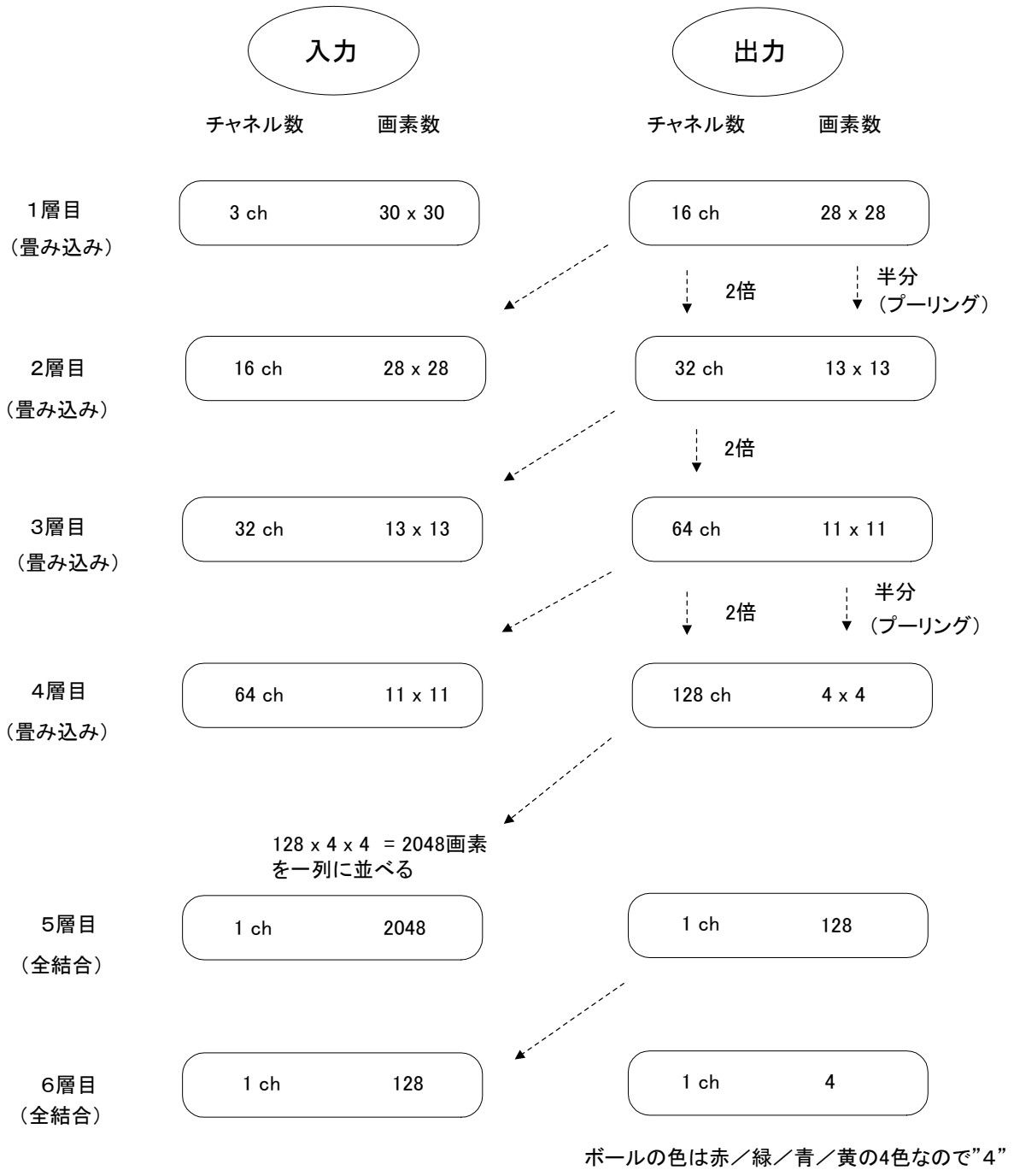


図1 画像認識の流れ～各層のチャネル数と画素数

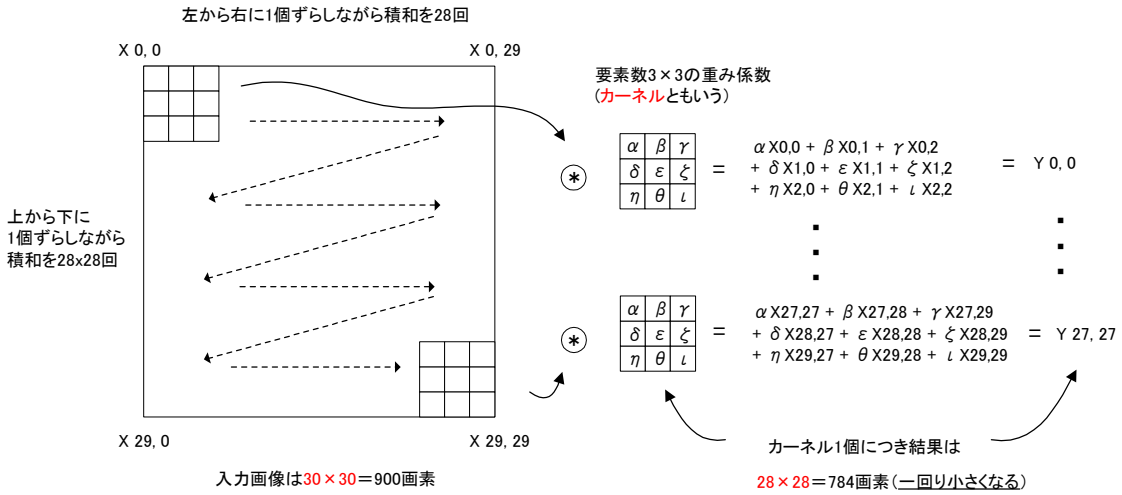


図2　この作業を「畳み込み」という

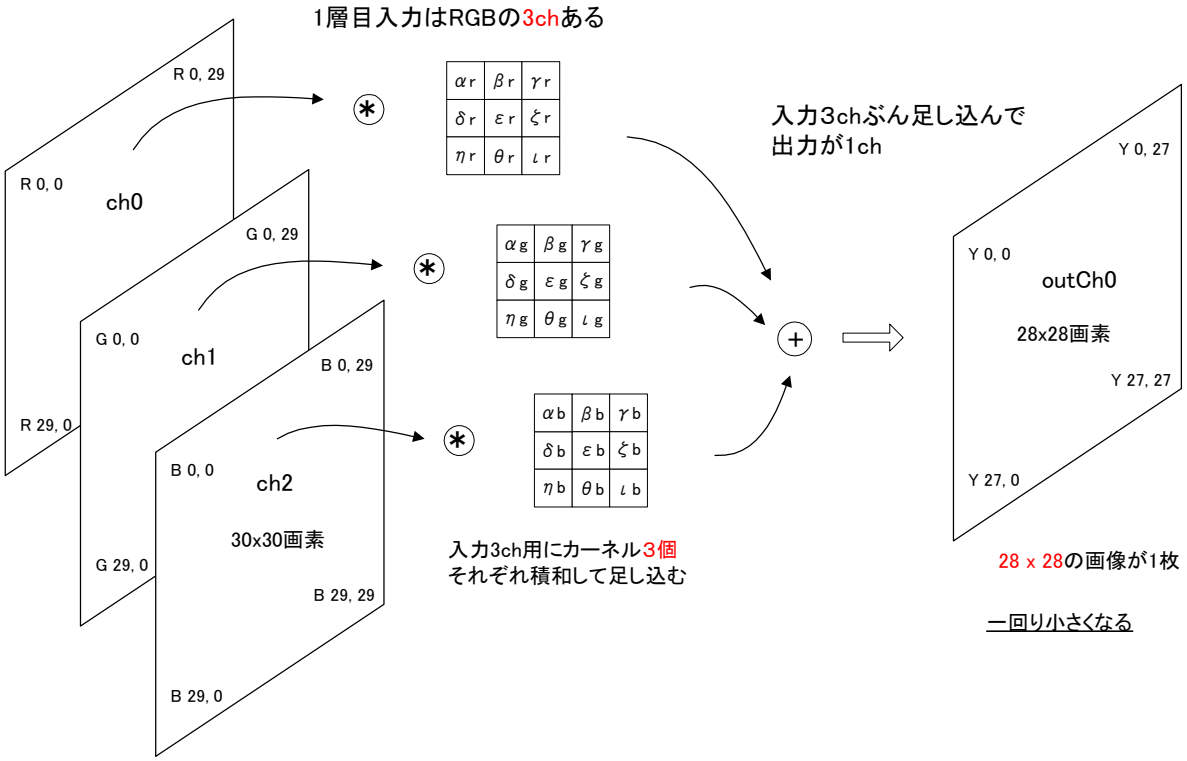


図3　入力3chをそれぞれ積和して足し込む

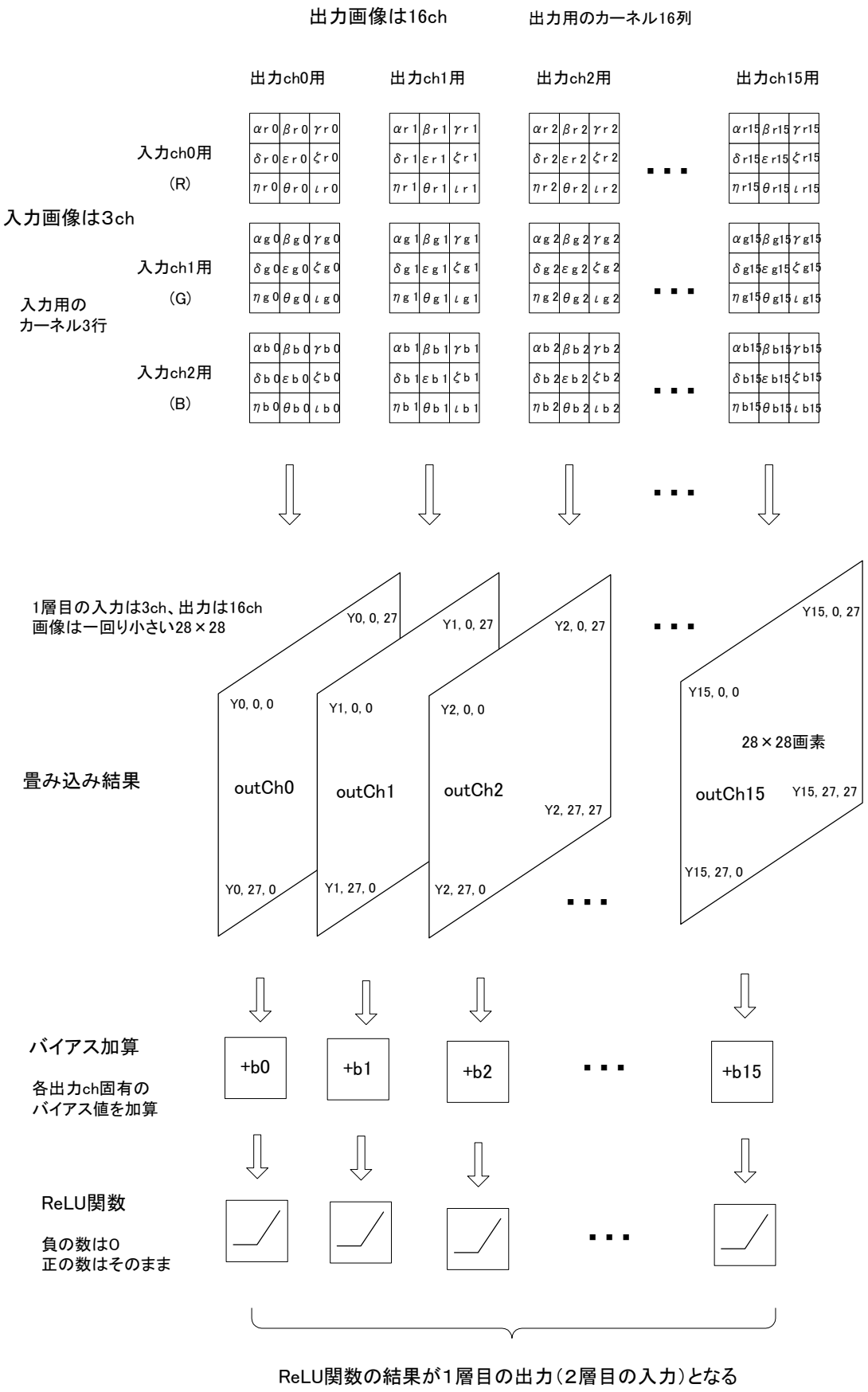


図4　入力にかかるカーネル3行、出力を計算するためのカーネル16列

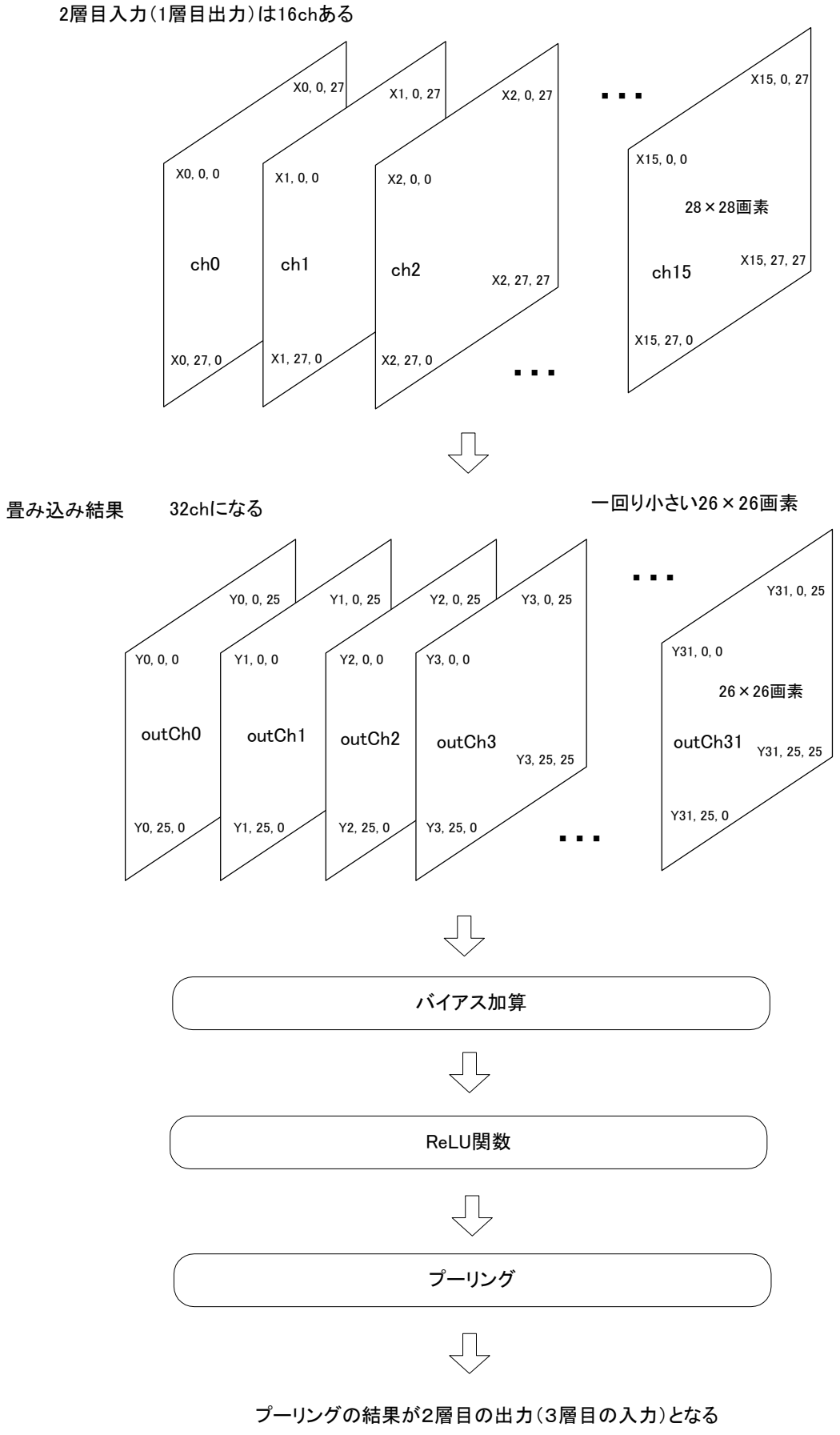
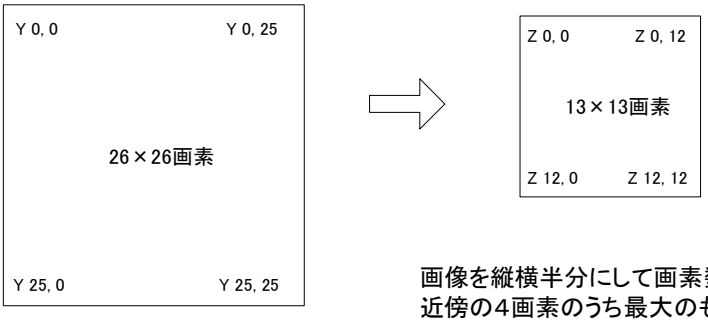


図5　畳み込み→バイアス加算→ReLUの後、2層目にはプーリングがある



画像を縦横半分にして画素数を1/4にする
近傍の4画素のうち最大のものを残す
<u>この作業をプーリングという</u>

図6 プーリングで26×26→13×13画素になる

2層目の出力（3層目の入力）はプーリングにより13×13になる

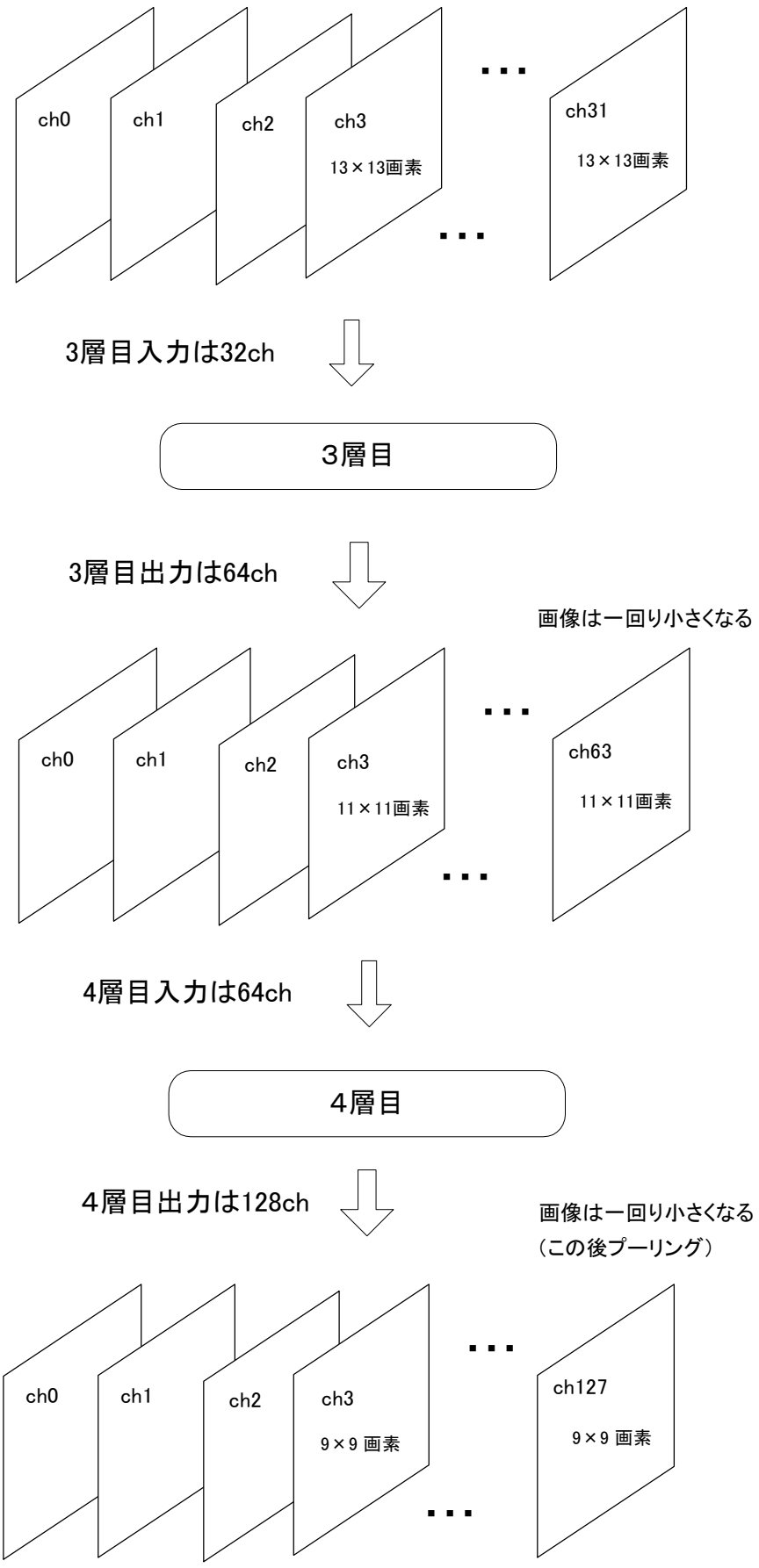


図7　3層目と4層目。この後プーリングがある

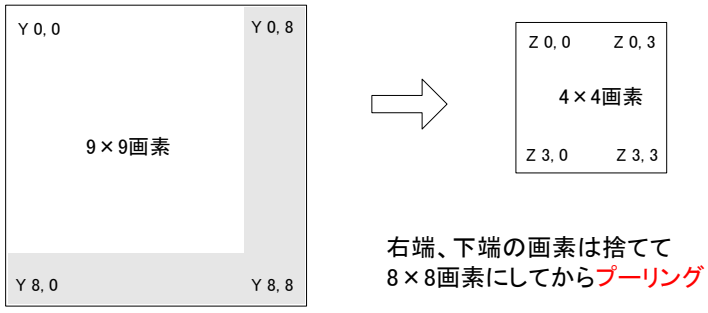
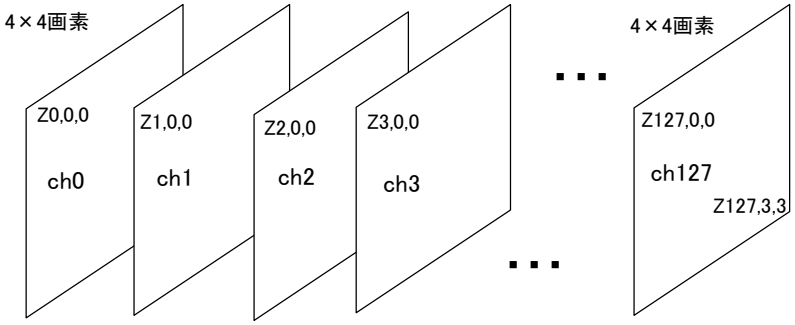
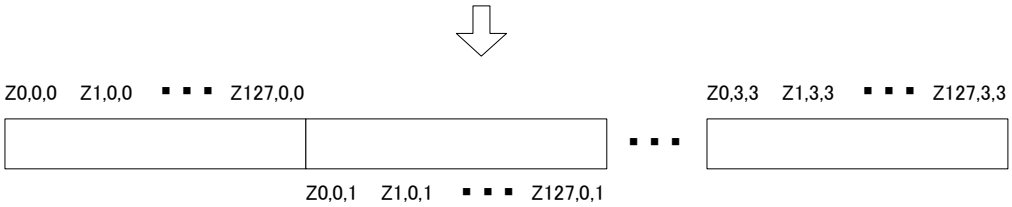


図8　4層目のプーリング。右と下の画素は捨てるので注意



4層目出力は4x4画素が128ch
これを一行2048画素に並び替える



128 x 4 x 4 = 2048 画素になる

図9　並べ方に注意　　Z c, x, y　⇒　c：チャネル番号、x, y：縦横の座標

図10　5層目は係数の数が多いので注意（コラムB）

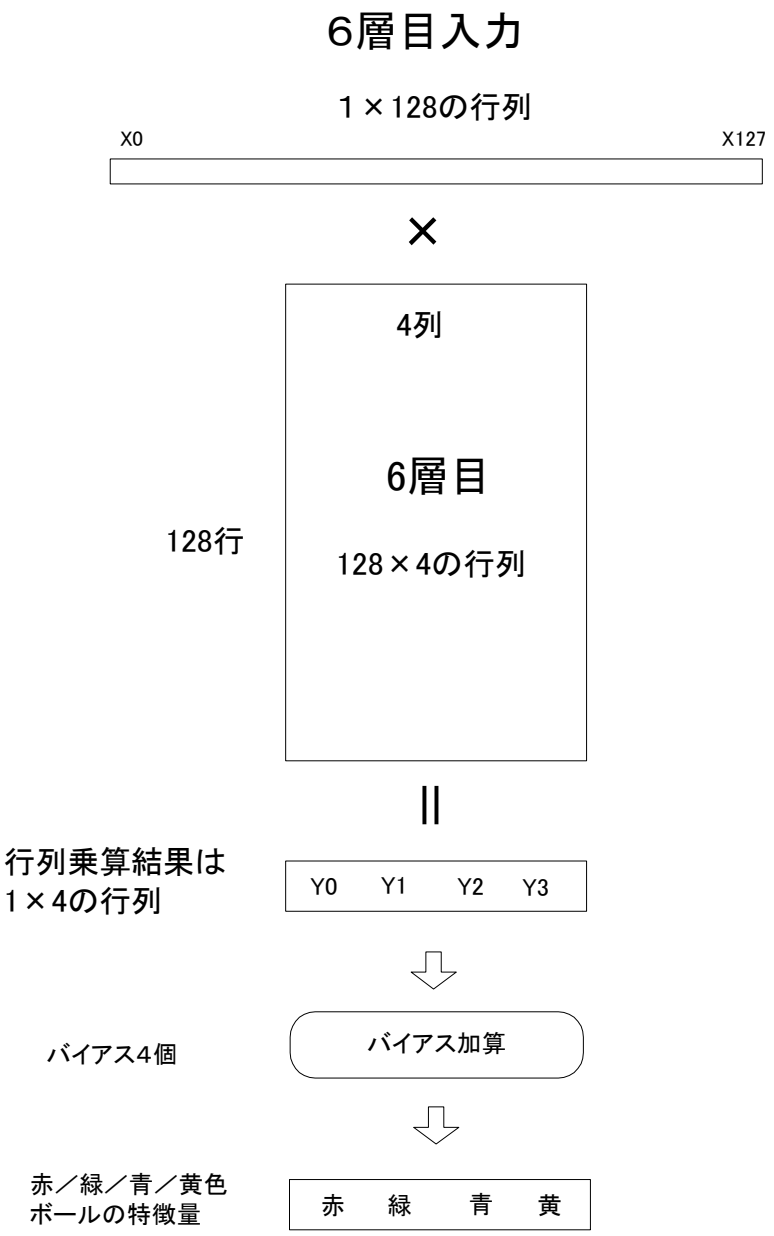
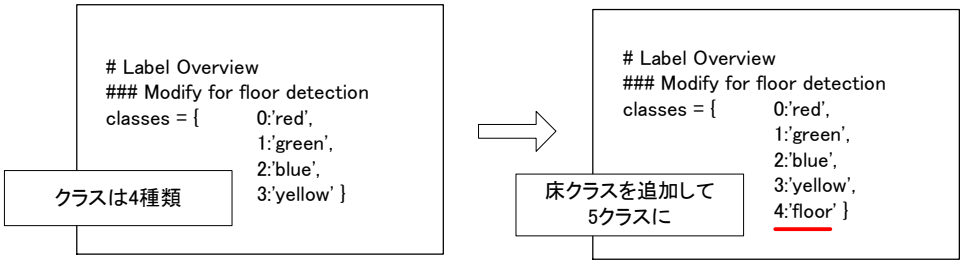
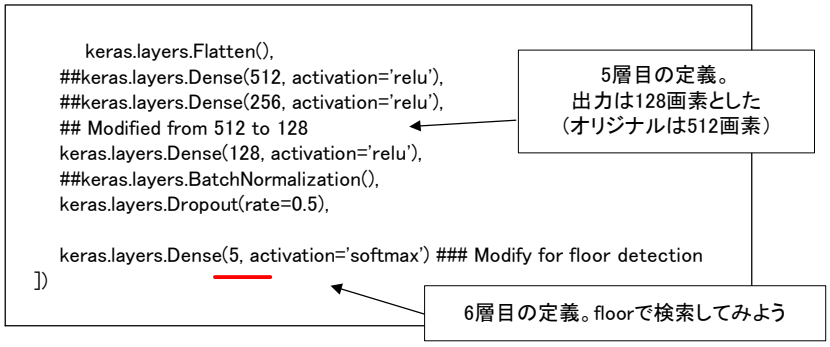


図11　最終的に4個の値が残る



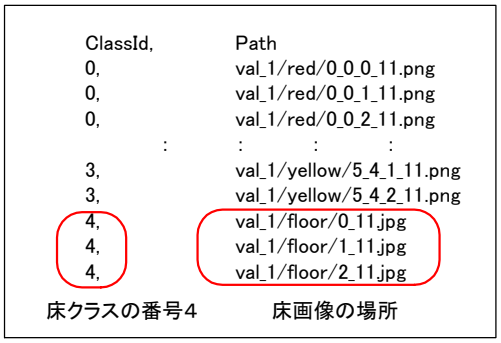
(a) balls_float.py　　(b) balls_float_5.py

リスト1　床クラスのラベル"floor"を追加する



リスト2　CNNのモデル定義の部分。6層目の出力画素数を"4"→"5"に変更（balls_float.py→balls_float_5.py）

リスト3　ラベルリストをTestBalls.csvからTestBalls_5.csvに変更（balls_float_pred.py→balls_float_pred_5.py）



リスト4　床画像のclassIDとPathをTestBalls_5.csvに追加