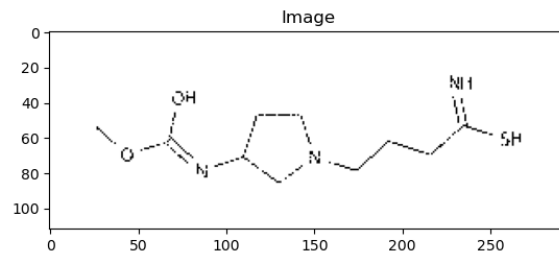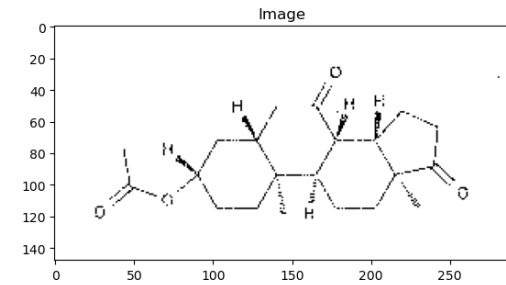
Encoder

Transform the images into features

Decoder

Transform the features to vectors
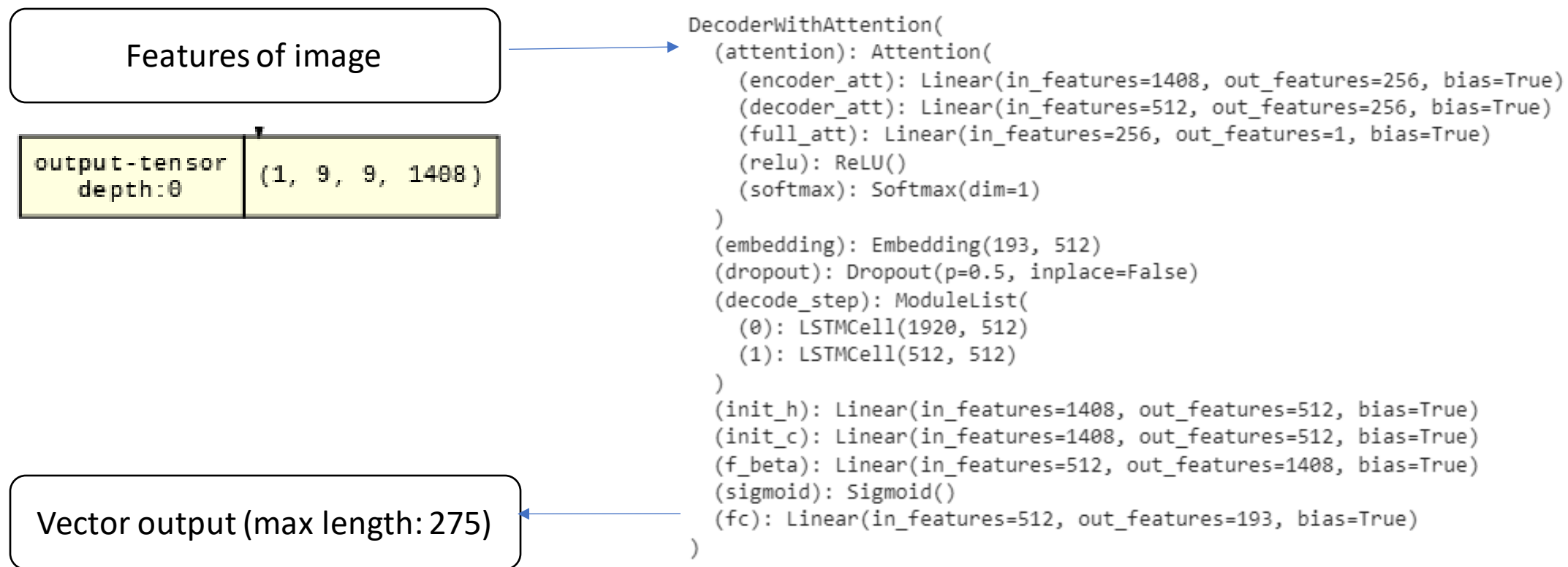
Tokenizer

Transform the vectors to a sequence of text

Chemistr AI – Architecture (Encoder)

Chemistr AI – Architecture (Decoder)

Features of image

output-tensor
depth:0   (1, 9, 9, 1408)

Vector output (max length: 275)

```
DecoderWithAttention(
  (attention): Attention(
    (encoder_att): Linear(in_features=1408, out_features=256, bias=True)
    (decoder_att): Linear(in_features=512, out_features=256, bias=True)
    (full_att): Linear(in_features=256, out_features=1, bias=True)
    (relu): ReLU()
    (softmax): Softmax(dim=1)
  )
  (embedding): Embedding(193, 512)
  (dropout): Dropout(p=0.5, inplace=False)
  (decode_step): ModuleList(
    (0): LSTMCell(1920, 512)
    (1): LSTMCell(512, 512)
  )
  (init_h): Linear(in_features=1408, out_features=512, bias=True)
  (init_c): Linear(in_features=1408, out_features=512, bias=True)
  (f_beta): Linear(in_features=512, out_features=1408, bias=True)
  (sigmoid): Sigmoid()
  (fc): Linear(in_features=512, out_features=193, bias=True)
)
```

Chemistr AI – Architecture (Tokenizer)

Vector output
(max length: 275)

Tokenizer

Transform the
vectors to a
sequence of text

Prediction:
InChI=1S/C11H16BrN4/c1-8-6-14-11(16-9)9(15-2)9(15-3)4-5-14-7-13-8/h4-5,8,14H,6H2,1-3H3