

LRP für CNNs

- Basisformel

$$R_i^{(l)} = \sum_j \frac{z_{ij}}{\sum_{i'} z_{i'j}} R_j^{(l+1)} \quad \text{mit} \quad z_{ij} = x_i^{(1)} w_{ij}^{(l,l+1)}$$

- Problematisch bei komplexeren Layern (*Beispiel: ConvLayer*)
- Betrachte alternative Implementierung

LRP für CNNs

■ Basisformel

$$\begin{aligned}\forall_k : z_k &= \epsilon + \sum_{0,j} a_j \cdot \rho(w_{jk}) && \text{(forward pass)} \\ \forall_k : s_k &= R_k / z_k && \text{(element-wise division)} \\ \forall_j : c_j &= \sum_k \rho(w_{jk}) \cdot s_k && \text{(backward pass)} \\ \forall_j : R_j &= a_j c_j && \text{(element-wise product)}\end{aligned}$$

- Problematisch bei komplexeren Layern (*Beispiel: ConvLayer*)
- Betrachte alternative Implementierung