

$$G_{kk'}^{[l]} = \sum_{i=1}^{n_H^{[l]}} \sum_{j=1}^{n_w^{[l]}} a_{ijk}^{[l]} a_{ijk'}^{[l]}$$

Remarque : les matrices de style de l'image de style et de l'image générée sont notées $G^{[l](S)}$ et $G^{[l](G)}$ respectivement.

□ **Fonction de coût de style** – La fonction de coût de style (en anglais *style cost function*), notée $J_{\text{style}}(S, G)$, est utilisée pour quantifier à quel point l'image générée G diffère de l'image de style S . Elle est définie de la manière suivante :

$$J_{\text{style}}^{[l]}(S, G) = \frac{1}{(2n_H n_w n_c)^2} \|G^{[l](S)} - G^{[l](G)}\|_F^2 = \frac{1}{(2n_H n_w n_c)^2} \sum_{k, k'=1}^{n_c} \left(G_{kk'}^{[l](S)} - G_{kk'}^{[l](G)} \right)^2$$

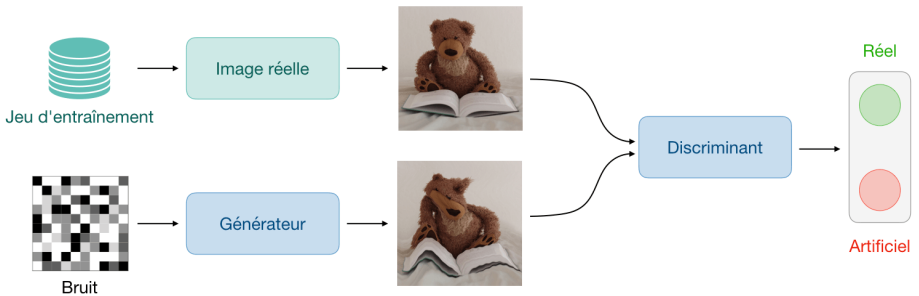
□ **Fonction de coût total** – La fonction de coût total (en anglais *overall cost function*) est définie comme étant une combinaison linéaire des fonctions de coût de contenu et de style, pondérées par les paramètres α, β , de la manière suivante :

$$J(G) = \alpha J_{\text{content}}(C, G) + \beta J_{\text{style}}(S, G)$$

Remarque : plus α est grand, plus le modèle privilégiera le contenu et plus β est grand, plus le modèle sera fidèle au style.

1.6.3 Architectures utilisant des astuces de calcul

□ **Réseau antagoniste génératif** – Les réseaux antagonistes génératifs (en anglais *generative adversarial networks*), aussi connus sous le nom de GANs, sont composés d'un modèle génératif et d'un modèle discriminatif, où le modèle génératif a pour but de générer des prédictions aussi réalistes que possibles, qui seront ensuite envoyées dans un modèle discriminatif qui aura pour but de différencier une image générée d'une image réelle.



Remarque : les GANs sont utilisées dans des applications pouvant aller de la génération de musique au traitement de texte vers image.

□ **ResNet** – L'architecture du réseau résiduel (en anglais *Residual Network*), aussi appelé ResNet, utilise des blocs résiduels avec un nombre élevé de couches et a pour but de réduire l'erreur de training. Le bloc résiduel est caractérisé par l'équation suivante :

$$a^{[l+2]} = g(a^{[l]} + z^{[l+2]})$$