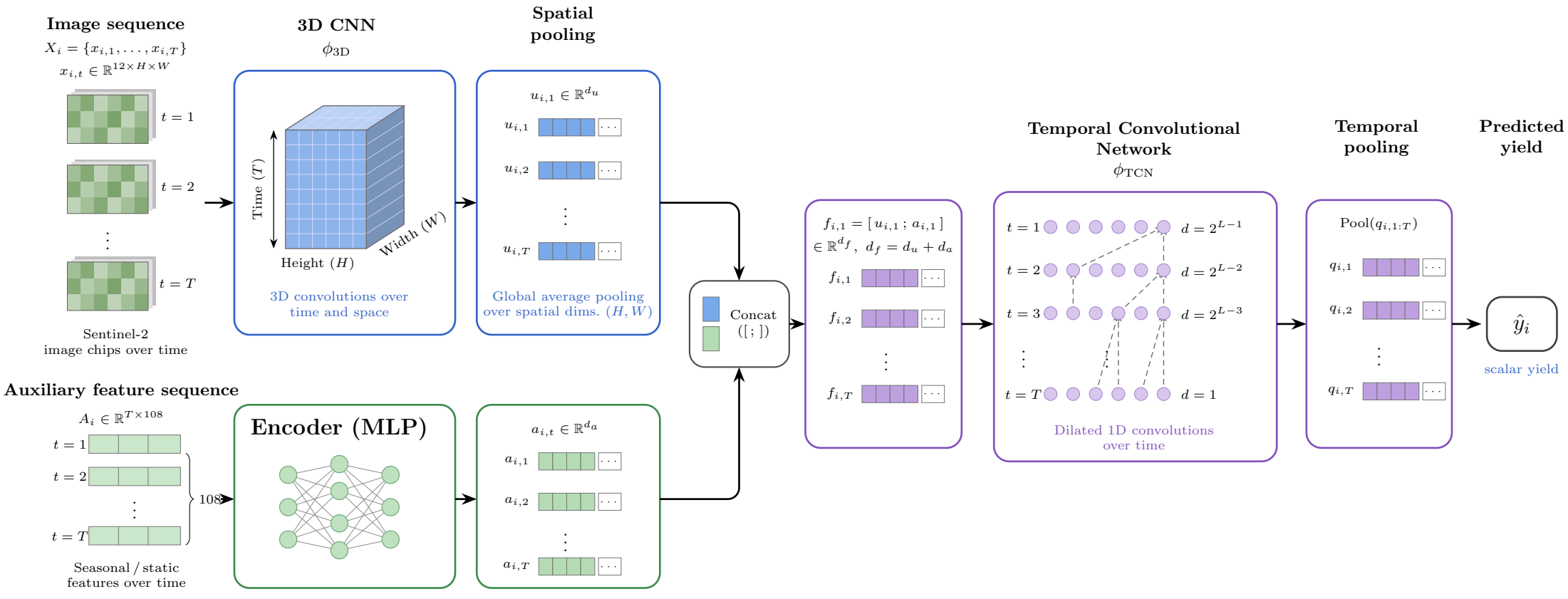
Image sequence
$X_i = \{x_{i,1}, \ldots, x_{i,T}\}$
$x_{i,t} \in \mathbb{R}^{12 \times H \times W}$
$t = 1$
$t = 2$
$t = T$
Sentinel-2
image chips over time
3D CNN
ϕ_{3D}
Time (T)
Height (H)
Width (W)
3D convolutions over
time and space
Spatial
pooling
$u_{i,1} \in \mathbb{R}^{d_u}$
$u_{i,1}$
$u_{i,2}$
$u_{i,T}$
Global average pooling
over spatial dims. (H, W)
Auxiliary feature sequence
$A_i \in \mathbb{R}^{T \times 108}$
$t = 1$
$t = 2$
$t = T$
108
Seasonal / static
features over time
Encoder (MLP)
$a_{i,t} \in \mathbb{R}^{d_a}$
$a_{i,1}$
$a_{i,2}$
$a_{i,T}$
Concat
([;])
$f_{i,1} = [\, u_{i,1} \,;\, a_{i,1} \,]$
$\in \mathbb{R}^{d_f}, \; d_f = d_u + d_a$
$f_{i,1}$
$f_{i,2}$
$f_{i,T}$
Temporal Convolutional
Network
ϕ_{TCN}
$t = 1$
$t = 2$
$t = 3$
$t = T$
$d = 2^{L-1}$
$d = 2^{L-2}$
$d = 2^{L-3}$
$d = 1$
Dilated 1D convolutions
over time
Temporal
pooling
$\text{Pool}(q_{i,1:T})$
$q_{i,1}$
$q_{i,2}$
$q_{i,T}$
Predicted
yield
$\hat{y}_i$
scalar yield