

Lab 3 - Création de simples auto-encodeurs avec Keras

Dans ce laboratoire, nous répondrons à quelques questions courantes sur les encodeurs automatiques et nous couvrirons quelques exemples de code pour les modèles suivants:

- un simple auto-encodeur basé sur une couche entièrement connectée
- un auto-encodeur profond entièrement connecté
- un autoencodeur à convolution profonde
- un modèle de débruitage d'image

Remarque: Vous aurez besoin de Keras version 2.0.0 ou supérieure pour les exécuter.

3.1 Que sont les auto-encodeurs?

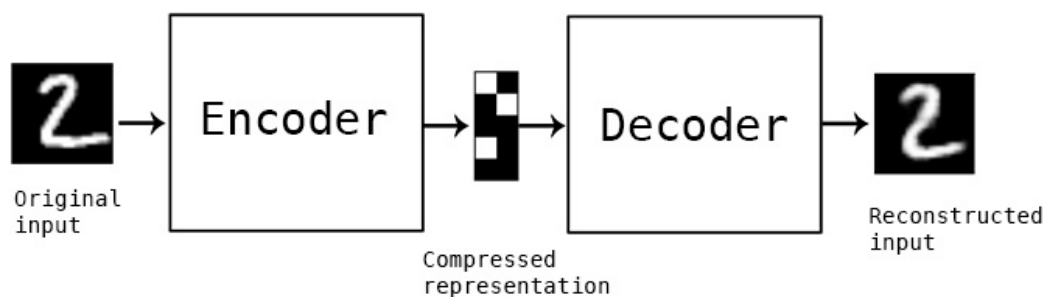


Figure 3.1 Architecture d'un auto-encodeur

Auto-encoding est un **algorithme de compression de données** dont les fonctions de compression et de décompression sont les suivantes:

1. spécifique aux **données**,
2. avec **perte**, et
3. appris **automatiquement** à partir d'exemples plutôt que conçu par un humain.

Les fonctions de compression et de décompression sont implémentées avec des réseaux de neurones.

1) Les auto-encodeurs sont spécifiques aux **données**, ce qui signifie qu'ils ne pourront compresser que des données similaires à celles sur lesquelles ils ont été formés. C'est différent, disons, de l'algorithme de compression MP3, qui ne contient que des hypothèses sur le **son** en général, mais pas sur des types de sons spécifiques. Un autoencodeur formé sur les images de visages ferait un travail plutôt médiocre de compression des images d'arbres, car les fonctionnalités qu'il apprendrait seraient spécifiques aux visages.

2) Les auto-encodeurs sont avec **perte**, ce qui signifie que les sorties décompressées seront dégradées par rapport aux entrées d'origine (similaire à la compression MP3 ou JPEG).

3) Les auto-encodeurs apprennent automatiquement à partir d'exemples de données. Il est facile de former des instances spécialisées de l'algorithme qui fonctionneront bien sur un type spécifique d'entrée.

Pour créer un encodeur automatique, vous avez besoin de trois composantes:

1. une **fonction d'encodage**,
2. une **fonction de décodage**, et
3. une **fonction de distance (*loss*)** entre la quantité de perte d'informations entre la représentation compressée de vos données et la représentation décompressée.

L'encodeur et le décodeur seront choisis pour être des fonctions paramétriques et différenciables par rapport à la fonction de distance, de sorte que les paramètres des fonctions d'encodage / décodage peuvent être optimisés pour minimiser la perte de reconstruction, en utilisant la **descente de gradient stochastique**.

C'est simple! Et vous n'avez même pas besoin de comprendre ces mots pour commencer à utiliser les encodeurs automatiques dans la pratique.

3.1.1 Sont-ils bons en compression de données?

Habituellement, pas vraiment. Dans la compression d'image par exemple, il est assez difficile de former un encodeur automatique qui fait un meilleur travail qu'un algorithme de base comme JPEG, et généralement le seul moyen d'y parvenir est de se limiter à un type d'image très spécifique. Le fait que les auto-encodeurs soient **spécifiques aux données** les rend généralement peu pratiques pour les problèmes de compression de données du monde réel: vous ne pouvez les utiliser que sur des données similaires à celles pour lesquelles ils ont été formés, et **les rendre plus généraux nécessite donc beaucoup de données de formation**.

Mais les avancées futures pourraient changer cela, qui sait!

3.1.2 À quoi servent les encodeurs automatiques?

Aujourd'hui, deux applications pratiques intéressantes des encodeurs automatiques sont le **débruitage des données (*denoising*)** (que nous présenterons plus loin dans ce laboratoire) et la réduction de dimensionnalité pour la visualisation des données. Avec des contraintes de dimensionnalité et de rareté appropriées, les auto-encodeurs peuvent apprendre des projections de données qui sont plus intéressantes que l'**ACP (*Principal component analysis -PCA*)** ou d'autres techniques de base.

L'une des raisons pour lesquelles les auto-encodeurs ont attiré tant de recherches et d'attention est qu'ils ont longtemps été considérés comme une voie potentielle pour résoudre le problème de l'**apprentissage non supervisé**, c'est-à-dire l'apprentissage de représentations utiles sans avoir besoin d'étiquettes.

3.2 Encodeur automatique le plus simple possible

Nous allons commencer simple, avec une seule couche neuronale entièrement connectée comme encodeur et décodeur.

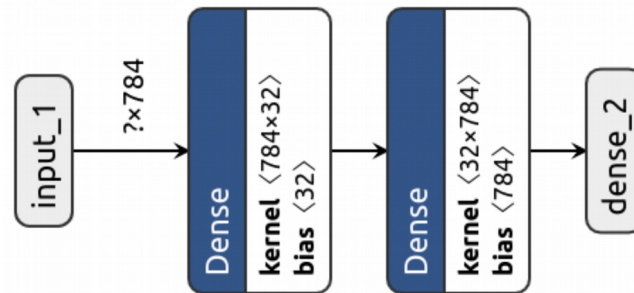


Figure 3.2 Modèle d'un simple auto-encodeur à 2 couches : **encoded** et **decoded**

Pour le comprendre lisez les commentaires associés aux lignes du code.

```
from keras.layers import Input, Dense
from keras.models import Model

# this is the size of our encoded representations
encoding_dim = 32
# 32 floats -> compression of factor 24.5, assuming the input is 784 floats

# this is our input placeholder
input_img = Input(shape=(784,))

# "encoded" is the encoded representation of the input
encoded = Dense(encoding_dim, activation='relu')(input_img)

# "decoded" is the lossy reconstruction of the input
decoded = Dense(784, activation='sigmoid')(encoded)

# this model maps an input to its reconstruction
autoencoder = Model(input_img, decoded)
# this model maps an input to its encoded representation
encoder = Model(input_img, encoded)

# create a placeholder for an encoded (32-dimensional) input
encoded_input = Input(shape=(encoding_dim,))

# retrieve the last layer of the autoencoder model
decoder_layer = autoencoder.layers[-1]

# create the decoder model
decoder = Model(encoded_input, decoder_layer(encoded_input))

autoencoder.compile(optimizer='adadelta', loss='binary_crossentropy')

autoencoder.summary()

from keras.datasets import mnist
import numpy as np
(x_train, _), (x_test, _) = mnist.load_data()

x_train = x_train.astype('float32') / 255.
x_test = x_test.astype('float32') / 255.
x_train = x_train.reshape((len(x_train), np.prod(x_train.shape[1:])))
x_test = x_test.reshape((len(x_test), np.prod(x_test.shape[1:])))

autoencoder.fit(x_train, x_train,
                epochs=20,
                batch_size=64,
                shuffle=True,
                validation_data=(x_test, x_test))
```

```

# encode and decode some digits
# note that we take them from the *test* set
encoded_imgs = encoder.predict(x_test)
decoded_imgs = decoder.predict(encoded_imgs)

# use Matplotlib
import matplotlib.pyplot as plt

n = 10 # how many digits we will display
plt.figure(figsize=(20, 4))
for i in range(n):
    # display original
    ax = plt.subplot(2, n, i + 1)
    plt.imshow(x_test[i].reshape(28, 28))
    plt.gray()
    ax.get_xaxis().set_visible(False)
    ax.get_yaxis().set_visible(False)

    # display reconstruction
    ax = plt.subplot(2, n, i + 1 + n)
    plt.imshow(decoded_imgs[i].reshape(28, 28))
    plt.gray()
    ax.get_xaxis().set_visible(False)
    ax.get_yaxis().set_visible(False)
plt.show()

```



Figure 3.3 Images avant encodage et après encodage-décodage

3.2.1 Exercice

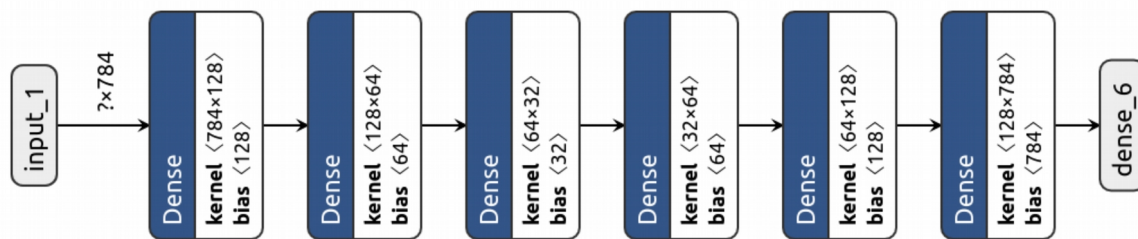
La fonction de `model.fit()` lance l'entraînement du modèle avec le nombre de phases (**epoch**) d'entraînement et la taille du **batch**

Modifiez ces paramètres avec `epochs=50`, et `batch_size=256`, et comparez les résultats d'affichage.

3.3 Un auto-encodeur profond

Nous n'avons pas à nous limiter à une seule couche comme **encodeur** ou **décodeur**, nous pourrions plutôt utiliser une pile de couches, telles que:

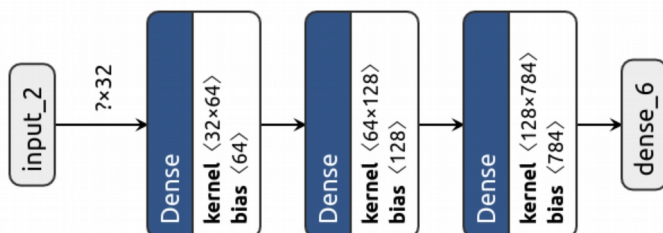
```
input_img = Input(shape=(784,))
encoded = Dense(128, activation='relu')(input_img)
encoded = Dense(64, activation='relu')(encoded)
encoded = Dense(32, activation='relu')(encoded)
decoded = Dense(64, activation='relu')(encoded)
decoded = Dense(128, activation='relu')(decoded)
decoded = Dense(784, activation='sigmoid')(decoded)
```



```
autoencoder = Model(input_img, decoded)
encoder = Model(input_img, encoded)
# separate Decoder model
# create a placeholder for an encoded (32-dimensional) input
encoded_input = Input(shape=(encoding_dim, ))

# retrieve the layers of the autoencoder model
decoder_layer1 = autoencoder.layers[-3]
decoder_layer2 = autoencoder.layers[-2]
decoder_layer3 = autoencoder.layers[-1]

# create the decoder model
decoder = Model(encoded_input,
decoder_layer3(decoder_layer2(decoder_layer1(encoded_input))))
```



```
autoencoder.compile(optimizer='adadelta', loss='binary_crossentropy')
```

```
autoencoder.fit(x_train, x_train,
epochs=10,
batch_size=64,
shuffle=True,
validation_data=(x_test, x_test))
```

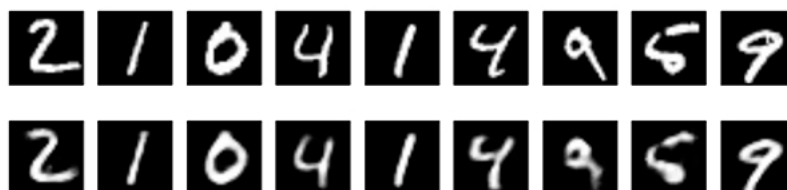


Figure 3.5 Images avant encodage et après encodage-décodage

3.3.1 Code complet de l'auto-encodeur profond

```
from keras.layers import Input, Dense
from keras.models import Model
# this is the size of our encoded representations
encoding_dim = 32
# 32 floats -> compression of factor 24.5, assuming the input is 784 floats
# this is our input placeholder
input_img = Input(shape=(784,))
# encoder part
encoded = Dense(128, activation='relu')(input_img)
encoded = Dense(64, activation='relu')(encoded)
encoded = Dense(32, activation='relu')(encoded)
# decoder part
decoded = Dense(64, activation='relu')(encoded)
decoded = Dense(128, activation='relu')(decoded)
decoded = Dense(784, activation='sigmoid')(decoded)

# separate models for autoencoder, encoder and decoder
autoencoder = Model(input_img, decoded)
encoder = Model(input_img, encoded)
# create a placeholder for an encoded (32-dimensional) input
encoded_input = Input(shape=(encoding_dim,))

# create the decoder model
# retrieve the last layer of the autoencoder model
decoder_layer1 = autoencoder.layers[-3]
decoder_layer2 = autoencoder.layers[-2]
decoder_layer3 = autoencoder.layers[-1]
decoder = Model(encoded_input,
decoder_layer3(decoder_layer2(decoder_layer1(encoded_input))))
decoder.summary()
decoder.save('autoenc_deep_decoder.h5')
autoencoder.compile(optimizer='adadelta', loss='binary_crossentropy')
autoencoder.summary()
autoencoder.save('autoenc_deep.h5')
from keras.datasets import mnist
import numpy as np
(x_train, _), (x_test, _) = mnist.load_data()

x_train = x_train.astype('float32') / 255.
x_test = x_test.astype('float32') / 255.
x_train = x_train.reshape((len(x_train), np.prod(x_train.shape[1:])))
x_test = x_test.reshape((len(x_test), np.prod(x_test.shape[1:])))
autoencoder.fit(x_train, x_train,
epochs=10,
batch_size=64,
shuffle=True,
validation_data=(x_test, x_test))

# encode and decode some digits
# note that we take them from the *test* set
encoded_imgs = encoder.predict(x_test)
decoded_imgs = decoder.predict(encoded_imgs)
import matplotlib.pyplot as plt
n = 10 # how many digits we will display
plt.figure(figsize=(20, 4))
for i in range(n):
    # display original
    ax = plt.subplot(2, n, i + 1)
    plt.imshow(x_test[i].reshape(28, 28))
    plt.gray()
    ax.get_xaxis().set_visible(False)
    ax.get_yaxis().set_visible(False)
    # display reconstruction
    ax = plt.subplot(2, n, i + 1 + n)
    plt.imshow(decoded_imgs[i].reshape(28, 28))
    plt.gray()
    ax.get_xaxis().set_visible(False)
    ax.get_yaxis().set_visible(False)
plt.show()
```

3.3.2 Exercice

La fonction de `autoencoder.fit()` lance l'entraînement du modèle avec le nombre de phases (**epoch**) d'entraînement et la taille du **batch**

Modifiez ces paramètres avec **epochs=50**, et **batch_size=256**, et comparez les résultats d'affichage.

3.4 Auto-encodeur convolutif

Puisque nos entrées sont des images, il est logique d'utiliser des réseaux de neurones convolutionnels (**convnets**) comme **encodeurs** et **décodeurs**. Dans les paramètres pratiques, les auto-encodeurs appliqués aux images sont toujours des encodeurs automatiques convolutifs - ils fonctionnent tout simplement beaucoup mieux.

Notre **encodeur** consiste en une pile de couches **Conv2D** et **MaxPooling2D** (le regroupement maximal étant utilisé pour le sous-échantillonnage spatial), tandis que le **décodeur** consiste en une pile de couches **Conv2D** et **UpSampling2D**.

```
from keras.layers import Input, Dense, Conv2D, MaxPooling2D, UpSampling2D
from keras.models import Model
from keras import backend as K

input_img = Input(shape=(28, 28, 1)) # adapt this if using `channels_first`
image data format

x = Conv2D(16, (3, 3), activation='relu', padding='same')(input_img)
x = MaxPooling2D((2, 2), padding='same')(x)
x = Conv2D(8, (3, 3), activation='relu', padding='same')(x)
x = MaxPooling2D((2, 2), padding='same')(x)
x = Conv2D(8, (3, 3), activation='relu', padding='same')(x)
encoded = MaxPooling2D((2, 2), padding='same')(x)

# at this point the representation is (4, 4, 8) i.e. 128-dimensional
x = Conv2D(8, (3, 3), activation='relu', padding='same')(encoded)
x = UpSampling2D((2, 2))(x)
x = Conv2D(8, (3, 3), activation='relu', padding='same')(x)
x = UpSampling2D((2, 2))(x)
x = Conv2D(16, (3, 3), activation='relu')(x)
x = UpSampling2D((2, 2))(x)
decoded = Conv2D(1, (3, 3), activation='sigmoid', padding='same')(x)

autoencoder = Model(input_img, decoded)
autoencoder.compile(optimizer='adadelata', loss='binary_crossentropy')
```



Pour l'entraîner, nous utiliserons les chiffres MNIST d'origine avec forme (**sample, 3, 28, 28**), et nous normaliserons simplement les valeurs de pixels entre 0 et 1.

```
from keras.datasets import mnist
import numpy as np

(x_train, _), (x_test, _) = mnist.load_data()

x_train = x_train.astype('float32') / 255.
x_test = x_test.astype('float32') / 255.
x_train = np.reshape(x_train, (len(x_train), 28, 28, 1)) # adapt this if using
`channels_first` image data format
x_test = np.reshape(x_test, (len(x_test), 28, 28, 1)) # adapt this if using
`channels_first` image data format
```

Entraînons ce modèle pour 50 époques.


```

autoencoder.fit(x_train, x_train,
                epochs=50,
                batch_size=128,
                shuffle=True,
                validation_data=(x_test, x_test))

```

Le modèle converge vers une perte de **0,094**, nettement meilleure que nos modèles précédents (ceci est en grande partie dû à la capacité entropique plus élevée de la représentation codée, 128 dimensions contre 32 précédemment). Jetons un coup d'œil aux chiffres reconstruits:

```

decoded_imgs = autoencoder.predict(x_test)

n = 10
plt.figure(figsize=(20, 4))
for i in range(n):
    # display original
    ax = plt.subplot(2, n, i)
    plt.imshow(x_test[i].reshape(28, 28))
    plt.gray()
    ax.get_xaxis().set_visible(False)
    ax.get_yaxis().set_visible(False)
    # display reconstruction
    ax = plt.subplot(2, n, i + n)
    plt.imshow(decoded_imgs[i].reshape(28, 28))
    plt.gray()
    ax.get_xaxis().set_visible(False)
    ax.get_yaxis().set_visible(False)
plt.show()

```

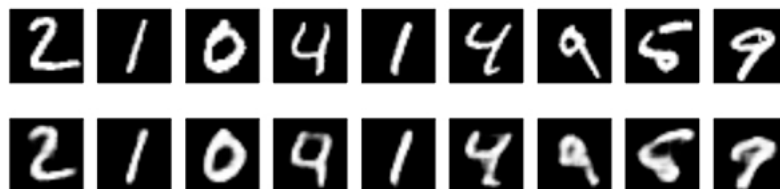


Figure 3.6 Images avant encodage et après encodage-décodage (modèle convolutive)

Nous pouvons également jeter un œil aux représentations codées à 128 ou 64 dimensions. Ces représentations sont 8x4x4, nous les remodelons donc en 4x32 afin de pouvoir les afficher sous forme d'images en niveaux de gris.

```

n = 10
plt.figure(figsize=(20, 8))
for i in range(n):
    ax = plt.subplot(1, n, i)
    plt.imshow(encoded_imgs[i].reshape(4, 4 * 8).T) # 8 => 4
    plt.gray()
    ax.get_xaxis().set_visible(False)
    ax.get_yaxis().set_visible(False)
plt.show()

```

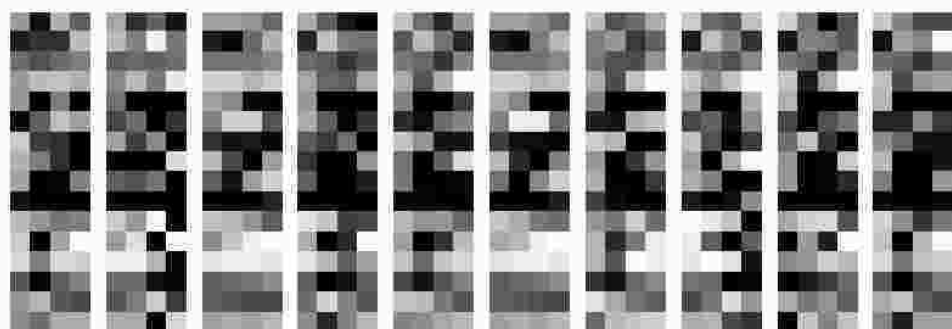


Figure 3.7 Images codées (modèle convolutive) : format réduit (4,4,4) : 16*4

3.4.1 Code complet de l'auto-encodeur convolutif

```
from keras.layers import Input, Dense, Conv2D, MaxPooling2D, UpSampling2D
from keras.models import Model
from keras import backend as K
input_img = Input(shape=(28, 28, 1))

# adapt this if using `channels_first` image data format
x = Conv2D(8, (3, 3), activation='relu', padding='same')(input_img)
x = MaxPooling2D((2, 2), padding='same')(x)
x = Conv2D(4, (3, 3), activation='relu', padding='same')(x)
x = MaxPooling2D((2, 2), padding='same')(x)
x = Conv2D(4, (3, 3), activation='relu', padding='same')(x)
encoded = MaxPooling2D((2, 2), padding='same')(x)

# at this point the representation is (4, 4, 8) i.e. 128-dimensional
x = Conv2D(4, (3, 3), activation='relu', padding='same')(encoded)
x = UpSampling2D((2, 2))(x)
x = Conv2D(4, (3, 3), activation='relu', padding='same')(x)
x = UpSampling2D((2, 2))(x)
x = Conv2D(8, (3, 3), activation='relu')(x)
x = UpSampling2D((2, 2))(x)
decoded = Conv2D(1, (3, 3), activation='sigmoid', padding='same')(x)

autoencoder = Model(input_img, decoded)
autoencoder.compile(optimizer='adadelta', loss='binary_crossentropy')

autoencoder.summary()

from keras.datasets import mnist
import numpy as np
(x_train, _), (x_test, _) = mnist.load_data()
x_train = x_train.astype('float32') / 255.
x_test = x_test.astype('float32') / 255.
x_train = np.reshape(x_train, (len(x_train), 28, 28, 1))
# adapt this if using `channels_first` image data format
x_test = np.reshape(x_test, (len(x_test), 28, 28, 1))
# adapt this if using `channels_first` image data format

autoencoder.fit(x_train, x_train,
               epochs=10,
               batch_size=64,
               shuffle=True,
               validation_data=(x_test, x_test), verbose=1)

autoencoder.save('autoenc_conv.h5')
# encode and decode some digits
# note that we take them from the *test* set
decoded_imgs = autoencoder.predict(x_test)

import matplotlib.pyplot as plt

n = 10 # how many digits we will display
plt.figure(figsize=(20, 4))
for i in range(n):
    # display original
    ax = plt.subplot(2, n, i + 1)
    plt.imshow(x_test[i].reshape(28, 28))
    plt.gray()
    ax.get_xaxis().set_visible(False)
    ax.get_yaxis().set_visible(False)

    # display reconstruction
    ax = plt.subplot(2, n, i + 1 + n)
    plt.imshow(decoded_imgs[i].reshape(28, 28))
    plt.gray()
    ax.get_xaxis().set_visible(False)
    ax.get_yaxis().set_visible(False)
plt.show()
```

```

encoder = Model(input_img, encoded)
encoded_imgs = encoder.predict(x_test)
plt.figure(figsize=(10, 4), dpi=100)
for i in range(n):
    ax = plt.subplot(1, n, i + 1)
    plt.imshow(encoded_imgs[i].reshape(4, 4*4).T) # 8 => 4
    plt.gray()
    ax.set_axis_off()
plt.show()

```

3.4.2 Exercice

La fonction de `autoencoder.fit()` lance l'entraînement du modèle avec le nombre de phases (**epoch**) d'entraînement et la taille du **batch**

Modifiez ces paramètres avec **epochs=40**, et **batch_size=256**, et comparez les résultats d'affichage.

3.5 Application au débruitage d'image

Mettons notre encodeur automatique convolutionnel **au travail** sur un problème de débruitage d'image. C'est simple: nous allons former l'auto-encodeur pour mapper des images de chiffres bruités à des images de chiffres propres.

Voici comment nous allons générer des chiffres bruités : nous appliquons simplement une matrice de bruit gaussienne et projetons les images sur les valeurs entre 0 et 1.

```
from keras.datasets import mnist
import numpy as np

(x_train, _), (x_test, _) = mnist.load_data()

x_train = x_train.astype('float32') / 255.
x_test = x_test.astype('float32') / 255.
x_train = np.reshape(x_train, (len(x_train), 28, 28, 1)) # adapt this if using
`channels_first` image data format
x_test = np.reshape(x_test, (len(x_test), 28, 28, 1)) # adapt this if using
`channels_first` image data format

noise_factor = 0.5
x_train_noisy = x_train + noise_factor * np.random.normal(loc=0.0, scale=1.0,
size=x_train.shape)
x_test_noisy = x_test + noise_factor * np.random.normal(loc=0.0, scale=1.0,
size=x_test.shape)

x_train_noisy = np.clip(x_train_noisy, 0., 1.)
x_test_noisy = np.clip(x_test_noisy, 0., 1.)
```

Voici à quoi ressemblent les chiffres bruités:

```
n = 10
plt.figure(figsize=(20, 2))
for i in range(n):
    ax = plt.subplot(1, n, i)
    plt.imshow(x_test_noisy[i].reshape(28, 28))
    plt.gray()
    ax.get_xaxis().set_visible(False)
    ax.get_yaxis().set_visible(False)
plt.show()
```



Figure 3.8 Images bruitées avant l'encodage

Si vous plissez les yeux, vous pouvez toujours les reconnaître, mais à peine. Notre encodeur automatique peut-il apprendre à récupérer les chiffres d'origine? Découvrons-le.

Par rapport au précédent auto-encodeur convolutif, afin d'améliorer la qualité de la reconstruction, nous utiliserons un modèle légèrement différent avec plus de filtres par couche:

Attention

L'entraînement de ce modèle avec les 32 filtres n'est pas possibles sur la carte Jetson Nano dont la RAM partagée (CPU et GPU) est limitée à 4Go.

Dans le code complet nous proposons seulement 8 filtres pour les couches **Conv2D**.

```
input_img = Input(shape=(28, 28, 1)) # adapt this if using `channels_first`
image data format

x = Conv2D(32, (3, 3), activation='relu', padding='same')(input_img)
x = MaxPooling2D((2, 2), padding='same')(x)
x = Conv2D(32, (3, 3), activation='relu', padding='same')(x)
```

```

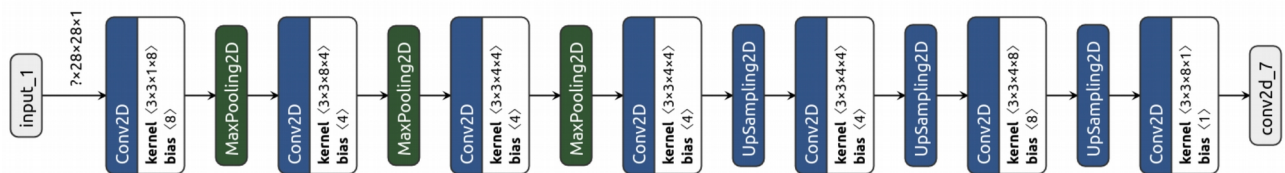
encoded = MaxPooling2D((2, 2), padding='same')(x)
# at this point the representation is (7, 7, 32)

x = Conv2D(32, (3, 3), activation='relu', padding='same')(encoded)
x = UpSampling2D((2, 2))(x)
x = Conv2D(32, (3, 3), activation='relu', padding='same')(x)
x = UpSampling2D((2, 2))(x)
decoded = Conv2D(1, (3, 3), activation='sigmoid', padding='same')(x)

autoencoder = Model(input_img, decoded)
autoencoder.compile(optimizer='adadelata', loss='binary_crossentropy')
Let's train it for 100 epochs:

autoencoder.fit(x_train_noisy, x_train,
                epochs=100,
                batch_size=128,
                shuffle=True,
                validation_data=(x_test_noisy, x_test),
                callbacks=[TensorBoard(log_dir='/tmp/tb', histogram_freq=0,
write_graph=False)])

```



Voyons maintenant les résultats. En haut, les chiffres bruités qui alimentent le réseau, et en bas, les chiffres reconstruits par le réseau.



Cela semble assez bien fonctionner

3.5.1 Code complet de l'auto-encodeur convolutif pour le débruitage

```

from tensorflow import keras
from keras.layers import Input, Dense, Conv2D, MaxPooling2D, UpSampling2D
from keras.models import Model
import sys

input_img = Input(shape=(28, 28, 1)) # adapt this if using `channels_first`
image data format

x = Conv2D(8, (3, 3), activation='relu', padding='same')(input_img)
x = MaxPooling2D((2, 2), padding='same')(x)
x = Conv2D(4, (3, 3), activation='relu', padding='same')(x)
x = MaxPooling2D((2, 2), padding='same')(x)
x = Conv2D(4, (3, 3), activation='relu', padding='same')(x)
encoded = MaxPooling2D((2, 2), padding='same')(x)

# at this point the representation is (4, 4, 8) i.e. 128-dimensional
x = Conv2D(4, (3, 3), activation='relu', padding='same')(encoded)
x = UpSampling2D((2, 2))(x)
x = Conv2D(4, (3, 3), activation='relu', padding='same')(x)
x = UpSampling2D((2, 2))(x)
x = Conv2D(8, (3, 3), activation='relu')(x)
x = UpSampling2D((2, 2))(x)
decoded = Conv2D(1, (3, 3), activation='sigmoid', padding='same')(x)

```

```

autoencoder = Model(input_img, decoded)
autoencoder.compile(optimizer='adadelat', loss='binary_crossentropy')

autoencoder.summary()
from keras.datasets import mnist
import numpy as np

(x_train, _), (x_test, _) = mnist.load_data()
x_train = x_train.astype('float32') / 255.
x_test = x_test.astype('float32') / 255.
x_train = np.reshape(x_train, (len(x_train), 28, 28, 1)) # adapt this if using
`channels_first` image data format
x_test = np.reshape(x_test, (len(x_test), 28, 28, 1)) # adapt this if using
`channels_first` image data format

noise_factor = 0.5 # set as argument
x_train_noisy = x_train + noise_factor * np.random.normal(loc=0.0, scale=1.0,
size=x_train.shape)
x_test_noisy = x_test + noise_factor * np.random.normal(loc=0.0, scale=1.0,
size=x_test.shape)

x_train_noisy = np.clip(x_train_noisy, 0., 1.)
x_test_noisy = np.clip(x_test_noisy, 0., 1.)

autoencoder.fit(x_train_noisy, x_train,
                epochs=10,
                batch_size=64,
                shuffle=True,
                validation_data=(x_test, x_test), verbose=1)

autoencoder.save('autoenc_conv_denoise.h5')
# Recreate the exact same model purely from the file
new_model = keras.models.load_model('autoenc_conv_denoise.h5')

# encode and decode some digits
# note that we take them from the *test* set
#decoded_imgs = autoencoder.predict(x_test)
decoded_imgs = new_model.predict(x_test)
import matplotlib.pyplot as plt
n = 10
plt.figure(figsize=(10, 4), dpi=100)
for i in range(n):
    # display noisy
    ax = plt.subplot(2, n, i + 1)
    plt.imshow(x_test_noisy[i].reshape(28, 28))
    plt.gray()
    ax.set_axis_off()
    # display reconstruction
    ax = plt.subplot(2, n, i + n + 1)
    plt.imshow(decoded_imgs[i].reshape(28, 28))
    plt.gray()
    ax.set_axis_off()

plt.show()

```

3.5.2 Exercices

Le paramètre `noise_factor` permet de préciser le **taux de bruitage**.

- Modifiez la ligne du code pour permettre sa modification par un argument du programme – `sys.argv[1]`.
- Lancez l'entraînement avec les valeurs plus importantes : 0.8, 0.9, voir 0.95 est observez le résultat de débruitage.

3.5.3 Sujet final – inférence - avec modèles entraînés sur Jetson Xavier

Le programme suivant est une application complète permettant d'entraîner et de sauvegarder les nouveaux modèles avec ou sans débruitage sur deux **datasets MNIST** : chiffres (*digit*) et vêtements (*fashion*).

Notre application permet également de charger les modèles externes pour leurs inférence sur les mêmes **datasets**.

Les arguments de l'application sont :

```
Give dataset[1,2,3,4] - digit,fashion,digit+noise,fashion+noise: 1
Train or inference [1,2]: 1
model_name to train/save or load/inference, .h5 will be added:
autoenc_conv_fashion_nano_1
Give width_factor to be multiplied by 8: 1
Give epoch_factor to be multiplied by 10: 1
Give batch_factor to be multiplied by 64:
```

Pour les modèles avec le bruit – le paramètre `noise_factor` [0.0 to 0.99] est demandé

3.5.3.1 Voici le code

```
from keras.layers import Input, Dense, Conv2D, MaxPooling2D, UpSampling2D
from keras.models import Model
from tensorflow import keras
from keras.datasets import fashion_mnist
import sys

ds=int(input("Give dataset[1,2,3,4] - digit,fashion,digit+noise,fashion+noise:
"))
ti=int(input("train or inference [1,2]: "))
mn=input("model_name to train/save or load/inference, .h5 will be added: ")
if ti==1:
    wf=int(input("Give width_factor to be multiplied by 8: "))
    ef=int(input("Give epoch_factor to be multiplied by 10: "))
    bf=int(input("Give batch_factor to be multiplied by 64: "))
if ds>2:
    nf=float(input("Give the noise_factor [0.0 to 0.99]: "))
if ti==1:
    input_img = Input(shape=(28, 28, 1)) # adapt this if using `channels_first`
image data format
    x = Conv2D(wf*8, (3, 3), activation='relu', padding='same')(input_img)
    x = MaxPooling2D((2, 2), padding='same')(x)
    x = Conv2D(wf*4, (3, 3), activation='relu', padding='same')(x)
    x = MaxPooling2D((2, 2), padding='same')(x)
    x = Conv2D(wf*4, (3, 3), activation='relu', padding='same')(x)
    encoded = MaxPooling2D((2, 2), padding='same')(x)
    # at this point the representation is (4, 4, 8) i.e. 128-dimensional
    x = Conv2D(wf*4, (3, 3), activation='relu', padding='same')(encoded)
    x = UpSampling2D((2, 2))(x)
    x = Conv2D(wf*4, (3, 3), activation='relu', padding='same')(x)
    x = UpSampling2D((2, 2))(x)
    x = Conv2D(wf*8, (3, 3), activation='relu')(x)
    x = UpSampling2D((2, 2))(x)
    decoded = Conv2D(1, (3, 3), activation='sigmoid', padding='same')(x)
    autoencoder = Model(input_img, decoded)
    autoencoder.compile(optimizer='adadelata', loss='binary_crossentropy')
    autoencoder.summary()
from keras.datasets import mnist
import numpy as np

if ds==1 or ds==3:
    (x_train, _), (x_test, _) = mnist.load_data()
if ds==2 or ds==4:
    (x_train, _), (x_test, _) = fashion_mnist.load_data()
x_train = x_train.astype('float32') / 255.
x_test = x_test.astype('float32') / 255.
```

```

x_train = np.reshape(x_train, (len(x_train), 28, 28, 1)) # adapt this if using
`channels_first` image data format
x_test = np.reshape(x_test, (len(x_test), 28, 28, 1)) # adapt this if using
`channels_first` image data format
if ds>2:
    x_train_noisy = x_train + nf * np.random.normal(loc=0.0, scale=1.0,
size=x_train.shape)
    x_test_noisy = x_test + nf * np.random.normal(loc=0.0, scale=1.0,
size=x_test.shape)
    x_train_noisy = np.clip(x_train_noisy, 0., 1.)
    x_test_noisy = np.clip(x_test_noisy, 0., 1.)
if ti==1 and ds<3:
    autoencoder.fit(x_train, x_train, epochs=ef*10,
batch_size=bf*64, shuffle=True, validation_data=(x_test, x_test), verbose=1)
    autoencoder.save(mn+'.h5')

if ti==1 and ds>2:
    autoencoder.fit(x_train_noisy, x_train, epochs=ef*10,
batch_size=bf*64, shuffle=True, validation_data=(x_test, x_test), verbose=1)
    autoencoder.save(mn+'.h5')
if ti==2:
    autoencoder = keras.models.load_model(mn+'.h5')
# encode and decode some digits
# note that we take them from the *test* set
# decoded_imgs = autoencoder.predict(x_test)
if ds<3:
    decoded_imgs = autoencoder.predict(x_test)
if ds>2:
    decoded_imgs = autoencoder.predict(x_test_noisy)

import matplotlib.pyplot as plt

n = 10 # how many digits we will display
plt.figure(figsize=(20, 4))
for i in range(n):
    # display original
    ax = plt.subplot(2, n, i + 1)
    if ds<3:
        plt.imshow(x_test[i].reshape(28, 28))
    if ds>2:
        plt.imshow(x_test_noisy[i].reshape(28, 28))

    plt.gray()
    ax.get_xaxis().set_visible(False)
    ax.get_yaxis().set_visible(False)
    # display reconstruction
    ax = plt.subplot(2, n, i + 1 + n)
    plt.imshow(decoded_imgs[i].reshape(28, 28))
    plt.gray()
    ax.get_xaxis().set_visible(False)
    ax.get_yaxis().set_visible(False)
plt.show()

if ti==1:
    encoder = Model(input_img, encoded)
    encoded_imgs = encoder.predict(x_test)
    plt.figure(figsize=(10, 4), dpi=100)
    for i in range(n):
        ax = plt.subplot(1, n, i + 1)
        plt.imshow(encoded_imgs[i].reshape(4, 4*4*wf).T) # nano max 4
        plt.gray()
        ax.set_axis_off()
    plt.show()

```


3.5.3.2 Un exemple d'exécution - entraînement avec les chiffres bruités

```
python3 L3.autoencoder.conv.all.py
```

```
..
Give dataset [1,2,3,4] - digit,fashion,digit+noise,fashion+noise: 3
train or inference [1,2]: 1
model_name to train/save or load/inference, .h5 will be added:
autoenc_conv_denoise_digit_nano
Give width_factor to be multiplied by 8: 1
Give epoch_factor to be multiplied by 10: 1
Give batch_factor to be multiplied by 64: 1
Give the noise_factor [0.0 to 0.99]: 0.8
```

Model: "model_1"

| Layer (type) | Output Shape | Param # |
|-------------------------------|-------------------|---------|
| input_1 (InputLayer) | (None, 28, 28, 1) | 0 |
| conv2d_1 (Conv2D) | (None, 28, 28, 8) | 80 |
| max_pooling2d_1 (MaxPooling2) | (None, 14, 14, 8) | 0 |
| conv2d_2 (Conv2D) | (None, 14, 14, 4) | 292 |
| max_pooling2d_2 (MaxPooling2) | (None, 7, 7, 4) | 0 |
| conv2d_3 (Conv2D) | (None, 7, 7, 4) | 148 |
| max_pooling2d_3 (MaxPooling2) | (None, 4, 4, 4) | 0 |
| conv2d_4 (Conv2D) | (None, 4, 4, 4) | 148 |
| up_sampling2d_1 (UpSampling2) | (None, 8, 8, 4) | 0 |
| conv2d_5 (Conv2D) | (None, 8, 8, 4) | 148 |
| up_sampling2d_2 (UpSampling2) | (None, 16, 16, 4) | 0 |
| conv2d_6 (Conv2D) | (None, 14, 14, 8) | 296 |
| up_sampling2d_3 (UpSampling2) | (None, 28, 28, 8) | 0 |
| conv2d_7 (Conv2D) | (None, 28, 28, 1) | 73 |
| Total params: 1,185 | | |
| Trainable params: 1,185 | | |
| Non-trainable params: 0 | | |

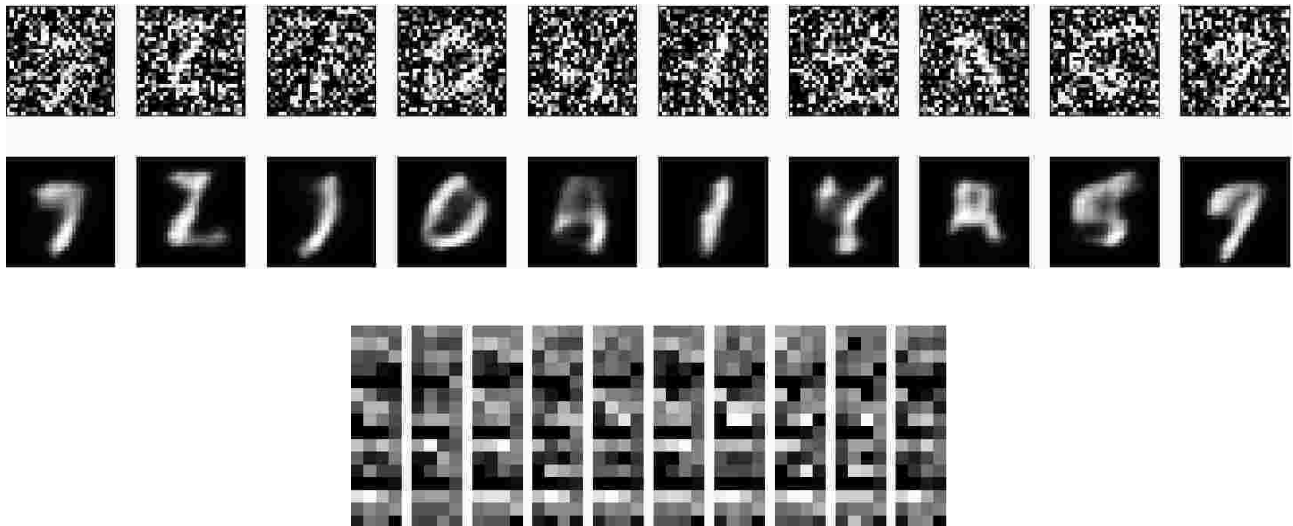
Train on 60000 samples, validate on 10000 samples

```
Epoch 1/10
2020-02-02 16:20:23.816637: I
tensorflow/stream_executor/platform/default/dso_loader.cc:42] Successfully
opened dynamic library libcudnn.so.7
60000/60000 [=====] - 14s 240us/step - loss: 0.2652 -
val_loss: 0.2622
Epoch 2/10
60000/60000 [=====] - 12s 192us/step - loss: 0.2200 -
val_loss: 0.2347
Epoch 3/10
60000/60000 [=====] - 11s 191us/step - loss: 0.2087 -
val_loss: 0.2081
Epoch 4/10
60000/60000 [=====] - 11s 191us/step - loss: 0.2031 -
val_loss: 0.2075
Epoch 5/10
60000/60000 [=====] - 11s 190us/step - loss: 0.1996 -
val_loss: 0.2026
Epoch 6/10
```

```

60000/60000 [=====] - 11s 189us/step - loss: 0.1968 -
val_loss: 0.2158
Epoch 7/10
60000/60000 [=====] - 11s 189us/step - loss: 0.1944 -
val_loss: 0.1893
Epoch 8/10
60000/60000 [=====] - 11s 189us/step - loss: 0.1922 -
val_loss: 0.1782
Epoch 9/10
60000/60000 [=====] - 11s 189us/step - loss: 0.1908 -
val_loss: 0.1780
Epoch 10/10
60000/60000 [=====] - 11s 188us/step - loss: 0.1897 -
val_loss: 0.1788
bako@bako-desktop: /mnt/ssd/tf2$

```



Le résultat de débruitage et la représentation interne du auto-encodeur.

3.5.3.3 Exercices

Exécutez l'application ci-dessus avec :

1. mode inférence et le modèle (digit) sans et avec bruit pour les modèles nano et xavier
2. mode entraînement pour créer le modèle fashion
3. modifiez le code pour pouvoir utiliser le dataset fashion pour le modèle entraîné pour les chiffres

Attention :

Le modèles entraînés sont disponibles sur le site Smartcomputerlab à coté de l'annonce de ce Lab.

3.6 Remarque finale

Si vous faites évoluer le processus de développement vers les **convnet** plus grands, vous pouvez commencer à créer les modèles pour le:

- débruitage de documents,
- coloriage des images ,

ou

- les modèles de débruitage audio.

Table of Contents

| | | |
|---------|--|----|
| 3.1 | Que sont les auto-encodeurs? | 1 |
| 3.1.1 | Sont-ils bons en compression de données? | 2 |
| 3.1.2 | À quoi servent les encodeurs automatiques? | 2 |
| 3.2 | Encodeur automatique le plus simple possible | 3 |
| 3.2.1 | Exercice | 4 |
| 3.3 | Un auto-encodeur profond | 5 |
| 3.3.1 | Code complet de l'auto-encodeur profond | 6 |
| 3.3.2 | Exercice | 7 |
| 3.4 | Auto-encodeur convolutif | 8 |
| 3.4.1 | Code complet de l'auto-encodeur convolutif | 10 |
| 3.4.2 | Exercice | 11 |
| 3.5 | Application au débruitage d'image | 12 |
| | Attention | 12 |
| 3.5.1 | Code complet de l'auto-encodeur convolutif pour le débruitage | 13 |
| 3.5.2 | Exercices | 14 |
| 3.5.3 | Sujet final - inférence - avec modèles entraînés sur Jetson Xavier | 15 |
| 3.5.3.1 | Voici le code | 15 |
| 3.5.3.2 | Un exemple d'exécution - entraînement avec les chiffres bruités | 17 |
| 3.5.3.3 | Exercices | 18 |
| 3.6 | Remarque finale | 18 |